



Mateo Andrés Manosalva Amaris

Edgar Santiago Ochoa Quiroga

Sergio Alejandro Bello Torres .....

## Problemas

### Problema1

Un bloque de masa  $m$  desliza bajo la acción de la gravedad por un plano inclinado formando un ángulo  $\theta$  respecto de la horizontal. Se puede demostrar que, si la fuerza de rozamiento  $F_r$  entre el bloque y el plano viene dada por  $F_r = -kmv^2$  con  $v$  la velocidad del bloque y  $k$  un coeficiente de rozamiento, entonces el tiempo  $T$  requerido para que el bloque recorra una distancia  $D$  partiendo del reposo está relacionado de la forma

$$e^{kD} = \cosh \left( T \sqrt{kg \sin \theta} \right),$$

siendo  $g = 9,8$  la aceleración de la gravedad. Si  $\theta = \frac{\pi}{4}$ ,  $k = 0,5 \pm 0,1$  y  $T = 2 \pm 0,2$ , hallar la precisión con la que se conoce  $D$ .

**Solución.** Despejando  $D$  de la ecuación, se tiene que:

$$D = \frac{\ln \left( \cosh \left( T \sqrt{kg \sin \theta} \right) \right)}{k}$$

Tomaremos  $D = \varphi(T, k)$ , de tal manera que el error relativo venga dado por:

$$\frac{\phi(\hat{T}, \hat{k}) - \phi(T, k)}{\phi(T, k)} \approx \Delta T \frac{\phi_T(T, k)}{\phi(T, k)} + \Delta k \frac{\phi_k(T, k)}{\phi(T, k)}$$

Calculando  $\phi_T$  y  $\phi_k$  obtenemos:

$$\phi_T(T, k) = \sqrt{\frac{g \sin \theta}{k}} \tanh \left( T \sqrt{kg \sin \theta} \right)$$

y

$$\phi_k(T, k) = \frac{T \sqrt{kg \sin \theta} \tanh \left( T \sqrt{kg \sin \theta} \right) - 2 \ln \left( \cosh \left( T \sqrt{kg \sin \theta} \right) \right)}{2k^2}$$

Con lo cual la estimación del error nos queda:

$$\frac{\phi(\hat{T}, \hat{k}) - \phi(T, k)}{\phi(T, k)} \approx \Delta T \frac{\sqrt{kg} \sin \theta \tanh(T\sqrt{kg} \sin \theta)}{\ln(\cosh(T\sqrt{kg} \sin \theta))} + \Delta k \left( \frac{T\sqrt{kg} \sin \theta \tanh(T\sqrt{kg} \sin \theta)}{2k \ln(\cosh(T\sqrt{kg} \sin \theta))} - \frac{1}{k} \right)$$

De esta forma, tomando  $k = 0,5$ ,  $T = 2$ ,  $g = 9,8$ ,  $\Delta T = 0,2$ ,  $\Delta k = 0,1$  y sabiendo que  $\sin\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}$ , al calcular  $\phi(T, k)$  y estimar el error relativo, obtenemos:

$$\phi(T, k) \approx 6,0605$$

y el error relativo es

$$\frac{\phi(\hat{T}, \hat{k}) - \phi(T, k)}{\phi(T, k)} \approx 0,2 \times \frac{1,8592}{3,0302} + 0,1 \times \left( \frac{3,7185}{3,0302} - 2 \right) \approx 0,0454$$

Así, obtenemos un error relativo de  $4,54 \times 10^{-2} \leq 5 \times 10^{-2}$ , lo cual significa que solo conocemos  $D$  con una precisión de una cifra significativa.

## Problema2

Utilizando el método de redondeo.

(a) Hallar el número de máquina más próximo a 125,6 y a 126 si trabaja con

- Base 10 y mantisa de 2 dígitos.

**Solución.** Para usar el método de redondeo, verificamos que para truncar 125,6 a 2 dígitos, el tercer dígito es  $5 \geq \frac{10}{2}$ , por lo tanto obtenemos 130 y la representación de 125,6 en punto flotante es:

$$\text{fl}(125,6) = 1,3 \times 10^2$$

De manera análoga, para 126, vemos que  $6 \geq \frac{10}{2}$ , así:

$$\text{fl}(126) = 1,3 \times 10^2$$

De donde concluimos que 130 es el número de máquina más próximo a 125,6 y 126.

- Base 2 y mantisa de 8 dígitos.

**Solución.** Primero, convertimos 125,6 y 126 a base 2; Para convertir 125 a base 2 tenemos:

División	Cociente	Residuo
$125 \div 2$	62	1
$62 \div 2$	31	0
$31 \div 2$	15	1
$15 \div 2$	7	1
$7 \div 2$	3	1
$3 \div 2$	1	1
$1 \div 2$	0	1

Ahora, convertimos 0,6 a base 2:

Producto	Resultado	Parte entera
$0.6 \times 2$	1.2	1
$0.2 \times 2$	0.4	0
$0.4 \times 2$	0.8	0
$0.8 \times 2$	1.6	1
$0.6 \times 2$	1.2	1
$\vdots$	$\vdots$	$\vdots$

Así, 125,6 en binario es:

$$(125,6)_{10} = (1111101, \overline{10011})_2$$

Ya que en el sistema binario, el único dígito diferente de 0 es 1, podemos representar el número con 9 dígitos, de esta forma, observamos que el décimo dígito de esta representación es 0, por lo tanto truncamos a 9 dígitos y así obtenemos

$$\text{fl}(125,6) = 1,11110110_2 \times 2^6$$

Es decir, el número de máquina más próximo a 125,6 en base 2 y mantisa de 8 dígitos es

$$\left(1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \frac{0}{2^5} + \frac{1}{2^6} + \frac{1}{2^7} + \frac{0}{2^8}\right) \times 2^6 = 125,5$$

Ahora convertimos 126 a base 2:

División	Cociente	Residuo
$126 \div 2$	63	0
$63 \div 2$	31	1
$31 \div 2$	15	1
$15 \div 2$	7	1
$7 \div 2$	3	1
$3 \div 2$	1	1
$1 \div 2$	0	1

Por lo tanto

$$(126)_{10} = (1111110)_2$$

Este numero es de 7 dígitos, por lo tanto se puede representar exactamente en punto flotante:

$$\text{fl}(126) = 1,1111000_2 \times 2^6$$

luego el número de maquina más próximo a 126 es 126.

(b) Verificar para  $x = 125,6$  la cota para el error relativo

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

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

**Solución.** Para la representación en base 10 y mantisa de 2 dígitos, el error relativo es:

$$\left| \frac{125,6 - 130}{125,6} \right| = 0,035 < \frac{1}{2} \times 10^{1-2} = 0,05$$

De manera similar, calculamos el error para base 2 y mantisa de 8 dígitos

$$\left| \frac{125,6 - 125,5}{125,6} \right| \approx 0,000796 < \frac{1}{2} \times 2^{1-8} = 2^{-8} = 0,00390625$$

En ambos casos, verificamos que se cumple la cota del error.

(c) ¿Cuál es, en cada caso, el valor que da la máquina como resultado de las operaciones  $126 + 126,5$  y  $126 - 125,6$ ? ¿Cuál es el error relativo de estos resultados?

**Solución.** • Para base 10 y mantisa de 2 dígitos se tiene:

$$\begin{aligned} \text{fl}(126 + 126,5) &= \text{fl}(\text{fl}(126) + \text{fl}(126,5)) \\ &= \text{fl}(1,3 \times 10^2 + 1,3 \times 10^2) \\ &= \text{fl}(2,6 \times 10^2) \\ &= 2,6 \times 10^2 \\ &= 260 \end{aligned}$$

Sin embargo  $126 + 126,5 = 252,5$ , por lo tanto el error relativo de este resultado es:

$$\left| \frac{252,5 - 260}{252,5} \right| \approx 0,0297$$

Calculemos ahora  $\text{fl}(126 - 125, 6)$

$$\begin{aligned}\text{fl}(126 - 125, 6) &= \text{fl}(\text{fl}(126) - \text{fl}(125, 6)) \\ &= \text{fl}(1, 3 \times 10^2 - 1, 3 \times 10^2) \\ &= \text{fl}(0) \\ &= 0\end{aligned}$$

El error relativo es, entonces:

$$\left| \frac{0,4 - 0}{0,4} \right| = 1$$

Aquí podemos ver que debido a que el sistema de punto flotante representa dos números distintos con el mismo número, aunado a la pequeña cantidad de dígitos de la mantisa, el error llega a propagarse de tal manera que al restar dos números cercanos se obtiene un error relativo del 100 %, lo cual significa que en algunos casos, no podemos conocer ninguna cifra significativa al efectuar la resta de dos números de máquina.

- Base 2 y mantisa de 8 dígitos.

Para convertir 126,5 a binario usaremos que  $(126)_{10} = (1111110)_2$  y  $(0,5)_{10} = (0,1)_2$ , de esta manera  $(126,5)_{10} = (1111110,1)_2$  y su representación en punto flotante es  $\text{fl}(126,5) = 1,11111010_2 \times 2^6$ . Así, al efectuar  $126+126,5$  tenemos:

$$\begin{aligned}\text{fl}(126 + 126,5) &= \text{fl}(\text{fl}(126) + \text{fl}(126,5)) \\ &= \text{fl}(1,11111000_2 \times 2^6 + 1,1111101_2 \times 2^6) \\ &= \text{fl}(11,1111001_2 \times 2^6) \\ &= 1,11111001_2 \times 2^7\end{aligned}$$

Es decir:

$$\left(1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \frac{1}{2^5} + \frac{0}{2^6} + \frac{0}{2^7} + \frac{1}{2^8}\right) \times 2^7 = 252,5$$

Hallamos entonces el error relativo

$$\left| \frac{252,5 - 252,5}{252,2} \right| = 0$$

Calculemos ahora  $126-125,6$ :

$$\begin{aligned}\text{fl}(126 - 125,6) &= \text{fl}(\text{fl}(126) - \text{fl}(125,6)) \\ &= \text{fl}(1,11111000_2 \cdot 2^6 - 1,11110110_2 \cdot 2^6) \\ &= 1,0000000_2 \cdot 2^{-1}\end{aligned}$$

Es decir  $\text{fl}(126 - 125,6) = 0,5$ , luego el error relativo es:

$$\left| \frac{0,6 - 0,5}{0,6} \right| = 0,1\bar{6}$$

### Problema3

Suponga que una compañía de computadores está desarrollando un nuevo sistema de punto flotante para usarlo con sus máquinas. Ellos necesitan su ayuda para responder unas preguntas acerca de su sistema. Siguiendo la terminología descrita en clase, el sistema de punto flotante de la compañía se especifica por  $(\beta, t, L, U)$ . Usted debe suponer que:

- Todos los valores de punto flotante son normalizados (excepto la representación de punto flotante de cero).
- Todos los dígitos en la mantisa de un valor de punto flotante son almacenados explícitamente.
- El cero se representa con una mantisa y exponente de ceros.

Preguntas:

- (a) ¿Cuántos valores diferentes de punto flotante no negativos pueden representarse por medio de este sistema de punto flotante?

**Solución.** Recordemos que en general la representación de punto flotante esta dada por

$$\text{fl}(x) = \pm \left( d_0 + \frac{d_1}{\beta} + \frac{d_2}{\beta^2} + \cdots + \frac{d_{t-1}}{\beta^{t-1}} \right) \beta^e = \pm (d_0.d_1d_2 \dots d_{t-1})\beta^e$$

Como son los no negativos y los valores del punto flotante están normalizados, Procedamos primero con los positivos, como todos los valores están normalizados, el primer dígito de la representación tiene que ser diferente de 0, es decir,  $1 \leq d_0 \leq \beta - 1$ , luego para  $d_i$  con  $1 \leq i \leq t - 1$ , como estos no tienen restricción podemos tomarlos tal que  $0 \leq d_i \leq \beta - 1$ . Por lo que de momento, por el principio fundamental del conteo tenemos  $(\beta - 1)(\beta)^{t-1}$  valores diferentes. Note que decimos de momento ya que no hemos considerado los posibles exponentes. Ahora tomando  $e$  un posible exponente, sabemos que  $L \leq e \leq U$ , luego la cantidad de posibles exponentes para cada uno de los valores encontrados seria  $U - L + 1$ . Así concluimos que la cantidad de valores no negativos diferentes en este sistema de punto flotante es

$$(\beta - 1)(\beta)^{t-1}(U - L + 1) + 1 = .$$

Note que ese mas 1 del final es añadido para contar el caso particular del 0, que es el único que tiene una mantisa y exponente de ceros.

- (b) La misma pregunta para el caso  $(\beta, t, L, U) = (8, 5, -100, 100)$ , el cual la compañía está contemplando en particular.

**Solución.** Como hallamos en el numeral anterior una formula para un sistema de punto flotante arbitrario, aquí basta simplemente con reemplazar los respectivos valores, es decir

$$\begin{aligned} (\beta - 1)(\beta)^{t-1}(U - L + 1) + 1 &= (8 - 1)8^{(5-1)}(100 - (-100) + 1) + 1 \\ &= 7 \cdot 8^4 \cdot 201 + 1 \\ &= 5763073. \end{aligned}$$

- (c) ¿Cuál es el valor aproximado (en base 10) del número más grande y el número positivo más pequeño que pueden ser representado en este sistema de punto flotante?

**Solución.** • Primero responderemos la pregunta de cual es el numero mas grande. Tenemos que tal numero seria uno donde cada  $d_i = \beta - 1$  y nuestro  $e = U$ , así para el sistema particular tenemos que ese numero es

$$x = (7,7777)8^{100}.$$

Ahora si lo vemos en representación decimal tenemos que

$$x = \left(7 + \frac{7}{8} + \frac{7}{8^2} + \frac{7}{8^3} + \frac{7}{8^4}\right) 8^{100} \approx 1,62957 \times 10^{91}.$$

- Ahora el numero mas pequeño positivo es cuando tenemos que  $d_0 = 1$ , mientras que  $d_i = 0$  para los demás, esto ya que esta normalizado, y consideramos  $e = L$ . Para nuestro caso particular tenemos que

$$y = (1,0000)8^{-100} = \left(1 + \frac{0}{8} + \frac{0}{8^2} + \frac{0}{8^3} + \frac{0}{8^4}\right) 8^{-100} \approx 4,90909 \times 10^{-91}.$$

- (d) Para el caso general, demuestre que los errores relativos producidos por realizar truncamiento y redondeo son

$$\left| \frac{fl(x) - x}{x} \right| \leq \begin{cases} \beta^{1-t} & \text{cuando se efectúa truncamiento,} \\ \frac{1}{2}\beta^{1-t} & \text{cuando se efectúa redondeo.} \end{cases}$$

**Demostración.** La prueba la tenemos que dividir en el caso del truncamiento y el caso del redondeo. Primero consideremos el caso del truncamiento. Tenemos que la expansión decimal de un numero  $x > 0$ , en un sistema de punto flotante arbitrario es

$$x = \left(d_0 + \frac{d_1}{\beta} + \frac{d_2}{\beta^2} + \cdots + \frac{d_{t-1}}{\beta^{t-1}} + \frac{d_t}{\beta^t} + \frac{d_{t+1}}{\beta^{t+1}} + \cdots\right) \beta^e = (d_0.d_1d_2\dots d_{t-1}d_td_{t+1}\dots)\beta^e.$$

Luego si realizamos el truncamiento en el termino  $t - 1$ , es decir, hacemos 0 los  $d_i$  cuando  $i > t - 1$ . tenemos que

$$fl(x) = \left(d_0 + \frac{d_1}{\beta} + \frac{d_2}{\beta^2} + \cdots + \frac{d_{t-1}}{\beta^{t-1}}\right) \beta^e.$$

Ahora note que si consideramos el error absoluto, obtenemos que

$$\begin{aligned}
 |fl(x) - x| &= \left| \left( -\frac{d_t}{\beta^t} - \frac{d_{t+1}}{\beta^{t+1}} - \dots \right) \beta^e \right| \\
 &= \left( \frac{d_t}{\beta^t} + \frac{d_{t+1}}{\beta^{t+1}} + \dots \right) \beta^e \\
 &= \left( d_t + \frac{d_{t+1}}{\beta} + \dots \right) \frac{\beta^e}{\beta^t} \\
 &\leq \beta^{1+e-t}
 \end{aligned}$$

Note que como  $d_i \leq \beta$  para todo  $i$ , ese ultimo paso para acotar es valido. Ahora note que como  $x > 0$ , tenemos que  $|x| \geq d_0 \beta^e \geq \beta^e$ . En efecto, ya que como el sistema esta normalizado, necesariamente  $d_0 \geq 1$  y ademas el numero  $d_0$  seguido de ceros es siempre menor que  $d_0$  seguido de cualquier cadena de dígitos permitida. Así tenemos que  $\frac{1}{|x|} \leq \frac{1}{\beta^e}$ . Multiplicando en la anterior desigualdad obtenida tenemos que

$$\left| \frac{fl(x) - x}{x} \right| \leq \frac{\beta^{1+e-t}}{\beta^e} = \beta^{1-t}.$$

Ahora consideremos el caso donde redondeamos. Note que podemos asumir que  $d_t \geq \frac{\beta}{2}$ , ya que en caso contrario no hay redondeo y simplemente es un truncamiento. En este caso nuestra aproximación es

$$fl(x) = \left( d_0 + \frac{d_1}{\beta} + \frac{d_2}{\beta} + \dots + \frac{d_{t-1} + 1}{\beta^{t-1}} \right) \beta^e = (d_0.d_1d_2\dots(d_{t-1} + 1))\beta^e.$$

Luego si consideramos nuevamente el error absoluto, obtenemos

$$\begin{aligned}
 |fl(x) - x| &= |(0.\underbrace{0\dots0}_{t-1 \text{ veces}}(\beta - 1 - d_t)(\beta - 1 - d_{t+1})\dots)\beta^e| \\
 &= ((\beta - 1 - d_t).(\beta - 1 - d_{t+1})\dots)\beta^{e-t} \\
 &\leq \frac{1}{2}\beta\beta^{e-t} = \frac{\beta^{1+e-t}}{2}.
 \end{aligned}$$

En efecto, ya que como  $d_t \geq \frac{\beta}{2}$ , tenemos que  $\beta - 1 - d_t \leq \beta - 1 - \frac{\beta}{2} = \frac{\beta}{2} - 1 \leq \frac{\beta}{2}$ . Como el  $x$  sigue siendo el mismo del anterior caso, seguimos teniendo que  $\frac{1}{|x|} \leq \frac{1}{\beta^e}$ . Procediendo de la misma manera obtenemos

$$\left| \frac{fl(x) - x}{x} \right| \leq \frac{\beta^{1+e-t}}{2\beta^e} = \frac{1}{2}\beta^{1-t}.$$

Note que el caso cuando  $x = 0$  no es necesario revisarlo ya que la representación es idéntica, mientras que el caso  $x < 0$  como estamos trabajando dentro de un valor absoluto es aplicar el



mismo razonamiento. De esta manera hemos demostrado que

$$\left| \frac{fl(x) - x}{x} \right| \leq \begin{cases} \beta^{1-t} & \text{cuando se efectúa truncamiento,} \\ \frac{1}{2}\beta^{1-t} & \text{cuando se efectúa redondeo.} \end{cases}$$

□□

## Problema4

Considere la expansión de Taylor para la función exponencial

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{i=0}^{\infty} \frac{x^i}{i!} = \lim_{N \rightarrow \infty} S(x, N)$$

donde  $S(x, N)$  es la suma parcial con  $N + 1$  términos.

- (a) Escriba un programa que grafique el error relativo de la suma,  $\frac{|S(x, N) - e^x|}{e^x}$  versus  $N$  (hasta  $N = 60$ ) para un valor dado de  $x$ . Pruebe su programa para  $x = 10, 2, -2$  y  $-10$ . De las gráficas, explique por qué ésta no es una buena manera para evaluar  $e^x$  cuando  $x < 0$ .

**Solución.** Implementamos en Python un código que grafica el error relativo de la suma parcial de Taylor, obtuvimos el resultado

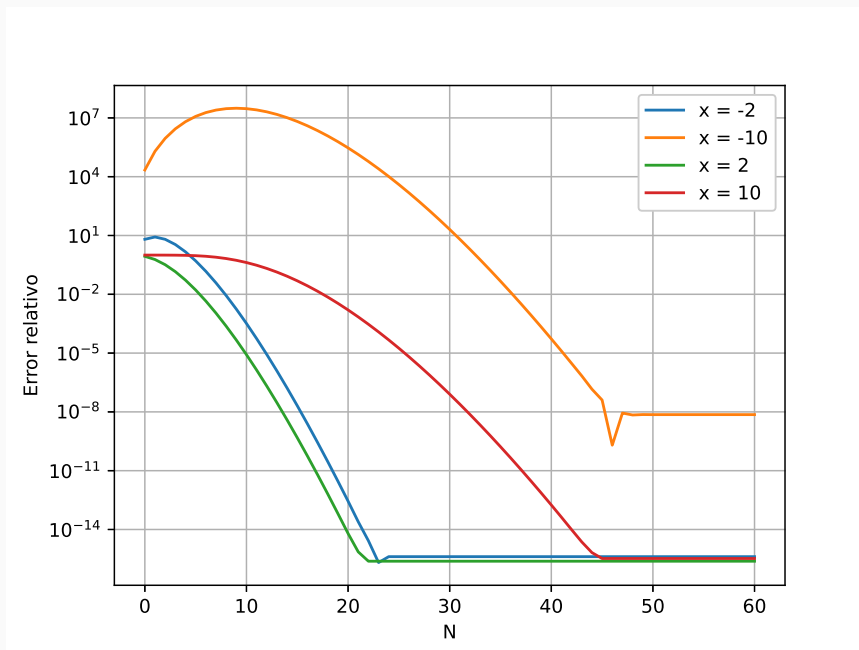


Figura 1: Error relativo

En la gráfica se evidencia que el error aumenta cuando los valores de  $x$  son negativos, note que si  $x$  es negativo entonces reescribimos la expresión como  $e^{-x}$  con  $x > 0$ , de esto se

sigue que

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \dots = \sum_{i=0}^{\infty} \frac{(-1)^i x^i}{i!} = \lim_{N \rightarrow \infty} \frac{1}{S(x, N)},$$

como al calcular el error necesitamos aproximar este término, entonces al truncar la serie alternante obtenemos resultados que oscilan alrededor del 0, por lo que la expresión  $1/S(x, N)$  toma valores grandes, en cambio si  $x$  es positivo nuestra serie truncada se aleja cada vez más de 0.

- (b) Modifique su programa tal que use la identidad  $e^x = \frac{1}{e^{-x}} = \frac{1}{S(-x, \infty)}$  para evaluar la función exponencial cuando  $x$  es negativa. Explique por qué esta técnica funciona mejor.

**Solución.** Para este caso obtuvimos la gráfica del error

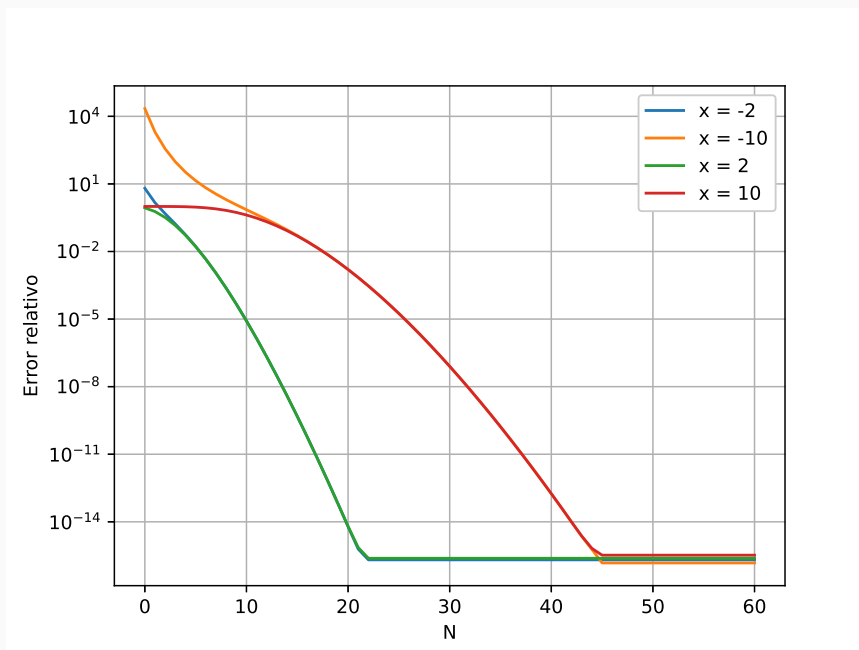


Figura 2: Error relativo

Esta técnica funciona mejor ya que si  $x < 0$ , entonces  $-x > 0$  y por lo tanto al calcular

$$e^x = \frac{1}{e^{-x}} \quad x < 0,$$

el término  $-x$  es positivo y podemos aproximar de manera adecuada  $e^{-x}$  por  $S(-x, N)$  como consecuencia del numeral anterior ya la serie no es alternante.

## Problema5

- (a) De manera similar a la desarrollada en clase, deduzca que una aproximación de  $f'(x_0)$  es

$$\frac{f(x_0 + h) - f(x_0 - h)}{2h}.$$

Muestre que esta aproximación tiene un error de  $O(h^2)$ . Más precisamente, el primer término del error es  $-\frac{h^2}{6}f'''(x_0)$  cuando  $f'''(x_0) \neq 0$ .

**Demostración.** Por el teorema de Taylor tenemos que

$$f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2}f''(x_0) + \frac{h^3}{6}f'''(x_0) + \dots,$$

además

$$f(x_0 - h) = f(x_0) - hf'(x_0) + \frac{h^2}{2}f''(x_0) - \frac{h^3}{6}f'''(x_0) + \dots,$$

luego el error de aproximación de  $f'(x_0)$  por la expresión planteada no considera los términos pares de la expansión de Taylor ya que tienen el mismo signo y se cancelan al realizar la resta como sigue

$$f(x_0 + h) - f(x_0 - h) = 2hf'(x_0) + \frac{h^3}{6}f'''(x_0) + \frac{h^3}{6}f'''(x_0) + O(h^5),$$

así

$$\begin{aligned}\frac{f(x_0 + h) - f(x_0 - h)}{2h} &= f'(x_0) + \frac{2h^3}{2h6}f'''(x_0) + \dots \\ &= f'(x_0) + \frac{h^2}{6}f'''(x_0) + \dots,\end{aligned}$$

se sigue que

$$\left| \frac{f(x_0 + h) - f(x_0 - h)}{2h} - f'(x_0) \right| \leq \frac{h^2}{6}f'''(x_0)$$

□

- (b) Adapte el programa de Matlab hecho en clase para visualizar el comportamiento del error de aproximación a medida que el paso  $h$  decrece desde  $h = 10^{-1}, \dots, 10^{-16}$ .

**Solución.** Implementamos el siguiente código:

```

1 x0 = 1.2;
2 f0 = sin(x0);
3 fp = cos(x0);
4 i = -16:0.5:-1;
5 h = 10.^i;
6 err_centrada = abs(fp - (sin(x0 + h) - sin(x0 - h)) ./ (2 * h));
7 d_err_centrada = fp / 6 * h.^2;
8 loglog(h, d_err_centrada, 'r-.', h, err_centrada, 'b-*', '
    LineWidth', 2);
9 xlabel('h', 'fontsize', 16);
10 ylabel('Error Absoluto', 'fontsize', 14);
11 legend('Error de truncamiento', 'Diferencias finitas', 'location',
    'southwest');
12 grid on;

```

con lo que obtuvimos la gráfica

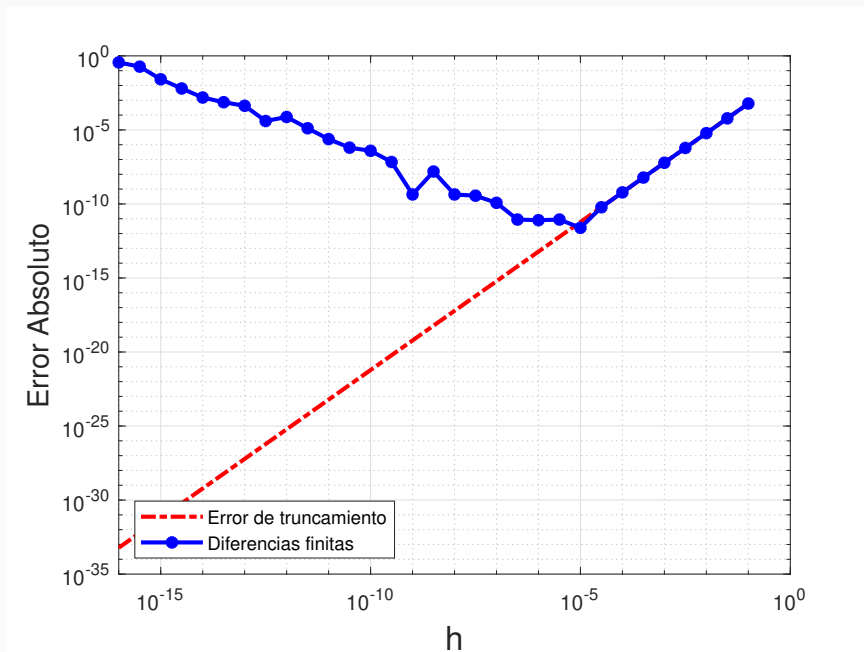


Figura 3: P5-1

(c) Muestre que

$$f'(x_0) = \frac{-f(x_0 + 2h) + 8f(x_0 + h) - 8f(x_0 - h) + f(x_0 - 2h)}{12h} + O(h^4)$$

De nuevo adapte su programa para visualizar el comportamiento del error de aproximación a medida que el paso  $h$  decrece desde  $h = 10^{-1}$ , ...,  $10^{-16}$ .

**Demostración.** De manera análoga al numeral a) tenemos que

$$8f(x_0 + h) - 8f(x_0 - h) = 16hf'(x_0) + 16\frac{h^3}{6}f'''(x_0) + \frac{8h^5}{60}f^{(5)}(x_0) + O(h^6),$$

por otro lado

$$-f(x_0 + 2h) = -f(x_0) - 2hf'(x_0) - \frac{4h^2}{2}f''(x_0) - \frac{8h^3}{6}f'''(x_0) - \frac{16h^4}{24}f^{(4)}(x_0) - \frac{32h^5}{120}f^{(5)}(x_0) + O(h^6),$$

$$f(x_0 - 2h) = f(x_0) - 2hf'(x_0) + \frac{4h^2}{2}f''(x_0) - \frac{8h^3}{6}f'''(x_0) + \frac{16h^4}{24}f^{(4)}(x_0) - \frac{32h^5}{120}f^{(5)}(x_0) + O(h^6),$$

sumando estas expresiones obtenemos que

$$8f(x_0 + h) - 8f(x_0 - h) - f(x_0 + 2h) + f(x_0 - 2h) = 12hf'(x_0) - \frac{2h^5}{5}f^{(5)}(x_0) + O(h^6).$$

En efecto

$$\frac{-f(x_0 + 2h) + 8f(x_0 + h) - 8f(x_0 - h) + f(x_0 - 2h)}{12h} = \frac{12hf'(x_0) + O(h^5)}{12h} = f'(x_0) + O(h^4),$$

de hecho no era necesario controlar el error de la quinta derivada para concluir este primer numeral, pero lo hicimos para poder adaptar el programa y visualizar de manera más precisa el error de truncamiento, obtenemos que el primer término de error es

$$\frac{h^4}{30}f^{(5)}(x_0)$$

□

Para este caso obtuvimos la siguiente gráfica del comportamiento del error

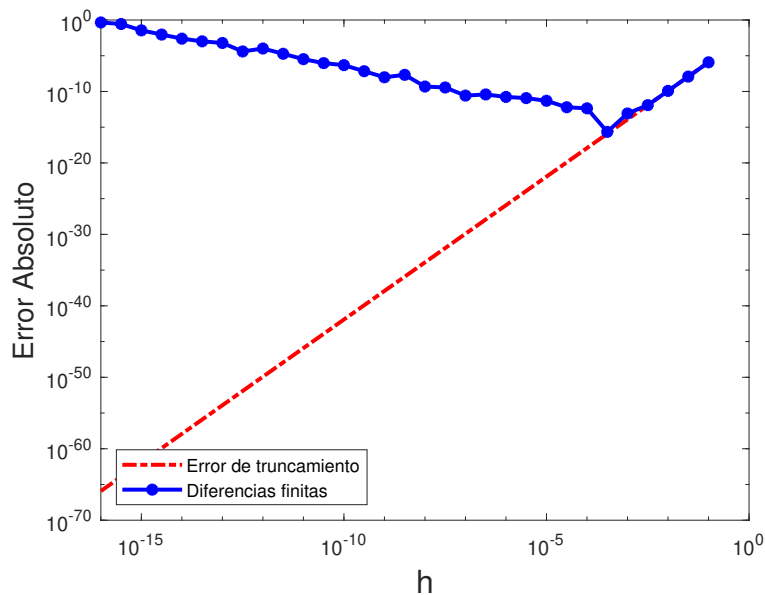


Figura 4: P5-2

## Problema6

Repaso de notación  $O$  mayúscula y  $o$  minúscula de una función. Si  $f(x) = O(x^2)$ ,  $g(x) = O(x^3)$  y  $h(x) = O(x^3)$  cuando  $x \rightarrow 0$ , ¿cuáles de las siguientes afirmaciones son verdaderas? Justifique sus respuestas.

- (a)  $f(x) = o(x)$  cuando  $x \rightarrow 0$

**Solución.** Cuando  $x \rightarrow 0$ , existe  $M > 0$  tal que  $|f(x)| \leq M|x^2|$ , como  $x \neq 0$ , tenemos que  $\left| \frac{f(x)}{x} \right| \leq M|x|$ , Luego  $\lim_{x \rightarrow 0} \left| \frac{f(x)}{x} \right| \leq 0$ . Por el valor absoluto y su continuidad, concluimos que

$$\lim_{x \rightarrow 0} \frac{f(x)}{x} = 0.$$

Es decir  $f(x) = o(x)$  cuando  $x \rightarrow 0$ . Así la afirmación es **Verdadera**.

- (b)  $f(x) = o(x^2)$  cuando  $x \rightarrow 0$

**Solución.** Sea  $f(x) = x^2$ , es claro que cuando  $x \rightarrow 0$ ,  $f(x) = O(x^2)$  Ahora note que  $\frac{f(x)}{x^2} = 1$ . Por lo tanto

$$\lim_{x \rightarrow 0} \frac{f(x)}{x^2} = 1.$$

Es decir que  $f(x) \neq o(x^2)$ . Así la afirmación es **Falsa**.

- (c)  $f(x) \cdot g(x) = O(x^5)$  cuando  $x \rightarrow 0$

**Solución.** Cuando  $x \rightarrow 0$  por hipótesis, existen  $M, N > 0$ , tales que  $|f(x)| \leq M|x^2|$  y  $|g(x)| \leq N|x^3|$ . Al ser cantidades positivas, la desigualdad se mantiene si las multiplicamos, es decir

$$|f(x)g(x)| \leq (M|x^2|)(N|x^3|).$$

Si tomamos  $K = MN > 0$ , tenemos que  $|f(x)g(x)| \leq K|x^5|$ , cuando  $x \rightarrow 0$ . Por lo que  $f(x)g(x) = O(x^5)$ . Así la afirmación es **Verdadera**.

(d)  $f(x) + g(x) = O(x^3)$  cuando  $x \rightarrow 0$

**Solución.** Sean  $f(x) = x^2$  y  $g(x) = x^3$ , es claro que cuando  $x \rightarrow 0$ ,  $f(x) = O(x^2)$  mientras que  $g(x) = O(x^3)$ . Supongamos que la afirmación es cierta. Luego existe  $M > 0$  tal que

$$|x^2 + x^3| \leq M|x^3|.$$

Así

$$\left| \frac{1}{x} + 1 \right| \leq M$$

Pero como esto se tiene cuando  $x \rightarrow 0$ , el lado derecho eventualmente se hará mas grande que cualquier constante fija. Por lo tanto  $f(x) + g(x) \neq O(x^3)$ . Así la afirmación es **Falsa**.

(e)  $g(x) - h(x) = 0$

**Solución.** Sean  $g(x) = 2x^3$  y  $h(x) = x^3$ . es claro que ambas son  $O(x^3)$  cuando  $x \rightarrow 0$  pero claramente

$$2x^3 - x^3 = x^3 \neq 0,$$

siempre que  $x \neq 0$ . que es el caso. Así la afirmación es **Falsa**.

(f)  $\frac{g(x)}{h(x)} = O(1)$  cuando  $x \rightarrow 0$

**Solución.** Sean  $g(x) = x^3$  y  $h(x) = x^4$ . Para  $0 < |x| < 1$  tenemos que  $|x^4| \leq |x^3|$ , luego cuando  $x \rightarrow 0$ , tenemos que  $g(x) = O(x^3)$  y  $h(x) = O(x^3)$ . Supongamos que la afirmación es cierta. En ese caso tendríamos que existe  $M > 0$  tal que

$$\left| \frac{1}{x} \right| = \left| \frac{x^3}{x^4} \right| \leq M.$$

Luego cuando  $x \rightarrow 0$  en algún punto el lado izquierdo sera mas grande que alguna constante fija, una contradicción. Mostrando así que  $\frac{g(x)}{h(x)} \neq O(1)$ . Así la afirmación es **Falsa**.