Temas: Programación en MATLAB: Sentencias, expresiones y variables. Estructuras de control. Operadores relacionales y lógicos. Programación de funciones.

Aritmética finita: Representación de números en computadora, dígitos significativos, errores de redondeo, precisión.

1. Evaluar el siguiente polinomio:

$$3.45x^7 - 1.32x^6 + 2.51x^5 + 0.73x^4 - 4.16x^3 + 2.27x^2 + 1.98x + 4.57$$

para valores de $x=0.23$, $x=-0.18$

- Calculando cada término por separado y sumándolos finalmente. ¿Cuántos productos por x se necesitaría realizar, si las potencias se calcularan como productos de x . (P/ej. $x^3 = x * x * x$) ?
- Calculando cada término por separado y sumándolos finalmente ¿Cuántos productos por x se necesitaría realizar, si las potencias se calcularan en forma anidada ? (P/ej. $x^4 = x * x^3$ y $x^3 = x * x^2$, etc.)
- ¿Cuántos productos por x se necesitaría realizar, si el polinomio se reescribiera de la siguiente forma?

$$4.57 + x(1.98 + x(2.27 + x(-4.16 + x(0.73 + x(2.51 + x(-1.32 + x(3.45)))))))$$

2. Escribir una función en MATLAB y evaluar con el polinomio y los puntos del ejercicio anterior.

- Que permita evaluar un polinomio de grado n . Los valores de los coeficientes se enviarán a la función en la forma de un vector, comenzando por el coeficiente de menor grado.
- Generalizar la función anterior para poder evaluar el polinomio en varios puntos simultáneamente.

3. Convertir de Binario a decimal

a) 11	d) 0.00000001	g) -1.0001
b) 1010101	e) -0.01010101	h) 101.0101
c) -100111	f) 0.00100111	i) 10.0111

4. Convertir de Decimal a Binario

a) -10	d) 0.1	g) 2.123
b) 127	e) 0.05	h) 7.102
c) 64	f) -0.10921	i) -11.023

5. Representar los siguientes números en cada uno de los formatos indicados: 1) Binario sin Signo, 2) Binario con Signo, 3) Complemento a 1, 4) Complemento a 2, 5) Exceso $2^{n-1} - 1$

a) 10100101	c) 10000001	e) 00011111
b) 00001101	d) 00000001	f) 11111111

6. Para cada secuencia binaria, indicar que valor decimal le corresponde, para cada formato indicado, considerando 8 bits: : 1) Binario sin Signo, 2) Binario con Signo, 3) Complemento a 1, 4) Complemento a 2, 5) Exceso $2^{n-1} - 1$

a) 16	c) 0	e) -5
b) -16	d) 10	f) -314

7. Representación Binaria

- Cuales son el máximo y el mínimo valor (en decimal) que se pueden representar en un sistema binario (sin signo) con 5 bits?
- Cuales son el máximo y el mínimo valor (en decimal) que se pueden representar en un sistema binario de complemento a 2 con 5 bits?
- Cuales son el máximo y el mínimo valor (en decimal) que se pueden representar en un sistema binario de con 5 bits y exceso 15?

8. Representación en coma flotante. Convertir de Binario a decimal

- Numeros binarios en formato flotante normalizado, con 16 bits, de los cuales uno es el bit de signo, 4 son bits de exponente (con exceso $z = 2^{4-1} - 1 = 7$) y 11 bits son de mantisa (mantisa= $0.1d_1d_2d_3\dots d_n$).
 - Signo = 0, Exponente = 0010, Mantisa = 01010110101
 - Signo = 0, Exponente = 1110, Mantisa = 11100000000
 - Signo = 1, Exponente = 1000, Mantisa = 10000000000

9. Convertir de Decimal a Binario

- A numeros enteros binarios con 8 digitos
 - 10
 - 127
 - 64
- A numeros binarios con 8 dígitos decimales
 - 0.1
 - 0.05
 - 0.10921
- Numeros binarios con 4 dígitos para la parte entera y 4 digitos para la parte decimal
 - 2.123
 - 7.102
 - 1.023
- A numeros binarios en formato flotante normalizado, con 16 bits, de los cuales uno es el bit de signo, 4 son bits de exponente (con exceso $z = 2^{4-1} - 1 = 7$) y 11 bits son de mantisa (mantisa= $0.1d_1d_2d_3\dots d_n$).
 - 2.5
 - 0.0121
 - 1021.2

10. Un número binario puede representarse en *punto flotante normalizado* (NBPFN) por medio de la siguiente expresión, donde **s** representa el *signo del número*, **f** la *fracción de la mantisa normalizada* y **e** el *exponente (con signo incluido)*.

En nuestro caso particular, MATLAB utiliza el formato **IEEE 754**. Por ejemplo, para el caso de números de doble precisión, tenemos:

$$NBPFN_{IEEE754} = (-1)^s * 2^{e-1023} * (1.f)$$

- a) Exprese la *cantidad de bits* y los rangos de **e** y **f** utilizados para la *representación binaria* de números en *simple precisión*.
- b) Idem con *doble precisión*.
- c) Debido a las limitaciones de representación del formato antes mencionado. ¿Cuál es la diferencia que existe entre el valor **1** y la representación del *número real* inmediatamente superior?

Observación: este es un ejercicio para buscar información en la bibliografía.

11. Supongamos que tenemos un sistema de representación de números en *punto flotante normalizado* que consta de solo **8 bits**, distribuidos de la siguiente manera: **1 bit** para el *signo del número*, **4 bits** para la *fracción de la mantisa normalizada* y **3 bits** para el *exponente* (*exceso* $2^{3-1} - 1 = 2^2 - 1 = 3$). (Asumimos que no hay representación de números especiales, excepto "0-0000-000"=0)
Para este sistema de representación, calcule y exprese en formato binario y decimal:

- a) El mayor número positivo que se puede representar.
- b) El número positivo mas pequeño que se puede representar.
- c) ¿Cuál es la diferencia que existe entre el valor **1** y la representación del *número real* inmediatamente superior ?

12. Dado los siguientes números reales:

i) 17.93, ii) -0.938, iii) 15.34, iv) 10.1, v) 3.57

- a) Transforme manualmente los siguientes números reales a binario y Realice el paso inverso para comprobar los resultados.
- b) Cuántos dígitos binarios necesitaría para representar exactamente cada número?
- c) Cuál es el error cometido en cada caso si la representación binaria esta limitada a 4 dígitos para la parte entera, 4 dígitos para la parte decimal, y un dígito de signo?

13. Calcular, manualmente (con ayuda de Matlab o la calculadora para hacer las cuentas) y usando las cuatro definiciones de la teoría (error absoluto vs. error relativo, 10^{-t} vs. $0.5 * 10^{-t}$), con cuantas cifras significativas es la aproximación $\hat{x}$ a x .

- a) $x = 3.141592$ a $\hat{x} = 3.1415$.
- b) $x = 3.141592$ a $\hat{x} = 3.14$.
- c) $x = 0.3141592$ a $\hat{x} = 0.314$.
- d) $x = 314.1592$ a $\hat{x} = 314$.
- e) $x = 0.0011174755$ a $\hat{x} = 0.00111748$.
- f) $x = 11.174755$ a $\hat{x} = 11.1748$.

14. Escribir un programa Matlab que reciba dos parametros (a=valor exacto y b=aproximacion), y calcule con cuantas cifras significativas es la aproximación de "a" por "b". Comparar con los resultados obtenidos en el ejercicio 13.

15. Una vez obtenido los resultados, comparar los valores entre si. Que conclusiones puede sacar comparando, en el ejercicio 13?:

- a) Puede estimar visualmente el valor a obtener cuando usa error absoluto? Justifique
 - b) Puede estimar visualmente el valor a obtener cuando usa error relativo? Justifique
 - c)Cuál sería la razón del uso del factor 0.5 ?
16. Dados los siguientes valores exactos, calcular su *valor redondeado*, su *error absoluto*, su *error relativo*, y las *cifras significativas* de la aproximación del valor exacto por su valor redondeado.
- a) Valor exacto = 0.42331082513075, redondear a **6** decimales.
 - b) Valor exacto = 0.31288191422707, redondear a **9** decimales.

17. Cuando dos números muy próximos entre sí se restan, se produce un fenómeno denominado *cancelación sustractiva*. Las siguientes funciones $f(x)$ y $g(x)$ son algebraicamente equivalentes. Calcule manualmente, con ayuda de Matlab o la calculadora, las expresiones con solo 6 cifras significativas (usando redondeo en cada operación), y además sin pérdida de cifras significativas. Calcular el error absoluto y relativo producido por usar menos cifras significativas. Comparar $g(x)$ y $f(x)$ obtenidos sin pérdida de cifras. Porque son diferentes?

a) $f(500)$ y $g(500)$ para

$$f(x) = x(\sqrt{x+1} - \sqrt{x}) \quad \text{para } x \gg 1$$

$$g(x) = x/(\sqrt{x+1} + \sqrt{x}) \quad \text{para } x \gg 1$$

Errores de truncamiento:

a) La serie de Maclaurin para la función "cos(x)" es

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \dots = \sum_{k=0}^n \frac{(-1)^k x^{2k}}{(2k)!}$$

Aplicar la serie a $x = 3 * \pi/4$ con $n = 1, 2, 3, 4, 5$ y calcular el error absoluto y relativo respecto a $\cos(x)$ calculado con Matlab.

b) La serie de Maclaurin para la función "sin(x)" es

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \dots = \sum_{k=1}^n \frac{(-1)^{k-1} x^{2k-1}}{(2k-1)!}$$

Aplicar la serie a $x = 3 * \pi/4$ con $n = 1, 2, 3, 4, 5$ y calcular el error absoluto y relativo respecto a $\sin(x)$ calculado con Matlab.

18. Ejercicio Adicional 1

- a) Representar 138,796 en formato de punto flotante normalizado con 16 bits. Donde el primer bit es para el signo, 11 bits para la mantisa (mantisa = $0,1d_1d_2d_3\dots d_n$) y los 4 bits restantes para el exponente (con exceso 8).

- b) Convertir el número representado en coma flotante a decimal.
- c) ¿Qué tipo de error se comete en a)?
- d) Calcular el error absoluto y relativo de representar 138.796 con el formato de punto flotante del inciso a)
- e) Calcular cuantos dígitos significativos tiene dicha representación.
- f) Si quisiera representar el mismo número en formato de punto flotante pero ahora con 8 bits; de los cuales uno es para el signo, 4 bits para la mantisa y 3 bits para el exponente (exceso 4). ¿Podría hacerlo? Justificar.