REPASO EXAMEN TEORÍA
Tema 1: Errores. Coma Flotante

NOVIEMBRE12

Problema 1 (3 puntos): Sea una representación en coma flotante en base 2 que almacena en memoria 3 bits de mantisa (b_1, b_2, b_3) y un exponente e que puede tomar los valores 0, 1, 2, 3.

Si $e=0$ el número máquina es $\hat{x} = \frac{b_1}{2} + \frac{b_2}{4} + \frac{b_3}{8}$.

Si $e \neq 0$ el número máquina es $\hat{x} = (1 + \frac{b_1}{2} + \frac{b_2}{4} + \frac{b_3}{8}) \times 2^{e-1}$.

i) Cuando $e=0$, ¿cuántos números máquina podemos representar?, calcular los dos números máquina menores y los dos números máquina mayores representados.

ii) Cuando $e \neq 0$, calcular el valor mínimo y máximo de esta representación. Indicar los bits de la mantisa y el valor del exponente a almacenar en memoria.

iii) Consideramos ahora la representación conjunta de los dos casos anteriores, cuando $e=0$ y cuando $e \neq 0$. Con los números calculados en los apartados anteriores, ¿cuánto valen $v_{\min}$ (valor mínimo distinto de cero) y $v_{\max}$ (valor máximo) de la representación?.

Calcular:

- El número máquina de x = 5.4. Indicar los bits de la mantisa y valor del exponente a almacenar en memoria.
- $\hat{x}_1$ el siguiente número máquina después de $x=1$, y el eps(1) (distancia entre el número máquina 1 y $\hat{x}_1$).
- Calcular una cota del error relativo.
- ¿Cuántas cifras significativas de precisión obtenemos con esta representación?. ¿Cuál es el número mínimo de bits de mantisa necesarios para obtener 3 cifras significativas de precisión?.

- Si e = 0, podemos representar 2^3 = 8 números máquina
  vmin1 = 0*2^-1 + 0*2^-2 + 0*2^-3 = 0
  vmin2 = 0*2^-1 + 0*2^-2 + 1*2^-3 = 2^-3 = 0.125
  vmax1 = 1*2^-1 + 1*2^-2 + 1*2^-3 = 0.875
  vmax2 = 1*2^-1 + 1*2^-2 + 0*2^-3 = 0.75

- Si e ≠ 0:
  vmin = (1 + 0*2^-1 + 0*2^-2 + 0*2^-3) x 2^(1-1) = 1x2^0 = 1
  vmax = (1 + 1*2^-1 + 1*2^-2 + 1*2^-3) x 2^(3-1) = 2^2 + 2 + 2^0 + 2^-1 = 7.5

- Teniendo en cuenta todos los números máquina representables, nos quedan:
  vmin = 0.125
  vmax = 7.5

Para representar el número real 5.4:
(1) Nos fijamos el exponente que necesitamos:
Si e = 3 → vmin = 1 x 2^(3-1) = 2^2 = 4 ← este es nuestro exponente

(2) A partir de e = 3, nos fijamos la mantisa que necesitamos:
$\hat{x}$ = (1 + b1*2^-1 + b2*2^-2 + b3*2^-3) x 2^2 = (2^2 + b1*2 + b2*2^0 + b3*2^-1) =
= (2^2 + 0*2 + 1*2^0 + 1*2^-1) = 5.5, donde m = 011 y e = 3

eps = eps(1) = (1 + 0*2^-1 + 0*2^-2 + 1*2^-3) x 2^0 - (1 + 0*2^-1 + 0*2^-2 + 0*2^-3) x 2^0 = 2^-3 = 0.125

Cota del error relativo de la representación:
eps/2 = 0.125/2 = 0.0625 = 6.25e-2

El número de cifras significativas de precisión correctas es 2, ya que -log10(6.25e-2) = 1

Para obtener 3 cifras significativas de precisión correctas necesitamos que se verifique la siguiente desigualdad
2^(-número bits mantisa) ≤ 10^(-número cifras significativas de precisión correctas)
2^-n ≤ 10^-3
-n log10(2) ≤ -3
n ≥ 3/log10(2) = 9.97
Luego, necesitamos un mínimo de 10 bits en la mantisa para garantizar 3 cifras significativas de precisión correctas.

NOVIEMBRE13

Problema 1. Se considera una representación en coma flotante en base 2. Cada palabra utiliza en memoria los siguientes 6 bits: 3 bits para el exponente, $e=(e_1\, e_2\, e_3)$, y 3 bits para la mantisa, $m=(b_1\, b_2\, b_3)$. Los números máquina $\hat{x}$ representados y su denominación son los siguientes:

Si $e=(0\,0\,0)_2=0$, $\quad \hat{x}=(0.b_1b_2b_3)_2\times 2^{-3}$ $\quad$ Número denormalizado

Si $e=(e_1\, e_2\, e_3)_2\neq 0$, $\quad \hat{x}=(1.b_1b_2b_3)_2\times 2^{e-4}$ $\quad$ Número normalizado

En esta representación:

- ¿Cuántos números denormalizados hay?, ¿cuántos normalizados? y, ¿cuántos números máquina?
- Calcular los números máquina (en formato decimal) y el contenido de los 6 bits a almacenar en memoria para los siguientes números: $v_{\min}$ valor mínimo, $v_{\max}$ valor máximo, $v_{\max}^{d}$ valor denormalizado máximo, $v_{\min}^{n}$ valor normalizado mínimo, y los números reales 1 y 9. Completar la siguiente tabla:

	Nº máquina	$e_1\, e_2\, e_3$	$b_1\, b_2\, b_3$
$v_{\min}$			
$v_{\max}^{d}$			
$v_{\min}^{n}$			
$v_{\max}$			
1			
9			

- Sea la sucesión $\hat{x}_n=2^{-n}$, $\quad n=0,1,2,3,...$ Calcular el menor n que verifica $1+\hat{x}_n=1$. Calcular el menor m que verifica $9+\hat{x}_m=9$.

- Tenemos un total de 2^3 = 8 números máquina denormalizados y 2^3x(2^3-1) = 8x7 = 56 número máquina normalizados y 8 + 56 = 64 números máquina en total

- vmin = (0.001)x2^-3 = 2^-3x2^-3 = 2^-6 = 0.015625 con e = 000, b = 001
  vmax_d = (0.111) x 2^-3 = (2^-1 + 2^-2 + 2^-3) x 2^-3 = (2^-4 + 2^-5 + 2^-6) = 0.109375, con e = 000, b = 111
  vmin_n = (1.000) x 2^(1-4) = 2^-3 = 0.125, con e = 001, b = 000
  vmax = (1.111) x 2^(7-4) = (1 + 2^-1 + 2^-2 + 2^-3) x 2^3 = 2^3 + 2^2 + 2 + 1 = 15, con e = 111, b = 111
  1 = (1.000) x 2^(4-4) = 1 x 2^0 = 1, con e = 100, b = 000
  9 = (1.001) x 2^(7-4) = (1 + 2^-3) x 2^3 = 2^3 + 1 = 9, con e = 111, b = 001

- eps(1) = 1.001 x 2^0 - 1.000 x 2^0 = 0.001 x 2^0 = 2^-3
  Luego, será eps/2 = 2^-4, es decir, n = 4

  eps(9) = 1.001 x 2^3 - 1.000 x 2^3 = 0.001 x 2^3 = 1
  Luego, será eps(9)/2 = 2^-1, es decir, m = 1

REPASO EXAMEN TEORÍA
Tema 2: Interpolación
Tema 3: Ajuste. Mejor Aproximación

ENERO 11

Problema 1:

a) Plantear el sistema lineal a resolver para hallar el spline cuadrático que interpola la siguiente tabla:

X	-1	0	1
Y	1	1	2

con la condición adicional de que s'(-1)=0.

b) Queremos hallar la función del tipo $u(t) = a + b\cos(t) + c\sin(t)$ que mejor ajuste los datos de la siguiente tabla:

t_k	$-\pi/2$	0	$\pi/2$	π
f_k	1	2	-1	1

con la restricción de que su derivada en cero debe valer 1, esto es, u'(0) = 1.

Plantear (sin resolver) el sistema sobredeterminado resultante.

a) Spline cuadrático → debe ser continuo en la función y en la primera derivada
Debemos construirnos dos polinomios genéricos de grado 2

$$s(x) = \begin{cases} a + bx + cx^2 & x \in [-1,0] \\ d + ex + fx^2 & x \in (0, 1] \end{cases}$$

Estudiamos continuidad en el punto conflictivo, es decir, en x = 0:

- Para la función:

$\lim_{x \to 0-} a + bx + cx^2 = a$ Luego, d = a

$\lim_{x \to 0+} d + ex + fx^2 = d$

- Para la primera derivada:

$\lim_{x \to 0-} b + 2cx = b$ Luego, e = b

$\lim_{x \to 0+} e + 2fx = e$

Sustituyendo en el spline, nos queda:

$$s(x) = \begin{cases} a + bx + cx^2, & x\in[-1,0] \\ a + bx + fx^2, & x\in(0,1] \end{cases}$$

Aplicamos los datos de la tabla + la condición adicional:

$s(-1) = a - b + c = 1$	$b = c \rightarrow b = 0$
$s(0) = a = 1$	$a = 1$
$s(1) = a + b + f = 2$	$f = 2 - a - b = 1$
$s'(-1) = b - 2c = 0$	$c - 2c = 0; -c = 0; c = 0$

Sustituyendo obtenemos finalmente:

$$s(x) = \begin{cases} 1 & x\in[-1,0] \\ 1 + x^2 & x\in(0,1] \end{cases}$$

b) Verificamos la restricción adicional $u'(0) = 1$:
$u'(t) = c\cos(t) - b\sin(t)$
$u'(0) = c = 1$

Sustituimos en la función original:
$u(t) = a + b\cos(t) + \sin(t)$

Planteamos el sistema lineal sobredeterminado:

$u(-\pi/2) = a + b\cos(-\pi/2) + \sin(-\pi/2) = 1$	$\rightarrow$	$a = 2$
$u(0) = a + b\cos(0) + \sin(0) = 2$	$\rightarrow$	$a + b = 2$
$u(\pi/2) = a + b\cos(\pi/2) + \sin(\pi/2) = -1$	$\rightarrow$	$a = -2$
$u(\pi) = a + b\cos(\pi) + \sin(\pi) = 1$	$\rightarrow$	$a - b = 1$

JULIO18

a) Hallar un polinomio de grado mínimo verificando $p(0)=1$, $p(1)=1$, $p'(1)=0$, $p(2)=3$. ¿De qué grado será el polinomio? Justificar.

El grado del polinomio será 3, ya que tenemos 4 datos.
Tabla de diferencias divididas:

x	p(x)	p(. , .)	p(. , . , .)	p(. , . , . , .)
$x0 = 0$	$A0 = p(0) = 1$	$A1 = (1 - 1) / (1 - 0) = 0$	$A2 = (0 - 0) / (1 - 0) = 0$	$A3 = (2 - 0)/(2 - 0) = 1$
$x1 = 1$	$p(1) = 1$	$p'(1) = 0$	$(2 - 0)/(2 - 1) = 2$	
$x2 = 1$	$p(1) = 1$	$(3 - 1) / (2 - 1) = 2$		
$x3 = 2$	$p(2) = 3$			

$p(x) = 1 + x(x - 1)^2$

b) Sea la siguiente tabla, se quiere ajustar sus datos con un polinomio de grado 2: p(x) = a+bx+cx^2

xk	0	1	2	3
fk	0	0	1	4

i) Plantear el sistema sobredeterminado a resolver para hallar el polinomio p(x) que mejor ajuste los datos de la tabla.

p(0) = a = 0
p(1) = a + b + c = 0
p(2) = a + 2b + 4c = 1
p(3) = a + 3b + 9c = 4

ii) Ahora queremos que p(x) VERIFIQUE de forma exacta las dos primeras condiciones y pase lo más cerca posible de las 2 restantes. Plantear el nuevo sistema sobredeterminado, resolverlo a través de las ecuaciones normales y dar la solución del p(x) correspondiente.

Verificamos las dos primeras condiciones:
p(0) = a = 0 → a = 0
p(1) = b + c = 0 → c = -b

Sustituyendo en el polinomio original nos queda:
p(x) = b(x - x^2)

Planteamos el sistema lineal sobredeterminado:
p(2) = -2b = 1
p(3) = -6b = 4

Pasamos a forma matricial: Hc = B, donde H = {matriz de coeficientes}, c = {vector de coeficientes} y B = {vector de términos independientes}

$$H = \begin{matrix} -2 \\ -6 \end{matrix} \qquad c = (b) \qquad B = \begin{matrix} 1 \\ 4 \end{matrix}$$

Planteamos las ecuaciones normales y resolvemos: H'Hc = H'B

$$(-2\ \ -6) \begin{matrix} -2 \\ -6 \end{matrix} \quad (b) \quad = \quad (-2 \quad -6) \begin{matrix} 1 \\ 4 \end{matrix}$$

40b = -26 → b = -26/40 = -13/20

p(x) = (13/20)(x^2 - x)

iii) ¿Cuál de los polinomios anteriores minimiza E = $\sum_k (f_k - p(x_k))^2$?

El error (por mínimos cuadrados) se minimiza en el primer polinomio, ya que tiene más parámetros libres y se ajustará mejor a toda la función que el segundo que solo tiene un grado de libertad (o parámetro libre).

JULIO17

2.1. - Dar la expresión mediante la fórmula de Newton generalizada de un polinomio p(x) de grado dos que interpola los datos: p(0)=0, p'(0)=1 y p(1)=10.

- Se considera la función a trozos s(x)

$$s(x) = \begin{cases} ax + x^2 & x \in [-1,0] \\ p(x) & x \in [0,1] \end{cases}$$

donde p(x) es el polinomio calculado anteriormente. Determinar el valor del parámetro a para s(x) sea una función spline cuadrado.

Tabla de diferencias divididas:

x	p(x)	p(. , .)	p(. , . , .)
x0 = 0	A0 = p(0) = 0	A1 = p'(0) = 1	A2 = (10 - 1)/(1 - 0) = 9
x1 = 0	p(0) = 0	(10 - 0)/(1 - 0) = 10	
x2 = 1	p(1) = 10		

p(x) = x + 9 x^2

Luego, el spline cuadrado nos queda:

$$s(x) = \begin{cases} ax + x^2 & x \in [-1,0] \\ x + 9x^2 & x \in [0, 1] \end{cases}$$

Estudiamos continuidad en la función y en la primera derivada:

- Para la función:
  $\lim_{x \to 0-}$ ax + x^2 = 0
  $\lim_{x \to 0+}$ x + 9x^2 = 0
- Para la primera derivada:
  $\lim_{x \to 0-}$ a + 2x = a   Luego, a = 1
  $\lim_{x \to 0+}$ 1 + 18x = 1

Sustituyendo en el spline nos queda:

$$s(x) = \begin{cases} x + x^2 & x \in [-1,0] \\ x + 9x^2 & x \in [0, 1] \end{cases}$$

2.2. - Se considera el problema de ajustar los datos de la siguiente tabla

x_i	-2	-1	0	1	2
y_i	1	3	8	3	1

por una función aproximante del tipo $u(x) = \dfrac{a}{1+bx^2}$. Transformar dicho problema en un problema de ajuste lineal y construir el sistema lineal sobredeterminado Hc=B (siendo c el vector de incógnitas) al que se llega. Se pide: Dar las expresiones de la matriz H de coeficientes y del término independiente B del sistema.

- Si se ajustaran los datos por una función del tipo $u(x)=1/(1+bx^2)$ ¿el error del ajuste sería mayor o menor que el producido en el caso anterior? Dar una respuesta razonada, sin calcular explícitamente los errores referidos.

Planteamos el sistema lineal sobredeterminado:

u(-2) = a / (1 + 4b) = 1 → a - 4b = 1
u(-1) = a / (1 + b) = 3 → a - 3b = 3
u(0) = a = 8 → a = 8
u(1) = a / (1 + b) = 3 → a - 3b = 3
u(2) = a / (1 + 4b) = 1 → a - 4b = 1

Pasamos a forma matricial: Hc = B, con H = {matriz de coeficientes}, c = {vector de coeficientes}
B = {vector de términos independientes}

$$H = \begin{matrix} 1 & -4 \\ 1 & -3 \\ 1 & 0 \\ 1 & -3 \\ 1 & -4 \end{matrix} \qquad c = \begin{matrix} a \\ b \end{matrix} \qquad B = \begin{matrix} 1 \\ 3 \\ 8 \\ 3 \\ 1 \end{matrix}$$

En el caso de ajustar los datos por una función del tipo u(x) = 1/(1+bx^2), al tener un grado menos de libertad, es decir, un parámetro menos libre (a = 1), entonces el error producido será mayor en este segundo caso.

REPASO EXAMEN TEORÍA
Tema 4: Ecuaciones No Lineales

ENERO 11

Problema 2: Se quiere aproximar la solución s de la ecuación $x^3+3x^2-1=0$ en el intervalo $[-1,-1/2]$. Para ello se utiliza el método de Newton, obteniéndose la sucesión $\{x_n\}$ a partir de un punto inicial x_0 del intervalo.

a) Probar que se cumple $|x_{n+1}-s| \le \left(\frac{2}{3}\right)^{2^n-1}(x_0-s)^{2^n}$, $\quad n=0,1,\ldots$

b) Suponiendo que el error inicial es $|x_0-s| \le 1/4$, determinar cuántas iteraciones se necesitan para obtener una aproximación de la solución s con un error no superior a 10^{-5}.

a) Necesitamos probar que |Ce0|^(2^n)/|C| ≤ (2/3)^((2^n) - 1), ya que sabemos que |en| = |xn+1 - s|
donde |e0| = |-1 - (-1/2)| = 0.5
|C| = M/2m, con M = max|f''(x)|[-1,-1/2] y m = min|f'(x)|[-1,-1/2]

Calculamos las derivadas primera y segunda:
f'(x) = 3x^2 + 6x → min → f'(-1) = 3 - 6 = -3 → 3
f'(-½) = 3(-1/2)^2 - 6/2 = - 2.25 → 2.25 ← este

f''(x) = 6x + 6 → max → f''(-1) = 0
f''(-0.5) = 3 ← este

C = 3/(2*2.25) = ⅔ = 0.666…

|en| ≤ $\frac{|e0*2/3|\text{^}(2\text{^}n)}{2/3}$ = |e0| ^(2^n) * |2/3|^(2^n) * |2/3|^-1 = |2/3| ^((2^n) - 1) * |e0|^(2^n),
donde e0 = |x0 - s|

b) $\frac{|1/4*2/3|\text{^}(2\text{^}n)}{2/3}$ < 10^-5

|1/6|^(2^n) < (⅔) * 10^-5
2^n * log10(⅙) < log10(2/3) - 5 ← aplicamos log10
2^n > (log10(2/3) - 5)/log10(1/6)
n > log2((log10(2/3) - 5)/log10(1/6)) = 2.73… ← aplicamos log2

Luego, el mínimo número de iteraciones será 3

JULIO19

Ejercicio 2 PROBLEMAS: Dada la ecuación $x^3+3x^2-3x-9=0$

a) Demostrar que tiene una única solución s en el intervalo [1.5, 2].

Tma Bolzano:
f(1.5) = 1.5^3 + 3*1.5^2 - 3*1.5 - 9 = -3.375
f(2) = 2^3 + 3*2^2 - 3*2 - 9 = 5

Como f(1.5)*f(2) < 0, entonces la función tiene una raíz en el intervalo [1.5, 2] por el teorema de Bolzano

b) Demostrar que el método de Newton-Raphson converge empezando en cualquier punto x0 de dicho intervalo.

El método de Newton converge si |Ce0| ≤ 1

e0 = |2 - 1.5| = 0.5

C = M/2m, donde M = max|f''(x)|[1.5,2] y m = min|f'(x)|[1.5,2]

Calculamos las derivadas:
f'(x) = 3x^2 + 6x - 3 → min → f'(1.5) = 27/4 + 9 - 3 = 12.75

f''(x) = 6x + 6 → max → f''(2) = 18

C = 18 / (2*12.75) = 12/17

Como |Ce0| = 12/17 * 0.5 = 6/17 < 1, entonces el método de Newton converge para cualquier x0 perteneciente al intervalo [1.5, 2]

c) Partiendo del punto medio x0 = 1.75, dar el resultado de la primera iteración del método. A partir de este único cálculo, estimar el error de la hipótesis inicial x0.

x1 = xn - f(xn) / f'(xn) = 1.75 - (1.75^3 + 3*1.75^2 - 3*1.75 - 9)/(3*1.75^2 + 6*1.75 - 3) = 1.732209…
|e0| = |x1 - x0| = |1.732209 - 1.75| = 0.017791

d) Sabiendo que la solución exacta es $s=\sqrt{3}$**, ¿cuál es el error relativo de la aproximación? ¿Cuántas cifras decimales significativas se han obtenido?**

Erel = |x - $\hat{x}$| / |x| = |1.732209 - sqrt(3)| / |sqrt(3)| = 9.13e-5
Ncif = -log10(9.13e-5) = 4.03 → Luego se obtienen 4 cifras decimales significativas correctas

JULIO18

Ejercicio 2 (30%) Dada la ecuación $e^{-x} = x + 2$:

a) Demostrar que la ecuación tiene una raíz s en el intervalo [-1,0].
b) Queremos aplicar el método de Newton para hallar la raíz del intervalo [-1,0]. ¿Qué error inicial $e_0 = |x_0 - s|$ máximo garantiza que el método converge?
c) Si tomamos $x_0 = -0.8$ como valor inicial, justificar si el método de Newton converge. Calcular la primera iteración x_1 del método partiendo de $x_0 = -0.8$.
d) Sin conocer la raíz s, a partir del valor x_1 del apartado anterior y sabiendo que $x_2 = -0.4433$, $x_3 = -0.4429$ estimar el error cometido (e_1, e_2, e_3) en las 3 iteraciones.

f(x) = e^-x - x - 2 = 0 → pasamos la función a un mismo término

a) Tma Bolzano:
f(-1) = e +1 -2 = 1.718282… > 0
f(0) = 1 - 0 - 2 = -1 < 0
Como f(-1) * f(0) < 0, entonces por el Tma Bolzano existe, al menos, una raíz en [-1,0]

b) Método converge para qué e0
El método de Newton converge si |Ce0| ≤ 1 → e0 ≤ 1/C, donde…
C = M/2m, con…
m = min|f'(x)|[-1,0] → f'(x) = -e^(-x) - 1 → f'(0) = 2; f'(-1) = 3.718282…
M = max|f''(x)|[-1,0] → f''(x) = e^(-x) → f''(0) = 1, f''(-1) = e = 2.718282…
C = 2.718282/(2*2) = 0.67957…

Luego, el máximo error inicial permitido será de 1/0.67957 = 1.4715…

c) x0 = -0.8 → ¿converge?
Dado que x0∈[-1,0] y que el máximo error permitido es mayor que la longitud del intervalo |-1 - 0| = 1, entonces el método de Newton converge
x1 = x0 - f(x0) / f'(x0) = -0.8 - (e^0.8 + 0.8 - 2)/(-e^0.8 - 1) = -0.4821…

d) e1 = |x2 - x1| = |-0.4433 + 0.4821| = 0.0388
e2 = |x3 - x2| = |-0.4429 + 0.4433| = 0.0004
Dado que el error es cuadrático, entonces podemos hacer: e2≃k*e1^2
k ≃e2/e1^2 = 0.0004/0.0388^2 = 0.2657
e3 ≃k*e2^2 = 0.2657*0.0004^2 = 0.000000043