Control 1 MA-33A-1

Profesor Dr. Gonzalo Hernandez - Auxiliar Gonzalo Rios

UChile - Departamento de Ingeniería Matemática

Semestre Otoño 2006

1) (25%) Sea una codificación binaria floating point tipo inicial, donde el número real máximo representable es $2^{63} - 2^{54}$.

   (a) Explique cuantos bits son necesarios para implementar esta codificación y como se distribuyen.

   (b) Sea $x = a_1 a_2 ... a_n$ donde n es el números de bits de la codificación usada y sea
   $a_i = \begin{cases} 1 & \text{si } i = 2j \vee i = 1 + 4j \ , j \in \mathbb{N} \cup \{0\} \\ 0 & \text{en otro caso} \end{cases}$
   Obtenga el real representado por x.

   (c) Sea $f(x) = e^{x^2}$. Para una aritmética finita de 5 cifras significativas de redondeo:

      i) Calcule el polinomio de Taylor de orden 4 de $f(x)$ en torno a $x_0 = 0$.

      ii) Calcule $p_4(0.5)$ con la misma aritmética.

      iii) Represente $f(0.5)$ y $p_4(0.5)$ según la codificación de punto flotante anterior. Calcule el error relativo y según esto concluya si $p_4(0.5)$ es o no una buena aproximación de $f(0.5)$

2) (25%) Sea la serie $S = \sum\limits_{k=1}^{\infty} \frac{1}{k^2}$. Denotemos por $S_n = \sum\limits_{k=1}^{n} \frac{1}{k^2}$ a la suma parcial n-ésima exacta, y por Z_n la suma parcial n-ésima calculada por el computador según el algoritmo:

   $Z_k = \begin{cases} 1 \ k = 1 \\ Z_{k-1} \oplus (1 \oslash (k \odot k)) \ \ k \geq 2 \end{cases}$

   El error absoluto está dado por $E_{abs} = |S - Z_n|$

(a) Utilizando la identidad $S - Z_n = S - Z_n + S_n - S_n$, indique como el error se puede dividir en un error de truncación y en un error de redondeo. Determine su comportamiento cuando $n \to \infty$

(b) Se puede demostrar que $Z_n = S_n + E_n$ donde $E_n \leq cn$, c constante. También se puede demostrar que:

$$\sum_{k=n+1}^{\infty} f(k) \leq \int_{n}^{\infty} f(x)dx$$

A partir de estos propuestos, determine una función $g(n)$ que acote el error $E_{abs} = |S - Z_n|$.

(c) Encuentre un valor de n óptimo que minimice la suma de las cotas encontradas anteriormente. Estime cuantos términos de la serie se deben sumar y cuantos dígitos decimales correctos tendrá el resultado calculado si $c = 4 \times 10^{-16}$.

(d) Sea la función $\phi(x, y) = g(x) \times g(y)$. Haga un análisis de propagación de error de la función $\phi(x, y)$. Considere $c > 0$.

3) (50%) Dado el SEL:

$$A = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 2 & 4 \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

(a) Métodos Directos:

i) Determine su descomposición $A = LU$ y resuelva el SEL utilizando esta factorización.

iii) Es A definida positiva ? Si lo es, determine su factorización de Cholesky y resuelva el SEL utilizando esta factorización.

(b) Métodos Iterativos:

i) Resuelva el SEL mediante el método Gauss-Seidel.

ii) Para el método SOR aplicado a matrices definidas positivas y tridiagonales, el valor óptimo de ω está dado por:

$$\overline{\omega} = \frac{2}{1 + \sqrt{1 - \rho(T_G)}}$$

Donde $\rho(T_G)$ es el radio espectral de la matriz del método de Gauss-Seidel T_G.

Calcule $\overline{\omega}$ y resuelva el SEL mediante el método SOR para $\overline{\omega}$.