Práctica voluntaria (1er. parcial) — Técnicas Numéricas
Profesor Francisco R. Villatoro — 7 de Octubre de 1999

PRIMER TEMA: Operaciones en coma flotante en Matlab

1. La aritmética que utiliza Matlab es compatible con el sistema IEEE a doble precisión, es decir, 53 bits binarios en la mantisa (el primero de los cuales se aprovecha para el signo) y un exponente entre $-1022 \leq e \leq 1023$.

   Los números flotantes en Matlab (IEEE) tienen la forma

   $$fl(x) = \sigma \cdot (1.a_1 a_2 \cdots a_{53})_2 \cdot 2^e,$$

   donde σ es el signo, $a_i \in \{0, 1\}$, y $-1022 \leq e \leq 1023$. ¿Cuántos dígitos binarios o bits requiere el exponente? ¿Cuántos bytes (grupo de 8 bits) se consumen en representar un número flotante?

   ¿Cuántos números con exponente $e = 0$ puede representar Matlab? ¿Cuál es el número positivo más pequeño representable (para $e = 0$ y para $e \neq 0$)? ¿y cuál es el más grande?

   En el sistema IEEE de doble precisión además del 0, se incluyen los números $\pm$`Inf` (infinito positivo y negativo) y `NaN` (el resultado no es un número). Prueba a calcular en Matlab `1/0`, `log(1/0)` y `1/0-log(1/0)`, ¿qué resultados obtienes? Justifica tu respuesta y pon otro ejemplo de tu invención de similares resultados.

2. Calcule el epsilon de la máquina para Matlab (comando `eps`).

   El siguiente programa

   ```
   e = 1;  while (1+e > 1), e=e/2; end,
   ```

   ¿calcula el epsilon de la máquina? Si no lo hace, cómo lo arreglarías y por qué.

   Escribe en representación flotante en binario en el sistema IEEE de Matlab el número epsilon de la máquina (ε) y el número $1 + \varepsilon$.

3. ¿Escribe una cota superior del error de normalización de un número flotante cuando se utiliza redondeo y otra cuando se utiliza truncado? Justifica tu respuesta. Calcula estas cotas para Matlab.

El error relativo de normalización de un número δ, donde $fl(x) = x\,(1+\delta)$, se puede acotar por

$$-2^{-n} \leq \delta \leq 2^{-n}, \quad \text{por redondeo,}$$

$$-2^{-n+1} \leq \delta \leq 0, \quad \text{por truncado.}$$

Justifica las expresiones anteriores.

4. Un ejemplo sencillo de los efectos de los errores en aritmética flotante.

   ¿Qué tiene que dar y que es lo que da realmente el programa siguiente?

   `x = ones(1,6)*987654321^2, mean(x), std(x),`

   ¿y este otro?

   `x = ones(1,7)*987654321^2, mean(x), std(x),`

   Justifica y razona las respuestas que has obtenido en Matlab.

5. ¿Qué hace el siguiente programa?

   `x = 10.^(0:20); x.*(sqrt(x+1)-sqrt(x))`

   ¿Cuál es el efecto de la diferencia cancelativa en el resultado? ¿Cómo calcularías el resultado evitando la diferencia cancelativa? Escribe un código en Maltab para ello. Compara los resultados de tu código con los del original y calcula los errores cometidos debido a la diferencia cancelativa.

6. Para el programa `x = 10.^(-(0:20)); (1 - cos(x) )./(x.^2)`, indica el efecto de la diferencia cancelativa e indica cómo lo evitarías (escribe un código Matlab). Calcula los errores cometidos debido a ella. Compara los resultados de tu código con los obtenidos utilizando la expresión

$$2\,\frac{\sin^2(x/2)}{x^2}.$$

   Razona el resultado que obtienes.

7. Escribe un programa Matlab que evalúe un polinomio utilizando la regla de Horner a partir de un vector con sus coeficientes que utilice el mismo formato que la función `polyval`, es decir, un fichero `horner.m` que contenga

```
function y = horner(p,x)
%%% donde p es un vector de coeficientes
%%% x el punto donde se evaluará el polinomio
%%%    AQUI VA EL ALG. HORNER
%%%    EL RESULTADO SERA CALCULADO EN y
```

Compara el número de operaciones (usando `flops`) de los códigos `polyval` y `horner`. ¿Qué conclusiones sacas?

8. Ejecuta el siguiente código para calcular una suma de números:

```
format long e;
a = 1./(200000:210000); b = 1./(210000:-1:200000);
sum(a), sum(b), sum(a)-sum(b),
```

¿son iguales los dos resultados? ¿Cuál es el más preciso? ¿Puedes estimar el tamaño de los errores cometidos en ambos casos?. Justifica tus respuestas y presenta los cálculos que consideres oportunos.

9. Para calcular la exponencial de un número se puede usar su desarrollo de Taylor

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

que se puede implementar en Matlab de la siguiente forma

```
n=20; %%% número de términos
potencias=0:n; factoriales=[1 cumprod(1:n)];
terminos=x.^potencias./factoriales;
exponencialdex = cumsum(terminos),
```

Calcula la exponencial para $x = -5'5$ con 7 dígitos de precisión. ¿Cuántos términos necesitas (`n`)? ¿Cómo estimarías a partir del desarrollo de Taylor ese número de términos?

Calcula uno partido la exponencial para $x = 5'5$ con 7 dígitos de precisión. ¿Cuántos términos necesitas (`n`)? ¿Cómo estimarías ese número a partir del desarrollo Taylor?

10. El ejemplo clásico de Wilkinson (1963) sobre la inestabilidad de métodos numéricos y la propagación de errores es el polinomio

$$p(x) = (x-1)(x-2)\cdots(x-19)(x-20) = x^{20} - 210\,x^{19} + \ldots,$$

cuyas raíces son reales y están bien separadas. Calculemoslas en Maltab:

```
lista = 1:20; factores = [lista*0+1; -lista]';
polinomio = [1];
for ind=1:20,
   polinomio = conv (polinomio, factores(ind,:));
end
roots (polinomio),
```

Ahora cambiaremos el coeficiente $-210\,x^19$ por $(-210 + 2^{-23})\,x^{19}$, y calcularemos de nuevo los ceros

```
polinomio(2) = polinomio(2) +2^(-23);
roots(polinomio),
```

¿cómo son los ceros del polinomio ahora? Comenta las diferencias entre los ceros de ambos polinomios.

¿Cómo estudiarías la sensitividad del polinomio

$$p(x, \alpha) = x^{20} - \alpha\, x^{19} + \ldots,$$

ante cambios en el valor de α? Es decir, cómo determinarías la influencia de un pequeño cambio en α sobre cada una de las raíces del polinomio.

11. Otro ejemplo clásico de cálculo de raíces de polinomios es

$$a\, x^2 + b\, x + c = 0,$$

para $a, b, c \in \mathbb{R}$ tiene como raíces

$$r_{\pm} = \frac{-b \pm \sqrt{b^2 - 4\, a\, c}}{2\, a}.$$

La evaluación de esta expresión, para una de las dos raíces reales cuando $b^2 \gg |4\, a\, c|$ involucra una diferencia cancelativa.

Un manera de evitar la diferencia cancelativa es utilizar las expresiones para las raíces

$$x_1 = -\frac{b + (b)\,\sqrt{b^2 - 4\, a\, c}}{2\, a}$$

$$x_2 = \frac{c}{a\, x_1},$$

a = 1, b= - $10^5, c = 1,$

true 99999.999990 0.000010000000001

a=6 10‘30, b = 5 $10^3 0, c = -410^3 0,$

a = $10^- 30, b = -10^3 0, c = 10^3 0, ...$

a = 1, b=-4, c = 3.9999999999999