Aritmética flotante

Métodos Numéricos

Prof. Juan Pablo Concha y Eduardo Uribe

Conferencia 3

Conferencia 3

Definición

Si p^* aproxima a p, el error absoluto es

$$E_a(p^*) = |p - p^*|,$$

y el error relativo (cuando $p \neq 0$) es

$$E_r(p^*) = \frac{|p - p^*|}{|p|}$$

Error relativo de $fl(x)$

Si $fl(x)$ se calcula para una mantisa de k dígitos:
Aproximando con corte:

$$E_r(fl(x)) = \frac{|x - fl(x)|}{|x|} \leq 10^{-k+1}$$

Aproximando con redondeo:

$$E_r(fl(x)) = \frac{|x - fl(x)|}{|x|} \leq 0{,}5 \times 10^{-k+1}$$

Operaciones elementales

Definición

$$
\begin{aligned}
x \oplus y &= fl(fl(x) + fl(y)) \\
x \ominus y &= fl(fl(x) - fl(y)) \\
x \otimes y &= fl(fl(x) \times fl(y)) \\
x \oslash y &= fl(fl(x)/fl(y))
\end{aligned}
$$

Ejemplo: $x = \frac{5}{7} = 0.\overline{714285}$; $y = \frac{1}{3} = 0{,}3\overline{3}$

Base:10; Mantisa: 5 dígitos ; Exponente: 1 dígito (sin signos)

$fl(x) = 0{,}71428 \times 10^0$; $fl(y) = 0{,}33333 \times 10^0$

	Resultado	Valor real	E. absoluto	E. relativo
$x \oplus y$	$0{,}10476 \times 10^1$	$22/21$	$0{,}190 \times 10^{-4}$	$0{,}182 \times 10^{-4}$
$x \ominus y$	$0{,}38095 \times 10^0$	$8/21$	$0{,}238 \times 10^{-5}$	$0{,}625 \times 10^{-5}$
$x \otimes y$	$0{,}23809 \times 10^0$	$5/21$	$0{,}524 \times 10^{-5}$	$0{,}220 \times 10^{-4}$
$x \oslash y$	$0{,}21428 \times 10^1$	$15/7$	$0{,}571 \times 10^{-4}$	$0{,}267 \times 10^{-4}$

Malos ejemplos!

Datos

$x = 0.\overline{714285}$ $\quad fl(x) = 0{,}71428 \times 10^0$

$u = 0{,}714251$ $\quad fl(u) = 0{,}71425 \times 10^0$

$v = 98765{,}9$ $\quad fl(x) = 0{,}98765 \times 10^5$

$w = 0{,}111111 \times 10^{-4}$ $\quad fl(x) = 0{,}11111 \times 10^{-4}$

Cálculos

	Resultado	Valor real	E. absoluto	E. relativo
$x \ominus u$	$0{,}30000 \times 10^{-4}$	$0{,}34714 \times 10^{-4}$	$0{,}471 \times 10^{-5}$	$0{,}136$
$(x \ominus u) \oslash w$	$0{,}27000 \times 10^{1}$	$0{,}31243 \times 10^{1}$	$0{,}424$	$0{,}136$
$(x \ominus u) \otimes v$	$0{,}29629 \times 10^{1}$	$0{,}34285 \times 10^{1}$	$0{,}465$	$0{,}136$
$x \oplus v$	$0{,}98765 \times 10^{5}$	$0{,}98766 \times 10^{5}$	$0{,}161 \times 10^{1}$	$0{,}163 \times 10^{-4}$

Observaciones

- Cancelación de resta (fuente principal de errores)
- Errores pueden aumentar al combinarse las operaciones.

Cancelación de resta

Problema

Cálcular los ceros de la ecuación $ax^2 + bx + c$ dados por

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-2c}{b \pm \sqrt{b^2 - 4ac}}$$

Ejemplo en el caso $b^2 \gg 4ac$

$x^2 + 62{,}1x + 1 = 0;\ x_1 = -0{,}01610723;\ x_2 = -62{,}08390$

Calculando con 4 dígitos y redondeando, con ambas expresiones (**para** x_1):

$$fl(x_1) = \frac{-62{,}1+\sqrt{(62{,}1)^2-4}}{2} = \frac{-62{,}1+62{,}06}{2} = -0{,}02 \quad E_r \approx 0{,}24$$

$$fl(x_1) = \frac{-2}{62{,}1+\sqrt{(62{,}1)^2-4}} = \frac{-2}{62{,}1+62{,}06} = -0{,}01610 \quad E_r \approx 0{,}00062$$

Cancelación de resta

Problema

Cálcular los ceros de la ecuación $ax^2 + bx + c$ dados por

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-2c}{b \pm \sqrt{b^2 - 4ac}}$$

Ejemplo en el caso $b^2 \gg 4ac$

$x^2 + 62{,}1x + 1 = 0$; $x_1 = -0{,}01610723$; $x_2 = -62{,}08390$

Calculando con 4 dígitos y redondeando, con ambas expresiones (**para** x_2):

$$fl(x_2) = \frac{-62{,}1 - \sqrt{(62{,}1)^2 - 4}}{2} = \frac{-62{,}1 - 62{,}06}{2} = -62{,}1 \quad E_r \approx 0{,}00032$$

$$fl(x_2) = \frac{-2}{62{,}1 - \sqrt{(62{,}1)^2 - 4}} = \frac{-2}{62{,}1 - 62{,}06} = -50 \quad E_r \approx 0{,}19$$

Ejercicios

1) Aproximar los números reales $x_1 = \dfrac{127}{7}$ y $x_2 = 56{,}786500$ empleando una aritmética de cuatro cifras significativas con la técnica de corte y la de redondeo.
2) Calcule la siguiente operación utilizando aritmética de punto flotante con corte al tercer dígito. Determine la exactitud del resultado obtenido.

$$\frac{\sqrt{5}(\pi + 24568)}{e^2 - 6}$$

3) Utilizando aritmética de siete dígitos decimales con redondeo efectuar los siguientes cálculos.
   a) Con $a = 1234{,}567,\ b = 45{,}67844,\ c = 0{,}0004$

   $$(a + b) + c, \qquad a + (b + c)$$

   b) Con $a = 1234{,}567,\ b = 1{,}234567,\ c = 3{,}333333$

   $$(a + b) \cdot c, \qquad a \cdot c + b \cdot c$$

Ejercicios

4) Use una aritmética de redondeo a cuatro cifras para determinar la aproximaciones más precisas de las raíces de las siguientes ecuaciones cuadráticas.

- $\frac{1}{3}x^2 - \frac{123}{4}x + \frac{1}{6} = 0$
- $1{,}002x^2 + 11{,}01x + 0{,}01265 = 0$

5) Suponga que dos puntos (x_0, y_0) y (x_1, y_1) están en una linea recta, con $y_1 \neq y_0$ Se tienen dos fórmulas para determinar la ordenada al origen de la recta:

$$x = \frac{x_o y_1 - x_1 y_0}{y_1 - y_0} \qquad \text{y} \qquad x = x_0 - \frac{(x_1 - x_0)y_0}{y_1 - y_0}$$

- Demuestre que ambas fórmulas son algebraicamente correctas.
- Utilice los datos $(x_0, y_0 = (1{,}31,\ 3{,}24))$ y $(x_1, y_1) = (1{,}93,\ 4{,}76)$ y la aritmética de redondeo a tres cifras para calcular la ordenada al origen de ambas formas. ¿Cuál método es mejor y por qué?