Métodos Numéricos 1 (MN1)

Unidade 1: Teoria dos Erros
Parte 4: Propagação de Erros Numéricos

Joaquim Bento Cavalcante Neto

joaquimb@lia.ufc.br

Grupo de Computação Gráfica, Realidade Virtual e Animação (CRAb)

Departamento de Computação (DC)

Universidade Federal do Ceará (UFC)

CRAb

UFC

Erros nas Parcelas

- Veremos a seguir como calcular os erros absoluto e relativo nas operações aritméticas com erros nas parcelas ou fatores usados
- A princípio, consideraremos só o erro de arredondamento/truncamento nas parcelas
- Sejam x e y, tais que:
  - $x = \bar{x} + Ea_x$
  - $y = \bar{y} + EA_y$

Adição: x+y

- Cálculo de x + y:

$$x + y = (\bar{x} + EA_x) + (\bar{y} + EA_y) = (\bar{x} + \bar{y}) + (EA_x + EA_y)$$

- Então, o erro absoluto na soma, EA_{x+y} é a soma dos erros absolutos das parcelas:

$$EA_{x+y} = EA_x + EA_y$$

- E o erro relativo será:

$$\begin{aligned} ER_{x+y} &= \frac{EA_{x+y}}{\bar{x}+\bar{y}} = \frac{EA_x}{\bar{x}}\left(\frac{\bar{x}}{\bar{x}+\bar{y}}\right) + \frac{EA_y}{\bar{y}}\left(\frac{\bar{y}}{\bar{x}+\bar{y}}\right) \\ &= ER_x\left(\frac{\bar{x}}{\bar{x}+\bar{y}}\right) + ER_y\left(\frac{\bar{y}}{\bar{x}+\bar{y}}\right) \end{aligned}$$

Subtração: x-y

- Cálculo de x - y:

$$x - y = (\bar{x} + EA_x) - (\bar{y} + EA_y) = (\bar{x} - \bar{y}) + (EA_x - EA_y)$$

- Então, o erro absoluto na diferença, EA_{x-y} é a diferença dos erros absolutos das parcelas:

$$EA_{x-y} = EA_x - EA_y$$

- E o erro relativo será:

$$\begin{aligned} ER_{x-y} &= \frac{EA_{x-y}}{\bar{x}-\bar{y}} = \frac{EA_x}{\bar{x}}\left(\frac{\bar{x}}{\bar{x}-\bar{y}}\right) - \frac{EA_y}{\bar{y}}\left(\frac{\bar{y}}{\bar{x}-\bar{y}}\right) \\ &= ER_x\left(\frac{\bar{x}}{\bar{x}-\bar{y}}\right) - ER_y\left(\frac{\bar{y}}{\bar{x}-\bar{y}}\right) \end{aligned}$$

Multiplicação: xy

- Erros absoluto e relativo:

$$
\begin{aligned}
xy &= (\bar{x} + EA_x)(\bar{y} + EA_y) \\
&= \bar{x}\,\bar{y} + \bar{x}\,EA_y + \bar{y}\,EA_x + \cancelto{\approx 0}{(EA_x)(EA_y)}
\end{aligned}
$$

$$EA_{x\,y} \approx \bar{x}\,EA_y + \bar{y}\,EA_x$$

$$ER_{x\,y} \approx \frac{\bar{x}\,EA_y + \bar{y}\,EA_x}{\bar{x}\,\bar{y}} = \frac{EA_x}{\bar{x}} + \frac{EA_y}{\bar{y}} = ER_x + ER_y$$

Divisão: x/y

- Cálculo de x / y:

$$\frac{x}{y} = \frac{\bar{x} + EA_x}{\bar{y} + EA_y} = \frac{\bar{x} + EA_x}{\bar{y}} \left(\frac{1}{1 + \frac{EA_y}{\bar{y}}} \right)$$

- Usando a série de Taylor para representar o termo $\dfrac{1}{1 + \frac{EA_y}{\bar{y}}}$ como uma série infinita, temos:

$$\frac{1}{1 + \frac{EA_y}{\bar{y}}} = 1 - \frac{EA_y}{\bar{y}} + \cancelto{\approx 0}{\left(\frac{EA_y}{\bar{y}} \right)^2 - \left(\frac{EA_y}{\bar{y}} \right)^3} + \ldots$$

Desprezaremos as potências maiores que 1

Divisão: x/y

- Erros absoluto e relativo:

$$\frac{x}{y} \approx \frac{\bar{x}+EA_x}{\bar{y}}\left(1-\frac{EA_y}{\bar{y}}\right) = \frac{\bar{x}}{\bar{y}} + \frac{EA_x}{\bar{y}} - \frac{\bar{x}\,EA_y}{\bar{y}^2} - \underbrace{\cancel{\frac{EA_x\,EA_y}{\bar{y}^2}}}_{\approx 0}$$

$$EA_{\frac{x}{y}} \approx \frac{EA_x}{\bar{y}} - \frac{\bar{x}\,EA_y}{\bar{y}^2} = \frac{\bar{y}\,EA_x - \bar{x}\,EA_y}{\bar{y}^2}$$

$$ER_{\frac{x}{y}} \approx \left(\frac{\bar{y}\,EA_x - \bar{x}\,EA_y}{\bar{y}^2}\right)\frac{\bar{y}}{\bar{x}} = \frac{EA_x}{\bar{x}} - \frac{EA_y}{\bar{y}} = ER_x - ER_y$$

Erro Completo

- A análise de erros só é completa quando consideramos, além dos erros nas parcelas ou fatores utilizados, também os erros de arredondamento/truncamento no resultado de cada operação efetuada uma a cada vez
- Exemplo: Supondo que x, y, z e t sejam representados exatamente (ER=0), qual erro total no cálculo de $u = (x + y) z - t$?
  - Será calculado erro relativo e RA será o erro relativo de arredondamento de cada operação

Exemplo: u = (x+y)z-t

- Seja s = x+y

$$ER_s = ER_x\left(\frac{\bar{x}}{\bar{x}+\bar{y}}\right) + ER_y\left(\frac{\bar{y}}{\bar{x}+\bar{y}}\right) + RA = 0 + RA$$

- Calculando m = s x z

$$|ER_s| = |RA| < \frac{1}{2} \times 10^{-t+1}$$

$$ER_m = ER_s + ER_z + RA = ER_s + \frac{EA_z}{\bar{z}} + RA = RA_s + 0 + RA$$

$$|ER_m| \leq |RA_s| + |RA| < \frac{1}{2} \times 10^{-t+1} + \frac{1}{2} \times 10^{-t+1} = 10^{-t+1}$$

Exemplo: u = (x+y)z-t

- Seja u = m - t

$$
\begin{aligned}
ER_u &= \frac{EA_m - EA_t}{\bar{m}-\bar{t}} + RA = \frac{EA_m}{\bar{m}-\bar{t}} + RA = \frac{EA_m}{\bar{m}}\left(\frac{\bar{m}}{\bar{m}-\bar{t}}\right) + RA \\
&= ER_m\left(\frac{\bar{m}}{\bar{m}-\bar{t}}\right) + RA
\end{aligned}
$$

- Então

$$
|ER_u| \leq |ER_m|\left|\frac{\bar{m}}{\bar{m}-\bar{t}}\right| + RA < 10^{-t+1}\left|\frac{\bar{m}}{\bar{m}-\bar{t}}\right| + \frac{1}{2}\times 10^{-t+1}
$$

$$
|ER_u| < \left(\frac{|\bar{m}|}{|\bar{m}-\bar{t}|} + \frac{1}{2}\right)\times 10^{-t+1}
$$

Exemplo

- Dado um sistema SPF(10,4,-5,5) que usa truncamento, calcule o erro absoluto do somatório abaixo em cada iteração:

$$\sum_{i=1}^{4} x_i + y_i, sendo\ x_i = 0.46709\ e\ y_i = 3.5678$$

- i=1
  - $S_1 = x_1+y_1 = 0.046709 \times 10^1 + 0.35678 \times 10^1 = 0.403489 \times 10^1$
  - $\overline{S_1} = \overline{x}_1+\overline{y}_1 = 0.0467 \times 10^1 + 0.3567 \times 10^1 = 0.4034 \times 10^1$
  - $|EA_{S1}| = |0.403489 \times 10^1 - 0.4034 \times 10^1| = 0.89 \times 10^{-3}$

Exemplo

- Dado um sistema SPF(10,4,-5,5) que usa truncamento, calcule o erro absoluto do somatório abaixo em cada iteração:

$$\sum_{i=1}^{4} x_i + y_i, sendo\ x_i = 0.46709\ e\ y_i = 3.5678$$

- i=2
  - $S_2 = x_1+y_1 + x_2+y_2 = 0.806978 \times 10^1$
  - $\overline{S_2} = \overline{x}_1+\overline{y}_1 + \overline{x}_2+\overline{y}_2 = 0.8068 \times 10^1$
  - $|EA_{S2}| = |0.806978 \times 10^1 - 0.8068 \times 10^1| = 0.178 \times 10^{-2}$

Exemplo

- Dado um sistema SPF(10,4,-5,5) que usa truncamento, calcule o erro absoluto do somatório abaixo em cada iteração:

$$\sum_{i=1}^{4} x_i + y_i, sendo\ x_i = 0.46709\ e\ y_i = 3.5678$$

  - i=3
    - $S_3 = x_1+y_1 + x_2+y_2 + x_3+y_3 = 0.1210467 \times 10^2$
    - $\overline{S_3} = \overline{x_1}+\overline{y_1} + \overline{x_2}+\overline{y_2} + \overline{x_3}+\overline{y_3} = 0.1210 \times 10^2$
    - $|EA_{S3}| = |\ 0.1210467 \times 10^2 - 0.1210 \times 10^2| = 0.467 \times 10^{-2}$

Exemplo

- Dado um sistema SPF(10,4,-5,5) que usa truncamento, calcule o erro absoluto do somatório abaixo em cada iteração:

$$\sum_{i=1}^{4} x_i + y_i, sendo\ x_i = 0.46709\ e\ y_i = 3.5678$$

- i=4
  - $S_4 = x_1+y_1 + x_2+y_2 + x_3+y_3 + x_4+y_4 = 0.1613956 \times 10^2$
  - $\overline{S_4} = \overline{x}_1+\overline{y}_1 + \overline{x}_2+\overline{y}_2 + \overline{x}_3+\overline{y}_3 + \overline{x}_4+\overline{y}_4 = 0.1613 \times 10^2$
  - $|EA_{S4}| = |\ 0.1613956 \times 10^2 - 0.1613 \times 10^2| = 0.956 \times 10^{-2}$

Cancelamento Subtrativo

- Dado um sistema SPF(10,4,-5,5) que usa o truncamento, calcular $z = x - y$ onde:

$$x = 0.235789 \text{ e } y = 0.235534$$

  - $\bar{x} = 0.2357$ e $\bar{y} = 0.2355$
  - $\bar{z} = 0.0002 = 0.2 \times 10^{-3}$
  - $|ER_z| = EA_z / \bar{z} = (EA_x - EA_y) / \bar{z} = (0.000089-0.000034)/0.0002 = 0.275$

- Quando $x \approx y$, a subtração causa perda de dígitos significantes (nesse exemplo, 3), o erro relativo pode ser grande e isto é então denominado de cancelamento subtrativo

Cancelamento Subtrativo

- Com t=8, sejam $a = e^{-\left(\frac{1}{100}\right)^2} = 0.99990001$ e $b = e^{-\left(\frac{1}{1000}\right)^2} = 0.99999900$ calcular c = a - b:
  - c = a - b = -0.00009899 ⇒ 4 dígitos significativos
- Se começássemos com aproximações mais precisas, então a diferença conteria mais dígitos significantes, como, usando t =16:

a = 0.9999000049998333
b = 0.9999990000005000
c = -0.0000989950006667

- c agora possui 12 dígitos significativos

Observação

- Se mudando a precisão de um cálculo alterar dramaticamente os resultados, então é quase certo que a computação seja seriamente afetada por erros de arredondamento ou de truncamento

- Deve-se ter cuidado com resultados!!!

Outras Limitações

- Algumas propriedades consagradas no conjunto dos números reais podem não ser verdadeiras, na exatidão da representação
  - Propriedades comutativa e associativa na adição
  - Propriedades comutativa e distributiva na multiplicação
- Exemplo: Dados $x, y, z \in \mathbb{R}$ e o SPF (3,2,-1,2):
  - $x = 5/3 = (0.12)_3 \times 3^1$
  - $y = 7/27 = (0.21)_3 \times 3^{-1}$
  - $z = 8/9 = (0.22)_3 \times 3^0$
  - $x + (y+z) = (x+y) + z$?

$x+(y+z) = 0.22 \times 3^1$

$(x+y)+z = 0.21 \times 3^1$