ENG10128 - Matemática Discreta	10/03/2011
Aula 3: Estimativa e medição de erros	
Prof.: Paulo Roberto Nunes de Souza	

8 Erros absolutos e relativos

Para se obter o erro em determinado valor utiliza-se o valor exato x e o valor aproximado $\overline{x}$. Com o conhecimento desses dois valores podemos calcular o erro absoluto EA_x e o erro relativo ER_x.

O erro absoluto EA_x é definido como:

$$EA_x = x - \overline{x}$$

Entretanto o grande problema surge do fato que não sabemos o valor exato da variável. Se este valor fosse conhecido, certamente o erro não existiria. Por este motivo, não nos preocuparemos com o valor exato do erro mas em um limitante superior deste erro.

Por exemplo, nós não temos a capacidade de representar o valor de π exatamente, pois o número possui uma dísima infinita não periódica. Porém, se considerarmos que $\pi \in (3,14; 3,15)$ poderemos estimar que $\mid EA_\pi \mid = \mid \pi - \overline{\pi} \mid < 0,01$

Considerando um outro exemplo, onde para um número x, $\overline{x} = 2112,9$ e $\mid EA_x \mid < 0,1 \mid$, ou seja, $x \in (2112,8; 2113,0)$ e para um número y, $\overline{y} = 5,3$ e $\mid EA_y \mid < 0,1 \mid$, ou seja, $y \in (5,2; 5,4)$. Os limitantes superiores para os erros absolutos são os mesmos. Podemos dizer que ambos os números estão representados com a mesma precisão?

Para se representar a precisão de uma medida é necessário comparar o erro absoluto com a ordem de grandeza do valor medido, para isto utiliza-se o erro relativo.

O erro relativo ER_x é definido como:

$$ER_x = \frac{EA_x}{\overline{x}} = \frac{x - \overline{x}}{\overline{x}}$$

No último exemplo teriamos:

$$\mid ER_x \mid = \frac{\mid EA_x \mid}{\mid \overline{x} \mid} = \frac{0,1}{2112,9} \approx 4,7 \times 10^{-5}$$

e

$$\mid ER_y \mid = \frac{\mid EA_y \mid}{\mid \overline{y} \mid} = \frac{0,1}{5,3} \approx 0,02$$

, portanto a representação de x tem mais precisão que a representação de y.

9 Erros de arredondamento e de truncamento

A representação de um número no sistema de pontos flutuantes depende de alguns parâmetros associados à máquina e às estruturas de dados utilizadas, tais como a base numérica adotada, o total de dígitos da mantissa, etc.

Seja um sistema utilizando aritmética de ponto flutuante de t dígitos na mantissa utilizando a base 10, e seja x, escrito na forma

$$x = \overbrace{f_x \times 10^e}^{\text{representável}} + \overbrace{g_x \times 10^{e-t}}^{\text{não representável}} \quad \text{onde } 0,1 \leq f_x < 1 \text{ e } 0 \leq g_x < 1$$

Por exemplo, se $t = 4$ e $x = 234,57$, então

$$x = 0,2345 \times 10^3 + 0,7 \times 10^{-1}, \text{ onde } f_x = 0,2345 \text{ e } g_x = 0,7$$

Desta forma ficam separadas a parte do número que pode ser representada $0,2345 \times 10^3$ e a parte que não pode ser representada $0,7 \times 10^{-1}$. Desta forma podemos calcular o limitante do erro nesta representação.

Se for utilizado o truncamento, ou seja, a parte não representável é simplesmente descartada, teremos que:

$$| EA_x |=| x - \overline{x} |=| f_x \times 10^e + g_x \times 10^{e-t} - f_x \times 10^e |=| g_x | \times 10^{e-t} < 10^{e-t}, \text{ visto que } | g_x |< 1$$

e

$$| ER_x |= \frac{| EA_x |}{| \overline{x} |} = \frac{| g_x | \times 10^{e-t}}{| f_x | \times 10^e} < \frac{10^{e-t}}{0,1 \times 10^e} = 10^{-t+1}, \text{ visto que } 0,1 \text{ é o menor valor possível para } f_x.$$

Se for utilizado o arredondamento na sua forma mais comum, teremos que:

$$\overline{x} = \begin{cases} f_x \times 10^e & \text{se } | g_x |< \frac{1}{2} \\ f_x \times 10^e + 10^{e-t} & \text{se } | g_x |\geq \frac{1}{2} \end{cases}$$

Portanto se $| g_x |< \frac{1}{2}$, g_x é desprezado, caso contrário, somamos 1 ao último dígito de f_x.

Então, se $| g_x |< \frac{1}{2}$ teremos

$$| EA_x |=| x - \overline{x} |=| f_x \times 10^e + g_x \times 10^{e-t} - f_x \times 10^e |=| g_x | \times 10^{e-t} < \frac{1}{2} \times 10^{e-t}$$

e

$$| ER_x |= \frac{| EA_x |}{| \overline{x} |} = \frac{| g_x | \times 10^{e-t}}{| f_x | \times 10^e} < \frac{0,5 \times 10^{e-t}}{0,1 \times 10^e} = \frac{1}{2} \times 10^{-t+1}.$$

E se $| g_x |\geq \frac{1}{2}$ teremos

$$| EA_x |=| x - \overline{x} |=| (f_x \times 10^e + g_x \times 10^{e-t}) - (f_x \times 10^e + 10^{e-t}) |$$

$$=| g_x \times 10^{e-t} - 10^{e-t} |=| (g_x - 1) | \times 10^{e-t} \leq \frac{1}{2} \times 10^{e-t}$$

e

$$| ER_x |= \frac{| EA_x |}{| \overline{x} |} \leq \frac{0,5 \times 10^{e-t}}{| f_x \times 10^e + 10^{e-t} |} < \frac{0,5 \times 10^{e-t}}{| f_x | \times 10^e} < \frac{0,5 \times 10^{e-t}}{0,1 \times 10^e} = \frac{1}{2} \times 10^{-t+1}.$$

Portanto, em qualquer dos casos teremos

$$| EA_x |\leq \frac{1}{2} \times 10^{e-t} \text{ e } | ER_x |< \frac{1}{2} \times 10^{-t+1}.$$

Apesar de incorrer em erros menores, o uso do arredondamento acarreta um tempo maior de execução e por esta razão o truncamento é mais utilizado.

10 Erros nas operações aritméticas de ponto flutuante

Além do erro inerente à representação de cada valor é importante entender como esses erro podem se propagar ou se acumular quando é feita uma operação aritmética. O erro total em uma operação é composto perro erro das parcelas ou fatores e pelo erro no resultado da operação.

Nos exemplos a seguir, iremos supor que as operações são efetuadas num sistema de 4 dígitos na mantissa utilizando a base 10 e com acumuladores de precisão dupla.

Dados $x = 0,937 \times 10^4$ e $y = 0,1272 \times 10^2$, obter $x + y$.

$$x + y = 0,937 \times 10^4 + 0,1272 \times 10^2 = 0,937 \times 10^4 + 0,001272 \times 10^4 = 0,938272 \times 10^4$$

Este é o resultado exato da operação $x+y = 0,938272 \times 10^4$, porém não é possível representar exatamente este valor no nosso sistema, portanto o valor a ser representado será $\overline{x+y} = 0,9382 \times 10^4$, se for utilizado o truncamento, ou $\overline{x+y} = 0,9383 \times 10^4$, se for utilizado o arredondamento.

Considerando um outro exemplo onde x e y são os mesmos do exemplo anterior, mas pretende-se obter xy:

$$xy = (0,937 \times 10^4) \times (0,1272 \times 10^2) = (0,937 \times 0,1272) \times 10^6 = 0,1191864 \times 10^6$$

Logo, $\overline{xy} = 0,1191 \times 10^6$, se for utilizado o truncamento, ou $\overline{xy} = 0,1192 \times 10^6$, se for utilizado o arredondamento.

Os dois exemplos mostram que, mesmo considerando que os valores estão representados sem erro, não se pode esperar que o resultado da operação seja isento de erro.

Utilizando raciocínio análogo ao da seção anterior pode-se obter que os erros decorrentes de uma operação serão:

$$\mid EA_{OP} \mid \leq 10^{e-t} \text{ e } \mid ER_{OP} \mid < 10^{-t+1}, \text{ no truncamento e}$$

$$\mid EA_{OP} \mid \leq \frac{1}{2} \times 10^{e-t} \text{ e } \mid ER_{OP} \mid < \frac{1}{2} \times 10^{-t+1}, \text{ no arredondamento.}$$

Veremos a seguir as fórmulas para os erros nas operações aritméticas, considerando apenas a propagação do erro nas parcelas ou fatores.

Considerando x e y, tais que $x = \overline{x} + EA_x$ e $y = \overline{y} + EA_y$.

10.1 Adição

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

Logo o erro absoluto na soma é

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

O erro relativo será o seguinte:

$$ER_{x+y} = \frac{EA_{x+y}}{\overline{x} + \overline{y}} = \frac{EA_x + EA_y}{\overline{x} + \overline{y}} = \frac{EA_x}{\overline{x}} \left(\frac{\overline{x}}{\overline{x} + \overline{y}} \right) + \frac{EA_y}{\overline{y}} \left(\frac{\overline{y}}{\overline{x} + \overline{y}} \right)$$

$$ER_{x+y} = ER_x \left(\frac{\overline{x}}{\overline{x} + \overline{y}} \right) + ER_y \left(\frac{\overline{y}}{\overline{x} + \overline{y}} \right)$$

10.2 Subtração

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

Logo o erro absoluto na subtração é

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

O erro relativo será o seguinte:

$$ER_{x-y} = \frac{EA_x - EA_y}{\overline{x} - \overline{y}} = ER_x \left(\frac{\overline{x}}{\overline{x} - \overline{y}} \right) + ER_y \left(\frac{\overline{y}}{\overline{x} - \overline{y}} \right)$$

10.3 Multiplicação

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

Considerando que $(EA_x)(EA_y)$ é um número muito pequeno, podemos desprezar este termo. Logo o erro absoluto na multiplicação é

$$EA_{xy} \approx \overline{x}(EA_y) + \overline{y}(EA_x).$$

O erro relativo será o seguinte:

$$ER_{xy} \approx \frac{\overline{x}(EA_y) + \overline{y}(EA_x)}{\overline{x} - \overline{y}} = \frac{EA_x}{\overline{x}} + \frac{EA_y}{\overline{y}} = ER_x + ER_y$$

10.4 Divisão

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

Representando o fator $\dfrac{1}{1 + \dfrac{EA_y}{\overline{y}}}$ sob a forma de uma série infinita, teremos

$$\frac{1}{1 + \dfrac{EA_y}{\overline{y}}} = 1 - \frac{EA_y}{\overline{y}} + \left(\frac{EA_y}{\overline{y}}\right)^2 - \left(\frac{EA_y}{\overline{y}}\right)^3 + \ldots$$

, despresando os termos da série com potências maiores que 1, teremos

$$\frac{x}{y} \approx \frac{\overline{x} + EA_x}{\overline{y}} \left(1 - \frac{EA_y}{\overline{y}}\right) = \frac{\overline{x}}{\overline{y}} + \frac{ER_x}{\overline{y}} - \frac{\overline{x}EA_y}{\overline{y}^2} - \frac{EA_xEA_y}{\overline{y}^2}$$

Considerando novamente que $(EA_x)(EA_y)$ é um número muito pequeno, podemos desprezar este termo. Com isso

$$\frac{x}{y} \approx \frac{\overline{x}}{\overline{y}} + \frac{ER_x}{\overline{y}} - \frac{\overline{x}EA_y}{\overline{y}^2}$$

Logo o erro absoluto na divisão é

$$EA_{x/y} \approx \frac{ER_x}{\overline{y}} - \frac{\overline{x}EA_y}{\overline{y}^2} = \frac{\overline{y}ER_x - \overline{x}EA_y}{\overline{y}^2}.$$

O erro relativo será o seguinte:

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