CÁLCULO NUMÉRICO

Profa. Dra. Yara de Souza Tadano *yaratadano@utfpr.edu.br*

Aulas 5 e 6

03/2014

Erros

Aritmética no Computador

- A aritmética executada por uma calculadora ou computador é diferente daquela dos cursos de álgebra e cálculo.

- Na **matemática tradicional** podemos representar números com infinitos dígitos.
  - **Ex.** $\sqrt{2} = 1,414213562\cdots$

- **Computacionalmente**, a representação é finita.
  - **Ex.** $\sqrt{2} = 1,41421$

Arredondamento x Truncamento

- As discrepâncias introduzidas pela representação finita dos números reais é denominada **erros de arredondamento**.
- Eles ocorrem quando números com uma **quantidade limitada** de algarismos significativos são usados para representar números exatos.

- Os **erros de truncamento** resultam de aproximações para representar procedimentos matemáticos exatos.
- Está associado ao método de aproximação empregado, como vimos quando fazemos aproximações usando **polinômios de *Taylor***.

Erro Absoluto

Valor verdadeiro = aproximação + erro absoluto

- O erro numérico é igual à discrepância entre o valor verdadeiro e a aproximação, que é chamado **erro absoluto**.

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

- onde, x é o valor verdadeiro e $\overline{x}$ é o valor aproximado.

- **PROBLEMA**: não leva em conta a ordem de grandeza.

Erro Relativo

- O **erro relativo** é a razão entre o erro absoluto e o valor verdadeiro:

$$\boxed{ER_x = \frac{x - \bar{x}}{\bar{x}}}$$

- O **erro relativo percentual** é dado por:

$$\boxed{\varepsilon_t = \frac{x - \bar{x}}{\bar{x}} \cdot 100\%}$$

Exemplo 1

- Suponha que você tenha a tarefa de medir os comprimentos de uma ponte e de um parafuso e obteve as medidas 9.999 cm e 9 cm, respectivamente. Se os valores verdadeiros forem 10.000 cm e 10 cm, respectivamente, calcule para cada caso:

- a) O erro absoluto verdadeiro (EA_x);
- b) O erro relativo percentual verdadeiro (ε_t).

Aproximação para o erro

- Para os métodos numéricos, o **valor verdadeiro** será conhecido **apenas** ao se lidar com funções que podem ser resolvidas **analiticamente**.
- Nas aplicações do mundo real, **não** conhecemos a resposta verdadeira. Nestes casos, encontramos um **limitante** para o erro, o que fornece "**o pior caso**" de erro.

$$\varepsilon_a = \frac{\textit{erro aproximado}}{\textit{aproximação}} \cdot 100\%$$

Aproximação para o erro

- Nos **métodos iterativos**, uma aproximação atual é feita com base em uma aproximação prévia. Esse processo é realizado repetidamente (iterativamente) para se calcular aproximações cada vez melhores.

- Nestes casos, o **erro relativo percentual** é determinado por:

$$\varepsilon_a = \frac{aproximação\ atual - aproximação\ prévia}{aproximação\ atual} \cdot 100\%$$

- A grande preocupação é em saber se o valor absoluto percentual é **menor** que uma tolerância percentual pré-estabelecida ε_s .

$$\boxed{\left|\varepsilon_a\right| < \varepsilon_s}$$

- Repete-se, então, os cálculos até que isto ocorra.

- É importante, também, relacionar esses erros ao número de algarismos significativos na aproximação. Pode ser mostrado que:

$$\boxed{\varepsilon_s = \left(0,5 \times 10^{2-n}\right)\%}$$

- Isto indica que o resultado é correto até pelo menos *n* algarismos significativos.

Exemplo 2

- Use expansões em série de *Taylor* com $x_0 = 0$ (em **série de *Maclaurin***) para aproximar $f(x) = e^x$.
- Começando com a versão mais simples $e^x = 1$, some um termo de cada vez para estimar $e^{0,5}$. Depois que cada termo for adicionado, calcule o **erro verdadeiro** e o **erro relativo percentual aproximado**. Observe que o valor verdadeiro é $e^{0,5} = 1,648721...$
- Adicione termos até que o valor absoluto do erro estimado aproximado ε_a esteja dentro do critério de erro pré-estabelecido ε_s que garanta **três algarismos significativos**.

Exemplo 2

- **Polinômio de Taylor**

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

Exemplo 2 - Resultado

Termos	Resultados	ε_t (%)	ε_a (%)
1	1	39,3	
2	1,5	9,91	33,3
3	1,625	1,46	7,69
4	1,645833	0,175	1,27
5	1,648438	0,0172	0,158
6	1,648698	0,00140	0,0159

ARITMÉTICA DE PONTO FLUTUANTE

Arredondamento e Truncamento

- Seja um sistema que opera em aritmética de ponto flutuante de k dígitos na base 10, e seja x, escrito na forma:

$$\boxed{x = f_x \times 10^{e} + g_x \times 10^{e-k}}$$

onde: $0,1 \le f_x \le 1$ e $0 \le g_x < 1$.

- **EXEMPLO**: Se $k = 4$ e $x = 234,57$, então:

$$x = 0,2345 \times 10^{3} + 0,7 \times 10^{-1}$$

onde: $f_x = 0,2345$ e $g_x = 0,7$.

Arredondamento e Truncamento

- Na representação de x nesse sistema $g_x \times 10^{e-k}$ **não** pode ser incorporado **totalmente** à mantissa. Então surge a questão de **como considerar** esta parcela na mantissa e definir o erro absoluto ou relativo máximo cometido.

Truncamento

- $g_x \times 10^{e-k}$ é desprezado e $\bar{x} = f_x \times 10^e$. Neste caso:

$$\boxed{|EA_x| = |x - \bar{x}| = |g_x| \times 10^{e-k} < 10^{e-k}}$$

- Visto que $|g_x| < 1$, e:

$$\boxed{|ER_x| = \frac{|EA_x|}{|\bar{x}|} = \frac{|g_x| \times 10^{e-k}}{|f_x| \times 10^e} < \frac{10^{e-k}}{0{,}1 \times 10^e} = 10^{-k+1}}$$

- Visto que 0,1 é o menor valor possível para f_x.

Arredondamento

- f_x é modificado para levar em conta g_x. A forma mais comum é o arredondamento simétrico.

$$\bar{x} = \begin{cases} f_x \times 10^e & , \quad se\ |g_x| < 0,5 \\ f_x \times 10^e + 10^{e-k} & , \quad se\ |g_x| \geq 0,5 \end{cases}$$

- Portanto, se $|g_x| < 0,5$, g_x é desprezado; caso contrário, somamos o número *1* ao último dígito de f_x .

Arredondamento

- Então, se $|g_x| < 0,5$:

$$\boxed{|EA_x| = |x - \bar{x}| = |g_x| \times 10^{e-k} < 0,5 \times 10^{e-k}}$$

e

$$\boxed{|ER_x| = \frac{|EA_x|}{|\bar{x}|} = \frac{|g_x| \times 10^{e-k}}{|f_x| \times 10^{e}} < \frac{0,5 \times 10^{e-k}}{0,1 \times 10^{e}} = 0,5 \times 10^{-k+1}}$$

Arredondamento

- Agora, se $|g_x| \geq 0,5$:

$$|EA_x| = |x - \bar{x}| = \left|\left(f_x \times 10^e + g_x \times 10^{e-k}\right) - \left(f_x \times 10^e + 10^{e-k}\right)\right|$$
$$= \left|g_x \times 10^{e-k} - 10^{e-k}\right| = \left|\left(g_x - 1\right) \times 10^{e-k}\right|$$
$$< 0,5 \times 10^{e-k}$$

e

$$|ER_x| \leq \frac{0,5 \times 10^{e-k}}{\left|f_x \times 10^e + 10^{e-k}\right|} < \frac{0,5 \times 10^{e-k}}{|f_x| \times 10^e} < \frac{0,5 \times 10^{e-k}}{0,1 \times 10^e} = 0,5 \times 10^{-k+1}$$

Arredondamento

- Portanto, em qualquer caso, teremos:

$$\boxed{\left|EA_x\right| < 0,5 \times 10^{e-k}} \quad \text{e} \quad \boxed{\left|ER_x\right| < 0,5 \times 10^{-k+1}}$$

- Apesar do uso de **arredondamento** implicar **erros menores**, exige **maior tempo** de execução e, portanto, **o truncamento** é mais **utilizado**.

ANÁLISE DE ERROS NAS OPERAÇÕES

- Dada uma sequência de operações, como, por exemplo, $u = [\ (x + y) - z - t\ ] / \mathrm{w}$, é **importante** a noção de como o **erro** em uma operação **propaga-se** ao longo das operações subsequentes.

- O **erro total** em uma operação é composto pelo erro das parcelas ou fatores e pelo erro no resultado da operação.

ADIÇÃO

- A **adição** em aritmética de ponto flutuante requer o alinhamento dos pontos decimais dos dois números.
- Para isto, a **mantissa** do número de **menor expoente** deve ser deslocada para a **direita**.
- Este **deslocamento** deve ser um número de casas decimais igual à **diferença entre os dois expoentes**.

- **Observe:** Ainda que as **parcelas ou fatores** de uma operação estejam representados **exatamente** no sistema, **não** se pode esperar que o **resultado** armazenado seja **exato**.

- Na maioria dos sistemas, o resultado exato da operação (denotado por OP) é **normalizado** e, em seguida, arredondado ou truncado para k dígitos, obtendo assim, o resultado aproximado (denotado por $\overline{OP}$) que é armazenado na máquina.
- Então, o **erro relativo** de uma operação (supondo que as parcelas ou fatores estão representados exatamente) será:
- **Truncamento:** $\boxed{|ER_{OP}| < 10^{-k+1}}$
- **Arredondamento:** $\boxed{|ER_{OP}| < 0,5 \times 10^{-k+1}}$

Erro nas operações aritméticas

- Veremos as fórmulas para os erros absoluto e relativo nas operações aritméticas.

- Vamos supor que o erro final é arredondado.

- Sejam x e y, tais que $x = \bar{x} + EA_x$ e $y = \bar{y} + EA_y$:

Erro nas operações aritméticas

- **ADIÇÃO:**
  - Erro absoluto:

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

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

  - Erro relativo:

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

Erro nas operações aritméticas

- **SUBTRAÇÃO:**
  - Erro absoluto:

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

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

  - Erro relativo:

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

Erro nas operações aritméticas

- **MULTIPLICAÇÃO:**
  - Erro absoluto:

$$x \cdot y = \left(\bar{x} + EA_x\right) \cdot \left(\bar{y} + EA_y\right) = \left(\bar{x} \cdot \bar{y}\right) + \bar{y} \cdot EA_x + \bar{x} \cdot EA_y + \underbrace{\left(EA_x \cdot EA_y\right)}_{\text{muito pequeno}}$$

$$\boxed{EA_{xy} = \bar{x}\ EA_y + \bar{y}\ EA_x}$$

  - Erro relativo:

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

Erro nas operações aritméticas

- **DIVISÃO:**
  - Erro absoluto:

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

Simplificação:

$$\frac{1}{1 + \frac{EA_y}{\bar{y}}} = 1 - \frac{EA_y}{\bar{y}} + \left(\frac{EA_y}{\bar{y}} \right)^2 - \left(\frac{EA_y}{\bar{y}} \right)^3 + \cdots$$

Desprezam-se os termos de potência >1

$$\frac{x}{y} \approx \frac{\bar{x}}{\bar{y}} + \frac{EA_x}{\bar{y}} - \frac{\bar{x} \cdot EA_y}{\bar{y}^2}$$

$$\boxed{EA_{x/y} \approx \frac{\bar{y}\ EA_x - \bar{x}\ EA_y}{\bar{y}^2}}$$

  - Erro relativo:

$$\boxed{ER_{x/y} = ER_x - ER_y}$$

Exemplo

- Suponha que x, y, z e t estejam representados exatamente, qual o erro total do cálculo de $u = (x + y)\, z - t$?
- Calcularemos o erro relativo e denotaremos por RA, o erro relativo de arredondamento no resultado da operação.