

Tópico 02: Aritmética de ponto flutuante, erros e efeitos numéricos

Professor: Ranniery Maia

Departamento de Informática e Matemática Aplicada - UFRN
Disciplina: DIM0404 - Cálculo Numérico para Ciência da Computação
Semestre: 2024.1

Natal, 20 de março de 2024.

Sumário

- 1 Plano de aula
- 2 Aritmética de ponto flutuante
- 3 Erros
- 4 Efeitos numéricos
- 5 Laboratório
- 6 Resumo

- **Objetivos.**

- Entender como arredondamento é formalmente feito.
- Entender como erros podem afetar resultados numéricos.

- **Objetivos específicos:** ao final o aluno deverá estar apto a:

- entender acerca de arredondamento e truncamento.
- Entender a diferença entre erros absoluto e relativo.
- Entender e analisar erros envolvidos nos processos numéricos.

- **Metodologia.**

- Exposição do assunto usando slides e quadro branco.
- Resolução de exercícios em sala.
- Laboratório: elaboração de algoritmos e codificação em *Python*.

- **Avaliação.**

- Participação do aluno.

- **Bibliografia:**

[Franco, 2007, Ruggiero and da Rocha Lopes, 1996]

Definição de arredondamento [Franco, 2007]: arredondar um número x , por outro com um número menor de dígitos significativos, consiste em encontrar um número \bar{x} , pertencente ao sistema de numeração, tal que $|\bar{x} - x|$ seja o menor possível.

Procedimento para arredondar um número

- Dado x e seja \bar{x} a sua versão arredondada em $F(\beta, t, m, M)$.
- Assumindo que no formato ponto flutuante o mesmo seja

$$x = \pm s \times \beta^e \quad (1)$$

as seguintes condições são necessárias para obter \bar{x} :

- 1 Para a mantissa s :

$$\beta^{-1} \left(1 - \frac{1}{2} \beta^{-t} \right) \leq s < 1 - \frac{1}{2} \beta^{-t} \quad (2)$$

- 2 Para o coeficiente e : $e \in [-m, M]$

- Se as condições acima forem satisfeitas, então:

- 1 Obtemos a nova mantissa s' como

$$s' = s + \frac{1}{2} \beta^{-t} = 0.d_1 d_2 \cdots d_t d_{t+1} \cdots \quad (3)$$

- 2 O número arredondado será.

$$\bar{x} = \{\text{ sinal de } x\} (0.d_1 d_2 \cdots d_t) \times \beta^e \quad (4)$$

Arredondamento: exercícios de sala

Considere o sistema $F(10, 3, 5, 5)$. Represente neste sistema os números abaixo.

- a $x = 1234.56$
- b $x = -0.00054962$
- c $x = 0.9995$
- d $x = 123456.7$
- e $x = -0.0000001$

Operações aritméticas em ponto flutuante

- Devido aos efeitos de arredondamento a ser feito **após cada operação** em um computador, as operações aritméticas não são nem associativas nem distributivas.
 - 1 $(x + y) + z \neq x + (y + z)$
 - 2 $\frac{xy}{z} \neq \left(\frac{x}{z}\right)y$
- Consequentemente, erros consideráveis podem ocorrer durante a execução de um algoritmo.
- Podemos obter resultados diferentes mesmo utilizando métodos numéricos matematicamente equivalentes.
- Devemos desenvolver algoritmos que levem em conta esses efeitos.

Aritmética de ponto flutuante: exercícios de sala

- 1 Vamos efetuar as seguintes operações considerando o sistema $F(10, 3, 4, 4)$.
- a $(11.4 + 3.18) + 5.05$ e $11.4 + (3.18 + 5.05)$
 - b $\frac{3.18 \times 11.4}{5.05}$ e $\left(\frac{3.18}{5.05}\right) \times 11.4$
 - c $3.18 \times (5.05 + 11.4)$ e $3.18 \times 5.05 + 3.18 \times 11.4$
- 2 Dados $x = 9370$ e $y = 12.72$, e assumindo o sistema $F(10, 3, 4, 4)$, obter:
- a $x + y$
 - b xy

- **Erro absoluto:** é a diferença entre o valor exato de um número x e de seu valor aproximado, \bar{x} .

$$\varepsilon_A = x - \bar{x} \quad (5)$$

- **Erro relativo:** é o erro absoluto dividido pelo valor exato.

$$\varepsilon_R = \frac{\varepsilon_A}{x} = \frac{x - \bar{x}}{x} \quad (6)$$

Efeitos numéricos

- Além dos erros causados pelas operações aritméticas, outros tipos de problemas podem resultar em resultados não confiáveis.
- Esses problemas são conhecidos como **efeitos numéricos**.
- Os principais são:
 - a Cancelamento.
 - b Propagação do erro.
 - c Instabilidade numérica.
 - d Mal condicionamento.

Efeitos numéricos: resumo

Cancelamento: acontece durante a subtração de dois números quase iguais.

Propagação do erro: corresponde à perda de dígitos significativos causada por uma soma intermediária grande.

Instabilidade numérica: ocorre quando os erros intermediários têm uma influência muito grande no resultado final.

- **Algoritmos estáveis:** quando os erros não afetam a qualidade final dos resultados.

Mal condicionamento: um problema é mal condicionado quando uma pequena mudança nos dados produz grandes mudanças nos resultados.

Vamos:

- 1 Elaborar um algoritmo para determinar a precisão de uma máquina.
- 2 Codificá-lo em linguagem *Python*.
- 3 Executá-lo.

- Em aritmética de ponto flutuante erros podem ocorrer devido a arredondamento de números.
- Devido aos efeitos de arredondamento, em aritmética de ponto flutuante as propriedades associativas e distributivas não são válidas.
- Existem dois tipos de erros: absoluto e relativo.
- Efeitos numéricos são problemas que podem gerar resultados não confiáveis.
- Os principais efeitos numéricos são: cancelamento, propagação do erro, instabilidade numérica e mal condicionamento.



Franco, N. B. (2007).

Cálculo numérico.

Pearson, São Paulo, SP, Brasil.



Ruggiero, M. A. G. and da Rocha Lopes, V. L. (1996).

Cálculo numérico: aspectos teóricos e computacionais.

Pearson, São Paulo, SP, 2 edition.