Cálculo Numérico: Notas de Aula: Ponto Flutuante e Erros

Prof: Felipe Figueiredo
http://sites.google.com/site/proffelipefigueiredo
Versão: 20150404

1.2.2 Aritmética de Ponto Flutuante

IEEE – Institute of Electrical and Electronic Engineers

Representação em ponto flutuante normalizada

$r = sinal \times mantissa \times base^{expoente} = \pm 0.ddddd...dd \times \beta^e$
Forma normalizada: um zero antes da vírgula, primeiro dígito depois da vírgula não-nulo.
$0.35 = 0.35 \times 10^0$
$5.47 = 0.547 \times 10^1$
$-123.456 = -12.3456 \times 10^1 = -1.23456 \times 10^2 = -0.123456 \times 10^3$ (normalizada)
Normalizar Ex: $2013.54 = 0.201354 \times 10^4$
Ex: $0.000502 = 0.502 \times 10^{-3}$

Sistema de ponto flutuante

$f(\beta, t, m, M)$, base β, t dígitos significativos, $-m \leq e \leq M$
Considere uma máquina com o sistema $f(10, 3, 5, 5)$

- Qual é o menor número positivo?

  $0.100 \times 10^{-5} = 10^{-6}$

  Representar um número menor que esse: **underflow**

- Qual é o maior número positivo?

  $0.999 \times 105 = 99900$

  Representar um número maior que esse: **overflow**

1.3.1 Erros Absolutos e Relativos

Erro absoluto: diferença entre o valor exato x e o valor aproximado $\bar{x}$

$$EA = x - \bar{x}$$

Cota superior para erro
$3.14 \leq \pi \leq 3.15$
Então:
$\pi \in (3.14, 3.15)$
$|EA_\pi| = |\pi - \bar{\pi}| < 0.01$
Erros absolutos são sempre iguais? Como comparar essas cotas para erros absolutos?
Seja $\bar{x} = 2112.9$ tal que $|EA| < 0.1$. Então $\bar{x} \in (2112.8, 2113)$.
Seja $\bar{y} = 5.3$ tal que $|EA| < 0.1$. Então $\bar{y} \in (5.2, 5.4)$.

$$ER = \frac{EA}{\bar{x}} = \frac{x - \bar{x}}{\bar{x}}$$

$|ER_x| < \dfrac{0.1}{2112.9} = 0.000047328 \approx 0.47 \times 10^{-4}$
$|ER_y| < \dfrac{0.1}{5.3} = 0.018867925 \approx 0.19 \times 10^{1}$
Portanto x é representado com maior precisão que y.

1.3.2 Erros de Arredondamento e Truncamento

E quanto o número não pode ser representado na máquina, por falta de precisão?
$f(10, 4, 5, 5)$ e $x = 234.57$
Normalizando, temos $x = 0.23457 \times 10^3$ (5 dígitos na mantissa!)
Truncamento: $\bar{x} = 0.2345 \times 10^3$
Arredondamento $\bar{x} = 0.2346 \times 10^3$
$x = 0.2345 \times 10^3 + 0.00007 \times 10^3 = 0.2345 \times 10^3 + 0.7 \times 10^{-1}$
`http://docs.oracle.com/cd/E19957-01/806-3568/ncg_goldberg.html`