Erros de arredondamento

Condição underflow e overflow. Para diferenciar estas duas condições vamos utilizar o seguinte SPF (2,3,-1,2). Neste caso, vamos representar os números x_1 = 0.38, x_2 =5.3 e x_3 =0.15 dados na base 10. Lembrando que:

De uma maneira geral, um número x é representado na base β por:

$$x = \pm\left[\frac{d_1}{\beta}+\frac{d_2}{\beta^2}+\frac{d_3}{\beta^3}+\ldots+\frac{d_t}{\beta^t}\right].\beta^{\exp}$$

$$0 \le d_i \le (\beta - 1); i = 2,3,\ldots,n.$$

$$\exp_{\min} \le \exp \le \exp_{\max}$$

$$\left[\frac{d_1}{\beta}+\frac{d_2}{\beta^2}+\frac{d_3}{\beta^3}+\ldots+\frac{d_t}{\beta^t}\right]$$

→ é chamada de mantissa e é a parte do número que representa seus dígitos significativos e t é o número de dígitos significativos do sistema de representação, **comumente chamado de precisão de máquina**.

Erros de arredondamento

Solução: Fazendo os cálculos, obtemos:

$(0.38)_{10} = 0.110 \times 2^{-1}$
$(5.3)_{10} = 0.101 \times 2^{3}$
$(0.15)_{10} = 0.100 \times 2^{-2}$

Assim no SPF (2,3,-1,2) apenas o primeiro número pode ser representado no sistema, pois para o segundo temos **OVERFLOW**, e para o terceiro **UNDERFLOW.**

Erros de arredondamento

Exemplo 2) Dado o número 12.20 que está na base 4, representá-lo na base 3.

Solução:

$$(12)_4 = 1\times 4^1 + 2\times 4^0 = (6)_{10}$$

$$(0.20)_4 = 2\times 4^{-1} + 0\times 4^{-2} = (0.5)_{10}$$

$$Logo,$$

$$(12.20)_4 = (6.5)_{10}$$

$$\begin{array}{r|l} 6 & 3 \\ \hline 0 & 2 \end{array}$$

0.5 x 3 = 1.5
0.5 x 3 = 1.5
.
.
.

Assim: $(6.5)_{10}$ = (20.11…) . Logo $(12.20)_4 = (20.11\ldots)_3$. Neste caso é importante ressaltar que o número na base 4 tem representação exata na base 10, mas não na base 3.

Erros de arredondamento

A representação de um número depende fundamentalmente da máquina utilizada pois seu Sistema estabelecerá a base numérica adotada, o comprimento de palavra, etc.

Se um dado número X não tem representação finita na base numérica deste Sistema, ou se o comprimento da palavra da máquina não comporta X, uma aproximação será obtida.

Erros de arredondamento

Regra geral para se arrendondar um número:

Dado x , seja x' sua representação SPF (β,t,m,M) adotando arredondamento. Se x=0, então x'=0. Se x≠0 , então escolhemos s e **e** tais que:

|x|= s x β^e , onde $$\beta^{-1}\left(1-\frac{1}{2}\beta^{-t}\right) \leq s < 1-\frac{1}{2}\beta^{-t} \qquad (1)$$

Se **e** está fora do intervalo [m,M] não temos condições de representar o número no sistema. Se **e** ϵ [m,M] então calculamos:

$$s+\frac{1}{2}\beta^{-t} = 0.d_1 d_2 \ldots d_{t+1} \ldots$$

e truncamos em t dígitos. Assim, o número arredondado será:

x' = (sinal x) $(0.d_1\ d_2 \ldots d_t)\beta^e$

Erros de arredondamento

Considere o sistema SPF (10,3,-5,5). Representar neste sistema os números

x_1 = 1234.56
x_2 = - 0.00054962
x_3 = 0.9995
x_4 = 123456.7 e
x_5 = -0.0000001

Solução: Primeiramente deve-se avaliar os valores permitidos para s. Desde que β=10 e t=3, usando (1), tem-se que:

$$10^{-1}\left(1-\frac{1}{2}10^{-3}\right)\leq s<1-\frac{1}{2}10^{-3}$$

Fazendo os cálculos, obtém-se:

$0.09995 \leq s < 0.9995$

Agora podemos representar os números no sistema dado:

Erros de arredondamento

a) Para x_1 =1234.56 , obtém-se:

$|x_1| = 1234.56 = 0.123456 \times 10^4$

$s + (1/2)x10^{-3} = 0.123456 + 0.0005 = 0.123956$

$x_1' = 0.123x10^4$

b) Para $x_2 = - 0.00054962$, obtém-se:

$|x| = 0.54962 \times 10^{-3}$

$s + (1/2)x10^{-3} = 0.54962 + 0.0005 = 0.55012$

$x_2' = - 0.550 \times 10^{-3}$

c) Para $x_3 = 0.9995$, observe que neste caso não podemos representar $|x_3| = 0.9995 \times 10^0$, pois neste caso s não pertence ao seu intervalo , e o número arredondado não estaria escrito na forma dos elementos do sistema. Assim, neste caso, considera-se:

Erros de arredondamento

$| x_3 | = 0.09995 \times 10^1$

$s + (1/2) \times 10^{-3} = 0.09995 + 0.0005 = 0.10045$

$x'_3 = 0.100 \times 10^1$

d) Para $x_4 = 123456.7$, obtém-se:

$| x_4 | = 0.1234567 \times 10^6$

e) Para $x_5 = -0.0000001$, obtém-se:

$| x_5 | = 0.1 \times 10^{-6}$

Observe que tanto em **d)** quanto em **e)** não podemos representar o número no sistema dado, pois em d) temos overflow e em e) temos, underflow.

Erros de arredondamento

Assim, em linhas gerais, para arredondar um número de base 10, devemos apenas observar o primeiro dígito a ser descartado.

Se este dígito é menor que 5 deixamos os dígitos inalterados; se é maior ou igual a 5 devemos somar 1 ao último dígito remanescente.

Erros de arredondamento

OPERAÇÕES EM PONTO FLUTUANTE

Considere uma máquina qualquer e uma série de operações aritméticas. Pelo fato do arredondamento ser feito após cada operação, temos, ao contrário do que é válido para números reais, que as operações aritméticas (adição, subtração, divisão e multiplicação) não são nem associativas e nem distributivas.

Vamos procurar ilustrar este fato através de exemplos.

Ex. 1) Efetuar as operações indicadas:

a) (11.4 + 3.18) + 5.05 e 11.4 + (3.18 + 5.05)

$$b)\ \frac{3.18 \times 11.4}{5.05} \quad e \quad \left(\frac{3.18}{5.05}\right) \times 11.4$$

c) 3.18 x (5.05 + 11.4) e 3.18 x 5.05 + 3.18 x 11.4

Erros de arredondamento

Para cada item, fazendo o arredondamento após cada uma das operações efetuadas, segue que:

a) (11.4 + 3.18) + 5.05 = 14.6 + 5.05 = **19.7**

enquanto

11.4 + (3.18 + 5.05) = 11.4 + 8.23 = **19.6**

$$b) \quad \frac{3.18\times 11.4}{5.05} = \frac{36.3}{5.05} = 7.19$$

enquanto

$$\left(\frac{3.18}{5.05}\right)\times 11.4 = 0.630\times 11.4 = 7.18$$

---

c) 3.18 x (5.05 + 11.4) = 3.18 x 16.5 = **52.3**

enquanto

3.18 x 5.05 + 3.18 x 11.4 = 16.1 + 36.3 = **52.4**

Erros de arredondamento

Exemplo) Avaliar o polinômio:

$P(x) = x^3 - 6x^2 + 4x - 0.1$

No ponto P(5.24) e comparar com o resultado exato.

Solução: para calcular o valor exato consideremos todos os dígitos de uma máquina, sem usar arredondamento em cada operação. Assim:

$P(5.24) = 143.8777824 - 164.7456 + 20.96 - 0.1 = -0.00776$ (valor exato)

Agora usando arrendondamento em cada operação efetuada, obtemos:

$P(5.24) = 5.24 \times 27.5 - 6 \times 27.5 + 4 \times 5.24 - 0.1$

$= 144. - 165. + 21.0 - 0.1$

$= -0.1$ (somando da esquerda para a direita)

$= 0.00$ (somando da direita para a esquerda)

Entretanto, observe que P(x) pode ser escrito como:

$P(x) = x(x(x-6)+4) - 0.1$

Erros de arredondamento

Assim

$P(5.24) = 5.24\ (5.24\ (5.24 - 6) + 4) – 0.1$

$P(5.24) = 5.24\ (-3.98 + 4) – 0.1$

$P(5.24) = 5.24\ (0.02) – 0.1$

$P(5.24) = 0.105 – 0.1$

$P(5.24) = 0.005$ (sinal errado)

Observando os exemplos, verifica-se que erros consideráveis podem ocorrer durante a execução de um algoritmo. Isto se deve ao fato que existe limitação da máquina e também porque os erros de arredondamento são introduzidso a cada operação efetuada. Em consequência, pode-se obter resultados diferentes mesmo que sejam utilizados métodos matemáticos equivalentes.

Assim devemos ser capazes de desenvolver um algoritmo tal que os efeitos da aritmética discreta do computador permaneçam inofensivos quando um grande número de operações sejam executadas.

Erros de truncamento

São erros provenientes da utilização de processos que deveriam ser infinitos ou muito grandes para determinação de um valor e que por razões práticas são truncados.

Estes processos infinitos são muito utilizados na avaliação de funções matemáticas, tais como, exponenciação, funções trigonométricas e várias outras.

Erros de truncamento

Considere a seguinte série:

$$seno(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \cdots$$

Uma máquina pode calcular a função seno(x) através do seguinte trecho de programa:

```
   FACT = 1
   SENO   = X
   SINAL = 1
   DO 10 I = 3,N,2
            FACT = FACT*I*(I-1)
            SINAL = - SINAL
            TERMO = SINAL*(X**I) / FACT
            SENO = SENO + TERMO

10 CONTINUE
```

Erros de truncamento

Para que o valor ao final do trecho de programa se tenha na variável SENO o valor de sen (x), o valor de N no comando DO deve ser bem grande, o que tornaria o cálculo ineficiente.

A solução adotada é de interromper os cálculos quando uma determinada precisão é atingida.

De uma maneira geral, pode-se dizer que o erro de truncamento pode ser diminuído até chegar a ficar da ordem do erro de arredondamento, a partir deste ponto não faz sentido diminuir-se mais, pois o erro de arredondamento será dominante.

Seguinte este raciocíonio o programa deve ser alterado para:

Erros de truncamento

```
   I = 3
   FACT = 1
   SENO   = X
   SINAL = 1

5    CONTINUE
           FACT = FACT*I*(I-1)
           SINAL = - SINAL
           TERMO = SINAL*(X**I) / FACT
           SENO = SENO + TERMO
           I = I +2
    IF (TERMO.GT.PRECISAO) GO TO 5
```

Onde,

PRECISAO é o valor de precisão da mantissa

Erros de truncamento

Exercícios) Calcular os erros de arredondamento e truncamento das operações a seguir. Considere a base decimal 10 e 4 dígitos de mantissa.

a) **t** = 0,937 x 10^4 e **y** = 0,1272 x 10^2 . Obter (**t+y**);
b) Considerando **t** e **y** do ítem a), obtenha (**ty**).