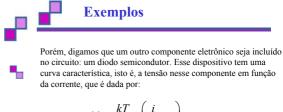
#### Cálculo numérico

Professor Walter Cunha

#### Representação numérica

Um conjunto de **ferramentas** ou **métodos** usados para se obter a solução de problemas matemáticos de forma **aproximada**. Esses métodos se aplicam principalmente a problemas que não apresentam uma solução exata, portanto precisam ser resolvidos **numericamente**.

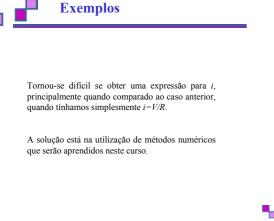
# Um circuito elétrico composto de uma fonte de tensão e um resistor. Desejamos obter a corrente que circula no circuito, dado o valor da tensão V e da resistência R. Lei de Kirchoff $V = R \ i$ $V = R \ i$ Ou seja, i = V/R $Para \ V = 10 \ V e \ R = 100 \ \Omega, \qquad i = 0,1 \ a.$



$$v(i) = \frac{kT}{q} \ln \left( \frac{i}{I_s} + 1 \right)$$

Resulta na seguinte equação:

$$V - R \cdot i - \frac{kT}{q} \ln \left( \frac{i}{I_s} + 1 \right) = 0$$



### Pontos Importantes

- > Escolher o método a ser utilizado, procurando aquele que é mais adequado para o seu problema. Que vantagens cada método oferece e que limitações eles apresentam;
- > Saber avaliar a qualidade da solução obtida. Para isso, é importante ele saber exatamente o que está sendo feito pelo computador ou calculadora, isto é, como determinado método é aplicado.



- Apresentar diversos métodos numéricos para a resolução de diferentes problemas matemáticos. Pretende-se deixar bem claro a importância desses métodos, mostrando:
- ✓ A essência de um método numérico;
- ✓ A diferença em relação a soluções analíticas;
- ✓ As situações em que eles devem ser aplicados;
- ✓ As vantagens de se utilizar um método numérico;
- ✓ As limitações na sua aplicação e confiabilidade na solução obtida.

# Exemplos

Qual o valor de  $\sqrt{2}$ 

 $\sqrt{2} = 1,41$  ou 1,4142 ou ainda 1,41421356237

Em uma máquina digital, como uma calculadora ou um computador, os números não são representados na base decimal. Eles são representados na base binária, ou seja, usam o número 2 como base. Na base binária existem somente 2 números: 0 e 1. Portanto, a base binária é usada porque essas máquinas utilizam-se de sinais elétricos, sendo o 0 para ausência de sinal e o número 1 a presença do sinal elétrico.

#### Ponto fixo e Ponto flutuante

Decimal	Ponto flutuante	Mantissa	Base	Exp
1532	0.1532 * 10 4	0.1532	10	4
15.32	0.1532 * 10 2	0.1532	10	2
0.00255	0.255 * 10 -2	0.255	10	-2
10	0.1 * 10 2	0.1	10	2



- Melhorar a familiarização e "intimidade" do aluno com a matemática, mostrando seu lado prático e sua utilidade no dia-a-dia de um engenheiro.
- Apresentar ao aluno maneiras práticas de se desenvolver e utilizar métodos numéricos.
- Treinar o aluno a aprender outros métodos numéricos por conta própria. No seu dia-a-dia profissional, ele pode se deparar com um problema cuja solução depende de um método numérico que não foi visto no curso.

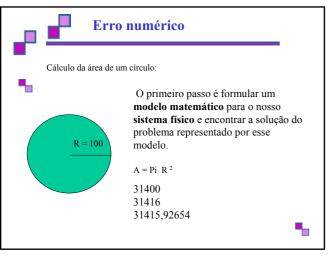


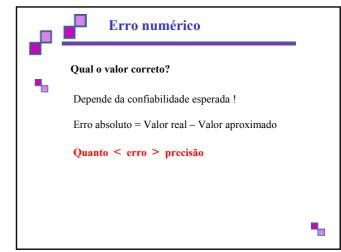
 $d_1d_2d_3...d_{-1}$  é uma fração na base b, também chamada de **mantissa**, com  $0 \le d_i \le b{-}1$ , para todo i = 1,2,3,...,t, sendo t o número máximo de dígitos da mantissa e é um expoente que varia em um intervalo dado pelos limites da máquina utilizada.



Em qualquer linguagem é possível especificar a representação que deve ser usada para os números a serem armazenados em uma dada variável. Na linguagem C, para **números inteiros** - *int* 2 bytes => (-32768, 32767).

No caso da representação de **ponto flutuante** - *float*  $\Rightarrow$  número máximo de dígitos na base binária igual a 24 (t=24) e expoente entre -126 e 127. Portanto, uma variável declarada como *float* pode armazenar números reais entre  $\sim 10^{-38}$  e  $\sim 10^{38}$ .





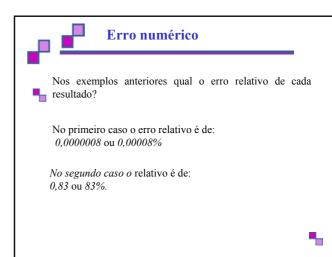


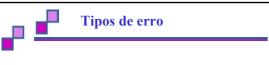
Se o resultado de uma operação é 2.123.542,7 e o valor esperado era 2.123.544,5.

O erro absoluto neste caso é 1.8. Comparada com o valor real, essa diferença é bem pequena, portanto, podemos considerar o resultado preciso.

Se o resultado de uma operação é 0.234 e o resultado esperado era 0.128. Desta vez o erro absoluto será igual a 0.106, porém o resultado é bastante impreciso.

Erro relativo = erro absoluto / valor real





Toda medida experimental apresenta uma incerteza.

A solução do problema será influenciada pela mesma.

Portanto, logo de início, existem diversos fatores que introduzem incertezas na solução numérica do problema. Esse tipo de erro é chamado de **erro inicial**.

**Erro inicial:** Erro gerado no momento de uma medição realizada através de um instrumento.



**Erro de arredondamento**: Erro gerado pela aproximação de um resultado devido a dificuldade de representação.

#### Tipos de erro

O calculo do valor de  $e^x$ . Como a *exponencial* é uma função que pode ser representada por uma série infinita dada por:

$$e^{x} = 1 + x + \frac{x^{2}}{2!} + \frac{x^{3}}{3!} + \Lambda + \frac{x^{n}}{n!} + \Lambda$$

Portanto, mais uma vez, teremos que fazer uma aproximação, que levará a um erro no resultado final de  $e^x$ .

Neste caso, faremos um truncamento dessa série.

**Erro de truncamento**: Erro gerado pelo truncamento de um resultado inexato.

#### Propagação e condicionamento

O conceito de propagação de erros é muito importante pois, além de determinar o erro final de uma operação numérica, ele também determina a **sensibilidade** de um determinado problema ou método numérico.

Se uma pequena variação nos dados de entrada de um problema levar a uma grande diferença no resultado final, considera-se que essa operação é mal-condicionada, ou seja, existe uma grande propagação de erros nessa operação. Por outro lado, se uma pequena variação nos dados de entrada leva a apenas uma pequena diferença no resultado final, então essa operação é bem-condicionada.

#### Erros na Aritmética de Ponto Flutuante

Computadores precisam representar os números com uma quantidade finita de algarismos.

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

$$X = f_{X * 10^{e}} + g_{X * 10^{e-t}}$$
  
onde  $0,1 \le f_{X} \le 1$  e  $0 \le g_{X} \le 1$ 

#### Propagação e condicionamento

Vamos supor que queremos calcular:  $\sqrt{2}$  -  $e^x$ 

O resultado da operação de subtração apresentará um erro que é proveniente dos erros nos valores de cada termo separadamente. Em outras palavras, os erros nos valores se **propagam** para o resultado.

Podemos concluir então que, ao se resolver um problema numericamente, a cada etapa e a cada operação realizada, devem surgir diferentes tipos de erros gerados das mais variadas maneiras, e estes erros se propagam e determinam o erro no resultado final obtido.

# Propagação e condicionamento

Operação mau-condicionada: Quando a operação não converge para um resultado confiável.

**Operação bem-condicionada**: Quando a operação converge para o resultado esperado.



Vamos considerar a base decimal (b=10) e uma mantissa de 4 algarismos (t=4).

O número 734,68, teríamos que truncar para 0,7346×10<sup>3</sup>

$$fx = 0.7346$$
 e  $gx = 0.8$ 

erro absoluto de 0,8×10-1

Generalizando:  $erro < b^{e-}$  (truncamento)

Erros na Aritmética de Ponto Flutuante

ou arredondar para  $0.7347 \times 10^3$  com um erro de  $0.2 \times 10^{-1}$ 

Generalizando:  $erro < \frac{1}{2} \times b^{e-t}$  (arredondamento)

Portanto, para uma representação numérica com t=53 (como no caso da maioria dos computadores) esse erro é muito pequeno.

Apesar de pequeno, é importante lembrar que ele se **propagará** nas operações aritméticas realizadas pelo computador.



#### Resumindo:

| EA | < b e-t (truncamento)

 $|EA| < \frac{1}{2}b^{e-t}$  (arredondamento)  $|ER| < b^{-t+1}$  (truncamento)

 $|ER| < \frac{1}{2}b^{-t+1}$  (arredondamento)



# Erros na Aritmética de Ponto Flutuante

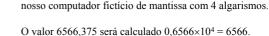
E o erro relativo?

ERx = Erro Absoluto / X

gy \* 10e-t = gy 10

 $\frac{gx * 10^{e-t}}{fx * 10^{e}} = \frac{gx * 10^{e-t}}{0.1 * 10^{e}} = gx * 10$ 





A soma de dois números exatos, o resultado da soma não será exata. Mesmo pequeno, ao se propagar esse erro precisa ser considerado.

Somando 6563 (=  $0.6563 \times 10^4$ ) e 3.375 (=  $0.3375 \times 10^1$ ) no

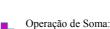
#### **Exercícios**

Seja um sistema que opera com em aritmética de ponto flutuante de t = 4 dígitos na base 10 calcule os erros absolutos e relativos por truncamento e arredondamento dos seguintes valores:

- a) 123,456 b) 374,3 + 3,345
- c) 124,34 + 0,1234
- d) 22,12 \* 0,123
- e) 0,3212 <sub>\*</sub>12,32







 $EA \Rightarrow x + y = X + EA_x + Y + EA_y = (X + Y) + EA_x + EA_y$ 

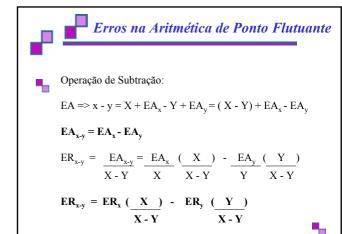
$$\mathbf{E}\mathbf{A}_{\mathbf{x}+\mathbf{y}} = \mathbf{E}\mathbf{A}_{\mathbf{x}} + \mathbf{E}\mathbf{a}_{\mathbf{y}}$$

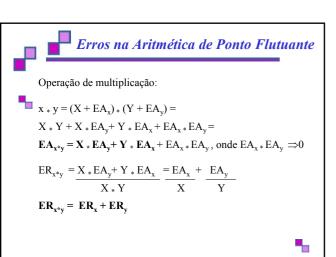
$$\frac{ER_{x+y}}{X+Y} = \frac{EA_{x+y}}{X+Y} = \frac{EA_{x}}{X} \left( \frac{X}{X+Y} \right) + \frac{EA_{y}}{Y} \left( \frac{Y}{X+Y} \right)$$

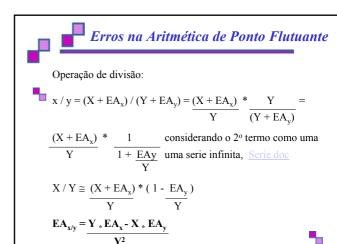
$$ER_{x+y} = ER_x (X X+Y) + ER_y (Y X+Y)$$

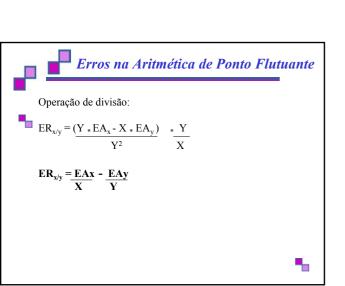


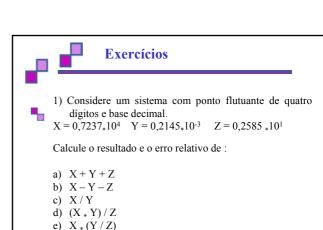
- Demonstre o erro absoluto e relativo para a operação de subtração
   Demonstre o erro absoluto e relativo para a
- Demonstre o erro absoluto e relativo para a operação de multiplicação

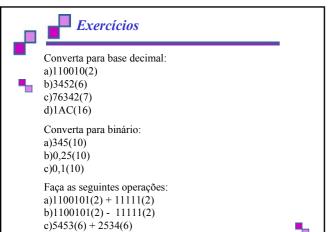












# Conversão para binário

Na conversão de números da base decimal para a base binária (usada por computadores) e vice-versa também ocorrem erros.

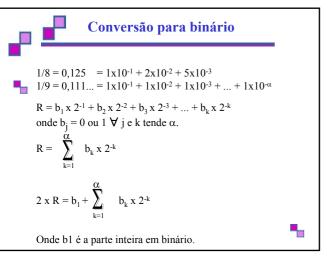
Um exemplo bastante peculiar é o número 0,1. Ao

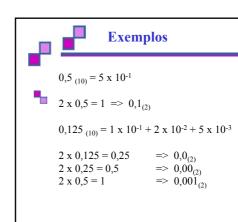
convertermos esse número da base decimal para a base

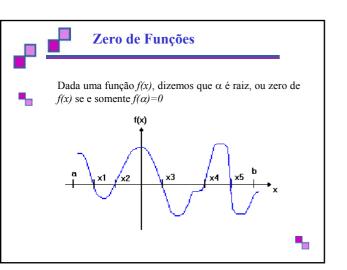
binária obtemos como resposta:  

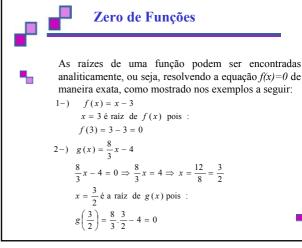
$$(0,1)_{10} = (0,0001100110011...)_{2}$$

O que gera um truncamento.  $\sum_{i=0}^{100} 0,1 \text{ não dará } 10$ 









#### Zero de Funções

$$3-) \quad h(x) = x^2 - 5x + 6$$

$$x^2 - 5x + 6 = 0$$

$$\Delta = 25 - 24 = 1$$
$$x = \frac{5 \pm \sqrt{1}}{2}$$

$$x_1 = 3$$

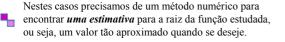
$$x_1 = 3$$
$$x_2 = 2$$

t anto 
$$x = 2$$
 quanto  $x = 3$  são soluções de  $h(x)$  pois :

$$h(3) = 3^2 - 5.3 + 6$$
  $h(2) = 2^2 - 5.2 + 6$ 

$$h(3) = 15 - 15 = 0$$
  $h(2) = 10 - 10 = 0$ 

#### Zero de Funções



Tais métodos devem envolver as seguintes etapas:

- Determinação de um intervalo em x que contenha pelo menos uma raiz da função f(x), ou seja, isolamento das raízes;
- Calculo da raiz aproximada através de um processo iterativo até a precisão desejada.



Porém, nem sempre é possível se encontrar analiticamente a raiz de uma função, como nos casos a seguir:

1-) 
$$f(x) = x^3 + 2x^2 - x + 1$$

$$2-) g(x) = sem(x) + ex$$

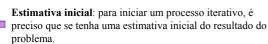
3-) 
$$h(x) = x + \ln(x)$$



#### Métodos iterativos:

Processos se caracterizam pela **repetição** de determinada operação. A idéia nesse tipo de processo é repetir determinado cálculo várias vezes, obtendo-se, a cada repetição ou **iteração**, um resultado mais preciso que aquele obtido na iteração anterior. A cada iteração utiliza-se o resultado da iteração anterior como parâmetro de entrada para o cálculo seguinte.





Convergência: a cada iteração feita, o resultado deve ser mais próximo daquele esperado, isto é, é preciso estar bem condicionado.

Critério de parada: É preciso parar as iterações em um determinado instante. Para isso, devemos utilizar um certo critério, que vai depender do problema a ser resolvido e da precisão que precisamos obter na solução.



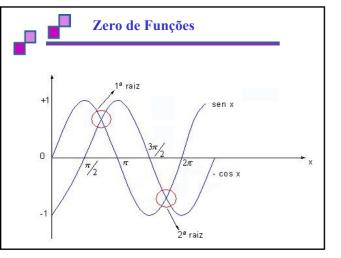
#### Zero de Funções

Isolamento de raízes:

Para determinarmos o número e a localização aproximada de raízes de uma função, a fim de obtermos uma estimativa inicial a ser usada nos processo iterativos, podemos examinar o comportamento dessa função através de um esboço gráfico.

Por exemplo, f(x) = sen(x) + cos(x)

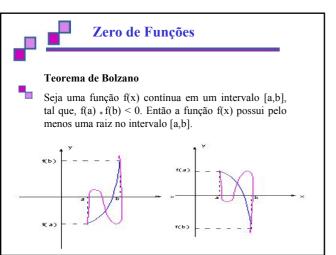
$$sen(x) + cos(x) = 0$$
,  $logo sen(x) = -cos(x)$ 





Pelo gráfico, vemos que a função g(x) irá interceptar a função h(x) entre  $\pi/2$  e  $\pi$  e entre  $3\pi/2$  e  $2\pi$ . Portanto, podemos afirmar que existe uma raiz de f(x) no intervalo  $[\pi/2, \pi]$  e no intervalo  $[3\pi/2, 2\pi]$ .

Porém, nem sempre o esboço gráfico é a forma mais prática de se obter um intervalo que contém pelo menos uma raiz da função f(x). Muitas vezes é preciso se utilizar um método algébrico. Para isso, vamos recorrer ao teorema de Bolzano.



#### Exercícios

Seja a função f(x) = x - ln(x) - 3, 2. Calcule o valor de f(x) para valores arbitrários de x de forma a descobrir o intervalo de uma raiz.

X	1	2	3	4	5
F(x)	-2,2	-1,89	-1,29	-0,58	0,19

Pelo teorema de Bolzano, concluímos que existe pelo menos uma raiz real no intervalo [4,5].

#### Exercícios

Determine o intervalo para pelo menos uma raiz das seguintes funções utilizando bozano:

1) 
$$x^3 - 9x + 3$$

2) 
$$x^3 - 2x^2 + 2x$$

3) 
$$\sqrt{x} - 5 * e^{-x}$$

#### Zero de Funções

#### Método da Dicotomia ou Bisseção

Seja uma função f(x) contínua em um intervalo [a,b], e  $\alpha$  uma raiz de f(x) isolada neste intervalo.

$$\left[a, \frac{a+b}{2}\right] e \left[\frac{a+b}{2}, b\right]$$

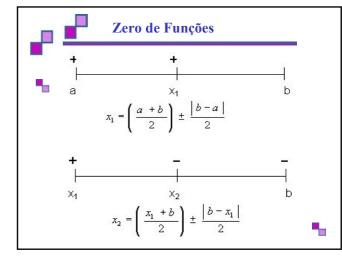
Inicialmente, subdividimos este intervalo em suas duas metades, e verificamos se a raiz está contida na primeira ou na segunda metade do intervalo inicial, usando o teorema de Bolzano.

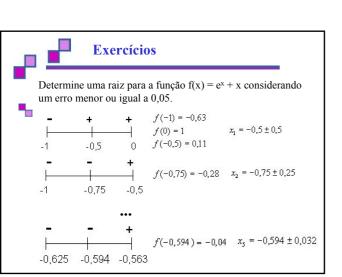
## Zero de Funções

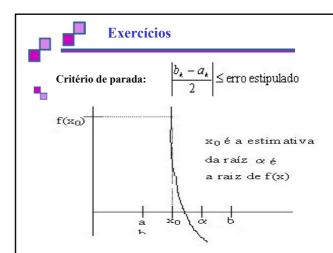
Método da Dicotomia ou Bisseção

Repetimos o processo para aquela metade que contém a raiz de f(x). Como todo processo numérico, é importante estimarmos o erro nesse resultado obtido.

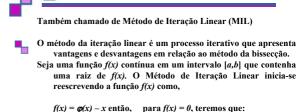
No caso do método da bisseção, o erro na estimativa será dado pela metade do comprimento do intervalo em estudo.







# Exercícios Estimativa do número de iterações k: $\begin{vmatrix} b_k - a_k | = \left| \frac{b_o - a_o}{2^k} \right| & e & \left| \frac{b_k - a_k}{2} \right| \le \varepsilon \log_o \quad 2\varepsilon \ge \left| \frac{b_o - a_o}{2^k} \right| \\ 2^{k+1} \ge \frac{\left| b_o - a_o \right|}{\varepsilon} & \Longrightarrow \quad \log(2^{k+1}) \ge \log\left(\frac{\left| b_o - a_o \right|}{\varepsilon}\right) \\ k \ge \frac{\log\left(\left| b_o - a_o \right|\right) - \log \varepsilon}{\log(2)} - 1$



Método das Aproximações Sucessivas

$$f(x) = \varphi(x) - x \text{ entable}, \quad \text{para} f(x) = 0,$$

$$f(x) = \varphi(x) - x = 0 \quad e \quad \varphi(x) = x$$

ou seja, no ponto x que corresponde à raiz de f(x), ao substituirmos o valor de x na função  $\varphi(x)$ , teremos como resultado o próprio valor de x. Portanto, a raiz de f(x) será o  $ponto\ fixo$  de  $\varphi(x)$ .

Método das Aproximações Sucessivas

 $f(x) = x^2 - x - 2$  pode ser reescrita como. Por exemplo, a função

$$f(x) = x^2 - 2 - x = \varphi(x) - x$$
, onde  $\varphi(x) = x^2 - 2$ .

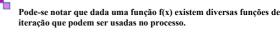
Portanto, para encontrarmos a raiz de f(x), podemos encontrar o valor numérico que ao substituirmos em  $\varphi(x)$  retorna o próprio valor de x. Para encontrarmos esse valor de x, vamos utilizar um processo iterativo, onde começamos a calcular o valor de  $\varphi(x)$  com um valor inicial de x, e recalculamos repetidamente o valor de  $\varphi(x)$ sempre usando o resultado de uma dada iteração como a nova estimativa de x, ou seja, fazendo:

$$x_{k+1} = \varphi(x_k) \tag{4.17}$$

onde, k é a ordem da iteração em que estamos (k = 0, 1, 2, 3, 4, ...). A função  $\varphi(x)$  é chamada de função de iteração.

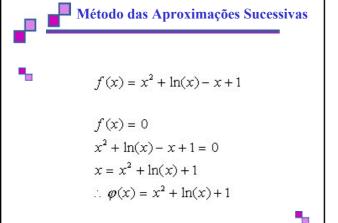


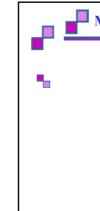
# Método das Aproximações Sucessivas

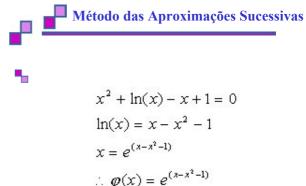


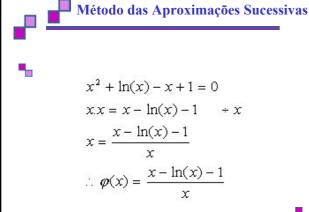
Exemplo: Encontre algumas funções de iteração a partir de

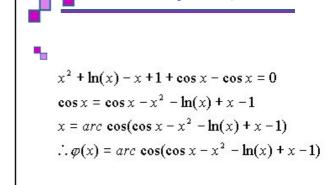
$$f(x) = x^2 + \ln(x) - x + 1.$$











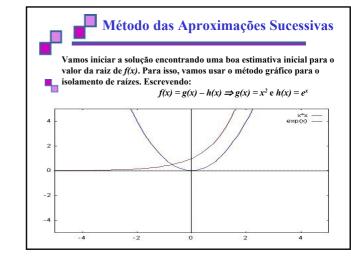
Método das Aproximações Sucessivas

Método das Aproximações Sucessivas

Exemplo: Encontre uma estimativa para a raiz de :

$$f(x) = x^2 - e^x$$

usando o método da iteração linear.



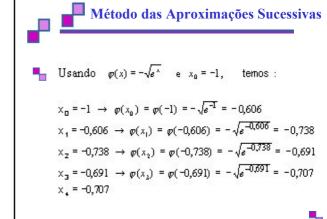
# Método das Aproximações Sucessivas

A partir do esboço gráfico acima, conclui-se que a raiz encontra-se no intervalo [-1,0]. Devemos agora escolher uma função de iteração  $\varphi(x)$ :  $f'(\chi) = 0$ 

$$x^2 - e^z = 0$$

$$x = \pm \sqrt{e^z}$$

$$x = \pm \sqrt{6}$$



### Método das Aproximações Sucessivas

Substituindo os valores de  $x_k$  em f(x) para cada iteração k, vemos que a cada etapa nos aproximamos mais da raiz de f(x), pois o valor dessa função fica mais próximo de zero a cada iteração:

x	$f(x) = x^2 - e^{x^2}$
-1	0,632 -0,178
-0,606 -0,738	0,067
-0,691 -0,707	-0,024 0,007
	ĺ

### Convergência no MIL

A cada iteração podemos nos aproximar ou nos afastar da solução real.

Portanto, antes de resolver um problema através desse método é preciso
se verificar se haverá ou não a convergência:

Seja uma função f(x) contínua em um intervalo [a,b] e  $\alpha$  uma raiz de f(x) contida em [a,b]. Seja  $\varphi(x)$  uma função de iteração obtida a partir de f(x). Se:

i) 
$$\varphi(x) \in \varphi'(x)$$
 forem continues em  $[a,b]$ ,

ii) 
$$|\varphi'(x)| \le 1 \text{ (para todo) } \forall x \in [a,b];$$

iii)  $\chi_a \in [a,b].$ 

Então:

$$\lim_{n\to+\infty} x_n = \alpha$$

Exemplo: Deseja-se encontrar a raiz de  $f(x) = x^2 + 0.96x - 2.08$ Para isto, pretende-se usar uma das seguintes funções de iteração:

$$g_{+}(x) = \frac{2,08}{x + 0.96}$$

$$g_1(x) = x^1 + 1.96 \times -2.08$$

Vamos iniciar verificando a condição (i) para a função  $\phi_l(x)$ 

$$\begin{split} \varphi_1(x) &= \frac{2.08}{x + 0.96} \\ \varphi_1(x) &= 2.08 \cdot (x + 0.96)^{-1} \\ \varphi_1'(x) &= 0.(x + 0.96)^{-1} - 2.08 \cdot (x + 0.96)^{-2} \cdot 1 \\ \varphi_1'(x) &= -\frac{2.08}{(x + 0.96)^2} \end{split}$$

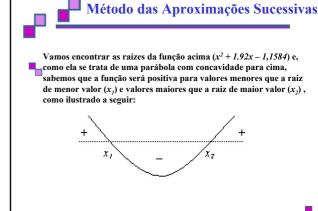
Ambas as funções  $\varphi_I(x)$  e  $\varphi_I'(x)$  são contínuas  $\forall x \in \mathbb{R} \text{ com } x \neq -0.96$ .

# Método das Aproximações Sucessivas Em seguida, vamos verificar a condição (ii) para $\varphi_i(x)$ . $\begin{vmatrix} \varphi_1'(x) & | & < 1 \\ - & \frac{208}{(x+0.96)^2} & | & < 1 \end{vmatrix}$ $\frac{|-208}{(x+0.96)^2} & < 1$ $\frac{208}{(x+0.96)^2} & < 1$

# Método das Aproximações Sucessivas

$$2,08 < (x + 0,96)^2$$
  
 $2,08 < x^2 + 1,92 x + 0,9216$   
 $x^2 + 1,92 x + 0,9216 > 2,08$   
 $x^2 + 1,92 x - 1,1584 > 0$ 

a condição (ii) do teorema da convergência é satisfeita.





As raízes da função são:

$$x^{2} + 1,92 x - 1,1584 = 0$$

$$\Delta = 8,32$$

$$x = \frac{-1,92 \pm \sqrt{8,32}}{2} = x_{1} = -2,40$$

$$x_{2} = 0,48$$

Portanto:

| \$\varphi\_i'(x) | < 1 para x<-2,40 ou x > 0,48. Finalmente, concluímos que as condições (i) e (ii) do Teorema de Convergência são

satisfeitas por:

$$\varphi_1(x) = \frac{2,08}{x + 0,96}$$
 no intervalo [1,2].