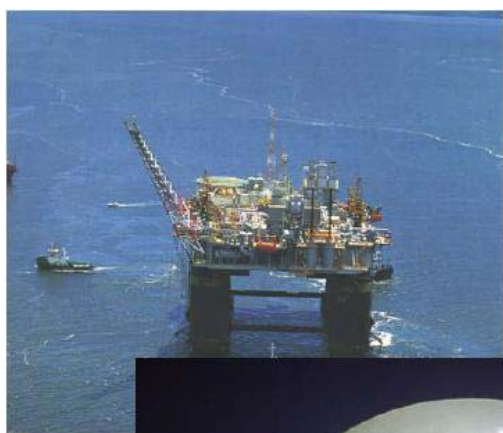


# APOSTILA DE CÁLCULO NUMÉRICO



$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{F}_{\text{int}}(\mathbf{u}, \dot{\mathbf{u}}) = \mathbf{F}_{\text{ext}}$$

**Professor: William Wagner Matos Lira**  
**Monitor: Ricardo Albuquerque Fernandes**

# 1 ERROS

## 1.1 Introdução

### 1.1.1 Modelagem e Resolução

A utilização de simuladores numéricos para determinação da solução de um problema requer a execução da seguinte sequência de etapas:

Etapa 1: Definir o problema real a ser resolvido

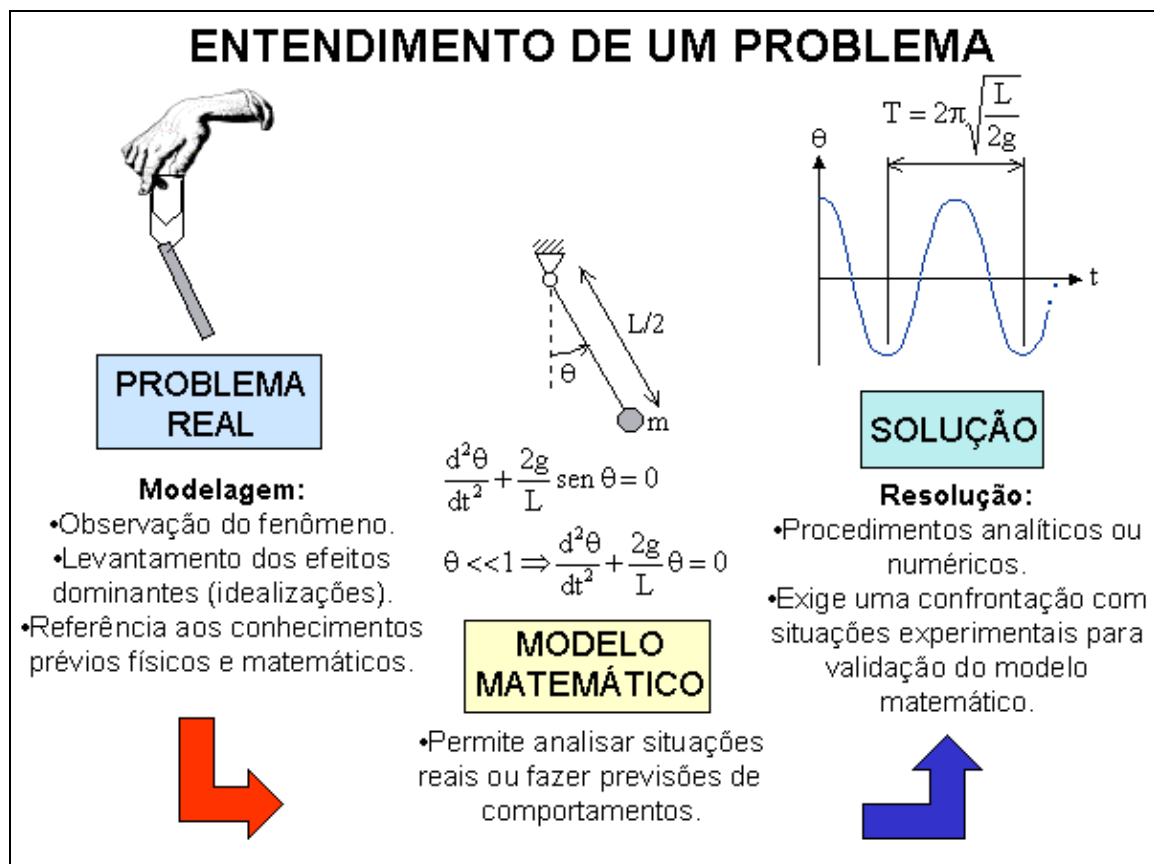
Etapa 2: Observar fenômenos, levantar efeitos dominantes e fazer referência a conhecimentos prévios físicos e matemáticos

Etapa 3: Criar modelo matemático

Etapa 4: Resolver o problema matemático

**Modelagem:** Fase de obtenção de um modelo matemático que descreve um problema físico em questão.

**Resolução:** Fase de obtenção da solução do modelo matemático através da obtenção da solução analítica ou numérica.



### 1.1.2 Cálculo Numérico

O cálculo numérico compreende:

- A análise dos processos que resolvem problemas matemáticos por meio de operações aritméticas;
- O desenvolvimento de uma sequência de operações aritméticas que levem às respostas numéricas desejadas (Desenvolvimento de algoritmos);
- O uso de computadores para obtenção das respostas numéricas, o que implica em escrever o método numérico como um programa de computador

Espera-se, com isso, obter respostas confiáveis para problemas matemáticos. No entanto, não é raro acontecer que os resultados obtidos estejam distantes do que se esperaria obter.

### 1.1.3 Fontes de erros

Suponha que você está diante do seguinte problema: você está em cima de um edifício que não sabe a altura, mas precisa determiná-la. Tudo que tem em mãos é uma bola de metal e um cronômetro. O que fazer?

Conhecemos também a equação

$$s = s_0 + v_0 t + \frac{gt^2}{2}$$

onde:

- $s$  é a posição final;
- $s_0$  é a posição inicial;
- $v_0$  é a velocidade inicial;
- $t$  é o tempo percorrido;
- $g$  é a aceleração gravitacional.

A bolinha foi solta do topo do edifício e marcou-se no cronômetro que ela levou 2 segundos para atingir o solo. Com isso podemos concluir a partir da equação acima que a altura do edifício é de 19,6 metros.

Essa resposta é confiável? Onde estão os erros?

Erros de modelagem:

- Resistência do ar,
- Velocidade do vento,
- Forma do objeto, etc.

*Estes erros estão associados, em geral, à simplificação do modelo matemático.*

Erros de resolução:

- Precisão dos dados de entrada  
(Ex. Precisão na leitura do cronômetro.  $p/t = 2,3$  segundos,  $h = 25,92$  metros, gravidade);
- Forma como os dados são armazenados;
- Operações numéricas efetuadas;
- Erro de truncamento (troca de uma série infinita por uma série finita).

## 1.2 Representação numérica

Motivação:

Exemplo 1:

Calcular a área de uma circunferência de raio 100 metros.

- a)  $31140 \text{ m}^2$
- b)  $31416 \text{ m}^2$
- c)  $31415,92654 \text{ m}^2$

Exemplo 2:

Calcular  $S = \sum_{i=1}^{3000} x_i$  para  $x_i = 0.5$  e para  $x_i = 0.11$

	S para $x_i = 0.5$	S para $x_i = 0.11$
<b>Calculadora</b>	15000	3300
<b>Computador</b>	15000	3299,99691

Por que das diferenças?

No caso do Exemplo 1 foram admitidos três valores diferentes para o número  $\pi$ :

- a)  $\pi = 3,14$
- b)  $\pi = 3,1416$
- c)  $\pi = 3,141592654$

Dependência da aproximação escolhida para  $\pi$ . Aumentando-se o número de dígitos aumentamos a precisão. Nunca conseguiremos um valor exato.

No caso do Exemplo 2 as diferenças podem ter ocorrido em função da base utilizada, da forma como os números são armazenados, ou em virtude dos erros cometidos nas operações aritméticas.

O conjunto dos números representáveis em qualquer máquina é finito, e portanto, *discreto*, ou seja não é possível representar em uma máquina todos os números de um dado intervalo  $[a,b]$ . A representação de um número depende da BASE escolhida e do número máximo de dígitos usados na sua representação.

Qual a base utilizada no nosso dia-a-dia?

Base decimal (Utiliza-se os algarismos: 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9).

Existem outras bases: 8 (base octal), 12, 60, porém, a base utilizada pela maioria dos computadores é a base binária, onde se utiliza os algarismos 0 e 1.

Os computadores recebem a informação numérica na base decimal, fazem a conversão para sua base (a base binária) e fazem nova conversão para exibir os resultados na base decimal para o usuário.

Exemplos:  $(100110)_2 = (38)_{10}$   
 $(11001)_2 = (25)_{10}$

### 1.2.1 Representação de um número inteiro

Em princípio, representação de um número inteiro no computador não apresenta qualquer dificuldade. Qualquer computador trabalha internamente com uma base fixa  $\beta$ , onde  $\beta$  é um inteiro  $\geq 2$ ; e é escolhido como uma potência de 2.

Assim dado um número inteiro  $x \neq 0$ , ele possui uma única representação,

$$x = \pm(d_n d_{n-1} \dots d_2 d_1 d_0) = \pm(d_n \beta^n + d_{n-1} \beta^{n-1} + \dots + d_1 \beta^1 + d_0 \beta^0)$$

onde  $d_i$  é um dígito da base em questão, no caso de uma base binária  $d_n = 1$  e  $d_{n-1}, \dots, d_0$  são iguais a 1 ou 0 que são os dígitos da base binária.

*Exemplos:*

a) Como seria a representação do número 1100 numa base  $\beta = 2$

$$(1100)_2 = 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 0 \times 2^0$$

Portanto  $(1100)_2 = (1100)_2$ .

b) Como seria a representação do número 1997 em uma base  $\beta = 10$ ?

$$1997 = 1 \times 10^3 + 9 \times 10^2 + 9 \times 10^1 + 7 \times 10^0$$

Logo,  $1997 = (1997)_{10}$ .

### 1.2.2 Representação de um número real

Se o número real  $x$  tem parte inteira  $x_i$ , sua parte fracionária  $x_f = x - x_i$  pode ser escrita como uma soma de frações binárias:

$$x_f = \pm(b_n b_{n-1} \dots b_2 b_1 b_0) = \pm(b_1 \beta^{-1} + b_2 \beta^{-2} + \dots + d_{n-1} \beta^{-(n-1)} + d_n \beta^{-n})$$

Assim o número real será representado juntando as partes inteiras e fracionárias, ou seja,

$$x = \pm(a_n a_{n-1} \dots a_2 a_1 a_0, b_m b_{m-1} \dots b_2 b_1 b_0)$$

onde,  $x$  possui  $n+1$  algarismos na parte inteira e  $m+1$  algarismos na parte fracionária.

*Exemplo:*

a) Como seria a representação do número 39,28 em uma base decimal?

$$(39,28)_{10} = (3 \times 10^1 + 9 \times 10^0) + (2 \times 10^{-1} + 8 \times 10^{-2})$$
$$(39,28)_{10} = (39,28)_{10}$$

b) Como seria a representação do número  $(14,375)_{10} = (?)_2$  em uma base binária?

$$(14,375)_{10} = (1110,011)_2$$

Precisamos saber fazer a conversão de bases que é o tópico seguinte.

## 1.3 Conversão entre as bases

Conforme dito anteriormente, a maioria dos computadores trabalha na base  $\beta$ , onde  $\beta$  é um inteiro  $\geq 2$ ; normalmente escolhido como uma potência de 2.

### 1.3.1 Binária para Decimal

*Exemplos:*

a)  $(1101)_2 = 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 = 8 + 4 + 0 + 1 = (13)_{10}$

b)  $(11001)_2 = 1 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 = 16 + 8 + 0 + 0 + 1 = (25)_{10}$

### 1.3.2 Decimal para Binária

Na conversão de um número escrito em base decimal para uma base binária são utilizados: o *método das divisões sucessivas* para a parte inteira e o *método das multiplicações sucessivas* para conversão da parte fracionária do número em questão.

- Método das divisões sucessivas (parte inteira do número)

- Divide-se o número (inteiro) por 2;
- Divide-se por 2, o quociente da divisão anterior;
- Repete-se o processo até o último quociente ser igual a 1.

O número binário é então formado pela concatenação do último quociente com os restos das divisões, lidos em sentido inverso.

- Método das multiplicações sucessivas (parte fracionária do número)

- Multiplica-se o número (fracionário) por 2;
- Do resultado, a parte inteira será o primeiro dígito do número na base binária e a parte fracionária é novamente multiplicada por 2;
- O processo é repetido até que a parte fracionária do último produto seja igual a zero

*Exemplos:*

a)  $(13)_{10} = (?)_2$

	Quociente	Resto
13/2	6	<b>1</b>
6/2	3	<b>0</b>
3/2	<b>1</b>	<b>1</b>

Resultado:  $(13)_{10} = (1101)_2$

b)  $(25)_{10} = (?)_2$

	Quociente	Resto
25/2	12	<b>1</b>
12/2	6	<b>0</b>
6/2	3	<b>0</b>
3/2	<b>1</b>	<b>1</b>

Resultado:  $(25)_{10} = (11001)_2$

c)  $(0,375)_{10} = (?)_2$

$$\begin{array}{rcl}
 0,375 & \xrightarrow{\times 2} & 0,750 \\
 \underline{\times 2} & & \underline{\times 2} \\
 0,750 & \xrightarrow{\times 2} & 1,500 \\
 \underline{\times 2} & & \underline{\times 2} \\
 & & 1,000
 \end{array}
 \rightarrow (0,375)_{10} = (0,011)_2$$

c)  $(13,25)_{10} = (?)_2$

Converte-se inicialmente a parte inteira do número:

	Quociente	Resto
13/2	6	1
7/2	3	0
3/2	1	1

... em seguida converte-se a parte fracionária:

$$(0,25)_{10} = (0,01)_2$$

0,25	0,50
x 2	x 2
0,50	1,0

Resultado:  $(13,25)_{10} = (1101,01)_2$

**Atenção:** Nem todo número real na base decimal possui uma representação finita na base binária. Tente fazer a conversão de  $(0,1)_{10}$ . Esta situação ilustra bem o caso de erro de arredondamento nos dados.

### 1.3.3 Exercícios Propostos

Faça as conversões indicadas abaixo:

a)  $(100110)_2 = (?)_{10}$

b)  $(1100101)_2 = (?)_{10}$

c)  $(40,28)_{10} = (?)_2$

d)  $(110,01)_2 = (?)_{10}$

e)  $(3,8)_{10} = (?)_2$

## 1.4 Arredondamento e aritmética de ponto flutuante

Um número é representado, internamente, num computador ou máquina de calcular através de uma sequência de impulsos elétricos que indicam dois estados: 0 ou 1, ou seja, os números são representados na base binária.



De uma maneira geral, um número  $x$  é representado na base  $\beta$  por:

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

onde:

$d_i$  - são números inteiros contidos no intervalo  $0 \leq d_i \leq \beta - 1$ ;  $i = 1, 2, \dots, t$ ;

$e$  - representa o expoente de  $\beta$  e assume valores entre  $I \leq e \leq S$  onde

$I, S$  - são, respectivamente, limite inferior e superior para a variação do expoente;

$\left[ \frac{d_1}{\beta} + \frac{d_2}{\beta^2} + \frac{d_3}{\beta^3} + \dots + \frac{d_t}{\beta^t} \right]$  é a chamada ***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 da máquina.

Um número real  $x$  no sistema de aritmética de ponto flutuante pode ser escrito também na forma:

$$x = \pm (0, d_1 d_2 d_3 \dots d_t) \cdot \beta^e$$

com  $d_1 \neq 0$ , pois é o primeiro algarismo significativo de  $x$ .

*Exemplos:*

a) Escrever os números reais  $x_1 = 0.35$ ,  $x_2 = -5.172$ ,  $x_3 = 0.0123$ ,  $x_4 = 0.0003$ , e  $x_5 = 5391.3$  onde estão todos na base  $\beta = 10$  em notação de um sistema de aritmética de ponto flutuante.

**Solução:**

$$\begin{aligned} 0.35 &= (3 \times 10^{-1} + 5 \times 10^{-2}) \times 10^0 = 0.35 \times 10^0 \\ -5.172 &= -(5 \times 10^{-1} + 1 \times 10^{-2} + 7 \times 10^{-3} + 2 \times 10^{-4}) \times 10^1 = -0.5172 \times 10^1 \\ 0.0123 &= (1 \times 10^{-1} + 2 \times 10^{-2} + 3 \times 10^{-3}) \times 10^{-1} = 0.123 \times 10^{-1} \\ 5391.3 &= (5 \times 10^{-1} + 3 \times 10^{-2} + 9 \times 10^{-3} + 1 \times 10^{-4} + 3 \times 10^{-5}) \times 10^4 = 0.53913 \times 10^4 \\ 0.0003 &= (3 \times 10^{-1}) \times 10^{-3} = 0.3 \times 10^{-3} \end{aligned}$$

b) Considerando agora que estamos diante de uma máquina que utilize apenas três dígitos significativos e que tenha como limite inferior e superior para o expoente, respectivamente, -2 e 2, como seriam representados nesta máquina os números do exemplo a)?

**Solução:** Temos então para esta máquina  $t = 3$ ,  $I = -2$  e  $S = 2$ . Desta forma  $-2 \leq e \leq 2$ . Sendo assim temos:

$$0.35 = 0.350 \times 10^0$$

$$-5.172 = -0.517 \times 10^1$$

$$0.0123 = 0.123 \times 10^{-1}$$

$$5391.3 = 0.53913 \times 10^4 \text{ Não pode ser representado por esta máquina. Erro de } \textbf{overflow}.$$

$$0.0003 = 0.3 \times 10^{-3} \text{ Não pode ser representado por esta máquina. Erro de } \textbf{underflow}.$$

Um erro de *overflow* ocorre quando o número é muito grande para ser representado, já um erro de *underflow* ocorre na condição contrária, ou seja, quando um número é pequeno demais para ser representado.

c) Numa máquina de calcular cujo sistema de representação utilizado de base binária, considerando que a máquina tenha capacidade para armazenar um número com dez dígitos significativos, com limites inferior e superior para o expoente de -15 e 15, respectivamente. Como que é representado o número  $(25)_{10}$  neste sistema ?

## 1.5 Erros

### 1.5.1 Erros absoluto, relativo e percentual

Erro absoluto: Diferença entre o valor exato de um número  $x$  e seu valor aproximado  $\bar{x}$  obtido a partir de um procedimento numérico.

$$E_{a_x} = x - \bar{x}$$

Em geral apenas  $x$  é conhecido, e o que se faz é assumir um limitante superior ou uma estimativa para o módulo do erro absoluto.

*Exemplos:*

a) Sabendo-se que  $\pi = (3,14; 3,15)$  tomaremos para  $\pi$  um valor dentro deste intervalo e teremos, então,  $|E_{a_x}| = |\pi - \bar{\pi}| < 0.01$ .

b) Seja  $x$  representado por  $\bar{x} = 2112,9$  de forma que  $|E_{a_x}| < 0,1$  podemos dizer que  $x \in (2112,8; 2113,0)$ .

c) Seja  $y$  representado por  $\bar{y} = 5,3$  de forma que  $|E_{a_y}| < 0,1$ , podemos dizer que  $y \in (5,2; 5,4)$

Temos que os valores para os respectivos erros absolutos nas letras b e c foram idênticos. Podemos afirmar que os valores de  $x$  e  $y$  foram representados com a mesma precisão?

O erro absoluto não é suficiente para descrever a precisão de um cálculo. Daí a maior utilização do conceito de *erro relativo*.

Erro relativo: Erro absoluto dividido pelo valor aproximado.

$$E_{r_x} = \frac{E_{a_x}}{\bar{x}} = \frac{x - \bar{x}}{\bar{x}}$$

Erro percentual: é o erro relativo em termos percentuais, ou seja:

$$E_{p_x} = E_{r_x} \times 100\%$$

*Exemplos:*

a) Seja  $x$  representado por  $\bar{x} = 2112,9$  de forma que  $|E_{a_x}| < 0,1$  podemos dizer que  $x \in (2112,8; 2113,0)$ .

$$|E_{r_x}| = \frac{|E_{a_x}|}{\bar{x}} < \frac{0,1}{2112,9} \cong 4,7 \times 10^{-5}$$

$$E_{p_x} = 4,7 \times 10^{-5} \cdot 100\% = 0,0047\%$$

b) Seja  $y$  representado por  $\bar{y} = 5,3$  de forma que  $|E_{a_y}| < 0,1$ , podemos dizer que  $y \in (5,2; 5,4)$

$$|E_{r_y}| = \frac{|E_{a_y}|}{\bar{y}} < \frac{0,1}{5,3} \cong 0,02$$

$$E_{p_y} = 0,02 \cdot 100\% = 2\%$$

Para valores próximos de 1, os erros absoluto e relativo, têm valores muito próximos. Entretanto, para valores afastados de 1, podem ocorrer grandes diferenças, e se deve

escolher um critério adequado para podermos avaliar se o erro que está sendo cometido é grande ou pequeno.

### 1.5.2 Exercícios Propostos

1. Suponha que tenhamos um valor aproximado de 0.00004 para um valor exato de 0.00005. Calcular os erros absoluto, relativo e percentual para este caso.
2. Suponha que tenhamos um valor aproximado de 100000 para um valor exato de 101000. Calcular os erros absoluto, relativo e percentual para este caso.
3. Considerando os dois casos acima, onde se obteve uma aproximação com maior precisão? Justifique sua resposta.

### 1.5.3 Erro de arredondamento e truncamento

Dar a representação dos números a seguir num sistema de aritmética de ponto flutuante de três dígitos para  $\beta = 10$ ,  $I=-4$  e  $S=4$

x	Representação por arredondamento	Representação por truncamento
1,25	$0,125 \times 10$	$0,125 \times 10$
10,053	$0,101 \times 10^2$	$0,100 \times 10^2$
-238,15	$-0,238 \times 10^3$	$-0,238 \times 10^3$
2,71828	$0,272 \times 10$	$0,271 \times 10$
0,000007	Exp < -4 (underflow)	Exp < -4 (underflow)
718235,82	Exp > 4 (overflow)	Exp > 4 (overflow)

Quando se utiliza o arredondamento os erros cometidos são menores que no truncamento, no entanto o arredondamento requer um maior tempo de execução e por esta razão o truncamento é mais utilizado. A demonstração de que no arredondamento incorremos em erros menores que no truncamento pode ser encontrada no livro de Cálculo Numérico da Márcia Ruggiero e Vera Lopes.

### 1.5.4 Propagação de erros

Será mostrado um exemplo que ilustra como os erros descritos anteriormente podem influenciar no desenvolvimento de um cálculo.

Suponhamos que as operações indicadas nos itens a) e b) sejam processadas numa máquina com 4 dígitos significativos.

- a)  $(x_2 + x_1) - x_1$
- b)  $x_2 + (x_1 - x_1)$

Fazendo  $x_1 = 0,3491 \times 10^4$  e  $x_2 = 0,2345 \times 10^0$  temos:

$$\begin{aligned} a) (x_2 + x_1) - x_1 &= (0,2345 \cdot 10^0 + 0,3491 \cdot 10^4) - 0,3491 \cdot 10^4 \\ &= 0,3491 \cdot 10^4 - 0,3491 \cdot 10^4 \\ &= 0,0000 \end{aligned}$$

$$\begin{aligned} b) x_2 + (x_1 - x_1) &= 0,2345 \cdot 10^0 + (0,3491 \cdot 10^4 - 0,3491 \cdot 10^4) \\ &= 0,2345 \cdot 10^0 + 0,0000 \\ &= 0,2345 \end{aligned}$$

A causa da diferença nas operações anteriores foi um arredondamento que foi feito na adição  $(x_2 + x_1)$  do item a), cujo resultado tem oito dígitos. Como a máquina só armazena 4 dígitos, os menos significativos foram desprezados.

Ao se utilizar uma máquina de calcular deve-se estar atento a essas particularidades causadas pelo erro de arredondamento, não só na adição, mas também nas demais operações.

## 2 ZEROS DE FUNÇÕES

### 2.1 Caracterização Matemática

- Conhecida uma função  $f(x)$ .
- Determinar o valor  $x^*$  tal que  $f(x^*)=0$ .
- Denomina-se  $x^*$  de zero da função  $f(x)$  ou raiz da equação  $f(x)=0$ .
- Solução analítica:
  - Equações algébricas (polinomiais) do 1º e 2º graus;
  - Certos formatos de equações algébricas do 3º e 4º graus;
  - Algumas equações transcendentais (não polinomiais).

### 2.2 Ilustração Através de Alguns Problemas de Engenharia

#### 2.2.1 Equilíbrio de Mecanismos

##### Exemplo:

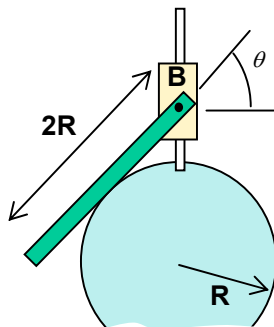
Mecânica Vetorial para Engenheiros – Estática

F. P. Beer & E. R. Johnston, Jr.

5ª Edição Revisada – 1994

MAKRON Books do Brasil Editora Ltda

Problema 4.60 (Página 254) – Uma haste delgada de comprimento  $2R$  e peso  $P$  está presa a um cursor em B e apoiada em um cilindro de raio  $R$ . Sabendo que o cursor pode se deslocar livremente ao longo de sua guia vertical, determine o valor de  $\theta$  correspondente ao equilíbrio. Despreze o atrito.



Incógnita: Ângulo  $\theta$  correspondente ao equilíbrio.  
Equação resultante durante o desenvolvimento da solução:

$$\cos^3\theta = \sin\theta$$

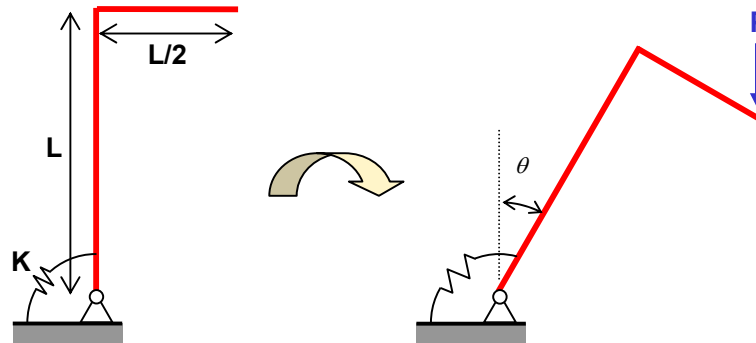
Reformatação do problema:

$$\cos^3\theta - \sin\theta = 0$$

Considerando  $f(\theta) = \cos^3\theta - \sin\theta$ , a solução da equação corresponde ao zero da função  $f(\theta)$ .

## 2.2.2 Equilíbrio de Corpos Rígidos com Apoio Deformável

**Exemplo:** Pórtico em L invertido com um apoio flexível de rotação.



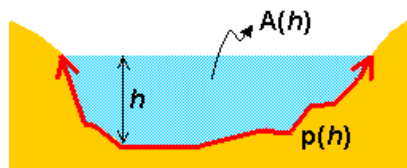
Incógnita: Ângulo  $\theta$  correspondente ao equilíbrio.  
 Equação resultante durante o desenvolvimento da solução:  
 $(K/PL) \cdot \theta = 0,5 \cdot \cos\theta + \sin\theta$   
 Reformatação do problema:  
 $(K/PL) \cdot \theta - 0,5 \cdot \cos\theta - \sin\theta = 0$   
 Considerando  $f(\theta) = (K/PL) \cdot \theta - 0,5 \cdot \cos\theta - \sin\theta$ , a solução da equação corresponde ao zero da função  $f(\theta)$ .

## 2.2.3 Equação de Manning

### EQUAÇÃO DE MANNING

**Exemplo:** Aplicação da Equação de Manning<sup>(\*)</sup> para verificação da capacidade de vazão de dutos.

(\*) Manual de Hidráulica – J. M. de Azevedo Netto – 8ª ed. atualizada – 1998 – Editora Edgard Blücher

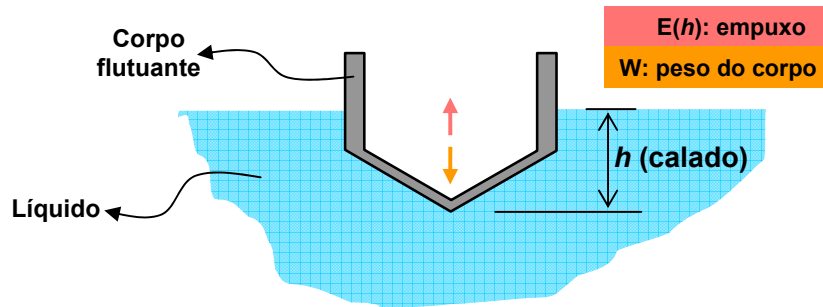


A: área molhada  
 R: raio hidráulico ( $A/p$ )  
 p: perímetro molhado  
 s: inclinação longitudinal do duto  
 n: parâmetro de rugosidade da superfície do duto  
 h: profundidade do duto  
 Q: vazão no duto

Incógnita: Profundidade  $h$  do duto.  
 Equação envolvida durante o desenvolvimento da solução:  
 $Q = \{A R^{2/3} s^{1/2}\} / n$   
 Reformatação do problema:  
 $\bar{Q} - \{A(h) R(h)^{2/3} s^{1/2}\} / n = 0$   
 Considerando  $f(h) = \bar{Q} - \{A(h) R(h)^{2/3} s^{1/2}\} / n$ , a solução da equação corresponde ao zero da função  $f(h)$ .

## 2.2.4 Equilíbrio de Corpos Flutuantes

**Exemplo:** Aplicação do Princípio de Arquimedes para a determinação do calado de embarcações.



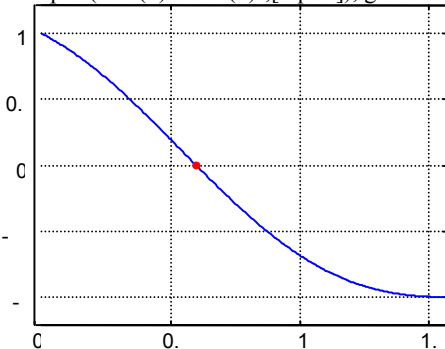
Incógnita: Profundidade  $h$  correspondente ao equilíbrio.  
Equação resultante durante o desenvolvimento da solução:  
 $\gamma_{\text{Sólido}} \cdot V_{\text{Sólido}} = \gamma_{\text{Líquido}} \cdot V_{\text{Líquido deslocado}}(h)$   
Reformatação do problema:  
 $\gamma_{\text{Sólido}} \cdot V_{\text{Sólido}} - \gamma_{\text{Líquido}} \cdot V_{\text{Líquido deslocado}}(h) = 0$   
Considerando  $f(h) = \gamma_{\text{Sólido}} \cdot V_{\text{Sólido}} - \gamma_{\text{Líquido}} \cdot V_{\text{Líquido deslocado}}(h)$ ,  
a solução da equação corresponde ao zero da função  $f(h)$ .

## 2.3 Algoritmos de Solução

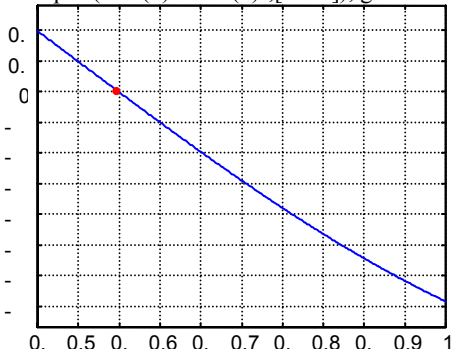
### 2.3.1 Método Gráfico

- Utilizar alguma sistemática para o **traçado do gráfico da função** estudada.
- O **intervalo inicial de observação** pode ser **criteriosamente definido** em função do entendimento físico do problema envolvido.
- O **zero da função** corresponde ao **ponto de contato** do gráfico da função com o eixo das abscissas.
- O **intervalo de observação** pode ser **refinado** até se atingir a precisão desejada.

```
> ezplot('cos(x)^3-sin(x)',[0 pi/2]), grid
```



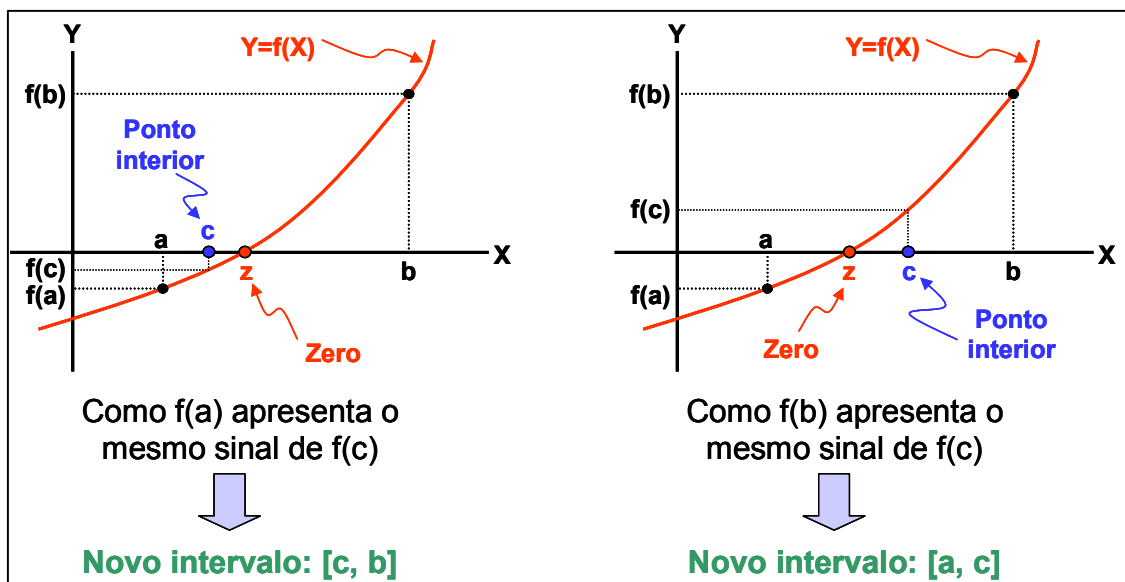
```
> ezplot('cos(x)^3-sin(x)',[0.5 1]), grid
```





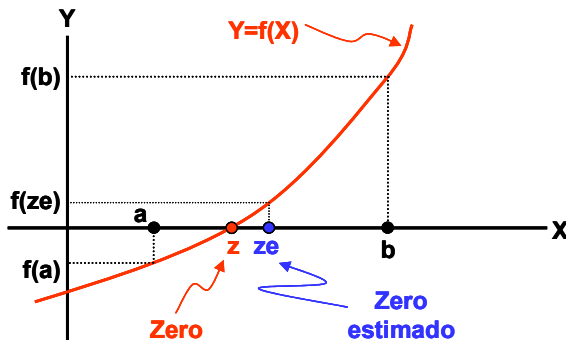
### 2.3.2 Métodos a Partir de um Intervalo (Bisseção e Cordas)

- **Pré-requisitos:**
  - Considere uma função  $f(x)$  contínua dentro de um intervalo  $[a, b]$ ;
  - Considere ainda que nos extremos do intervalo  $[a, b]$  a função estudada apresente sinais contrários, ou seja,  $f(a) \cdot f(b) < 0$ .
- **Resultado:**
  - Garante-se a existência de pelo menos um zero dessa função dentro do intervalo  $[a, b]$ .
- **Idéia:**
  - Encontrar um intervalo menor que o intervalo original e que atenda aos pré-requisitos acima mencionados;
  - Repetir o procedimento anterior até que se atinja o critério de tolerância de determinação do zero da função.
- **Estratégia de diminuição do intervalo:**
  - Nenhum cuidado especial é necessário para garantir o primeiro pré-requisito uma vez que toda função contínua em um intervalo, também será contínua em qualquer subintervalo menor;
  - Para garantir que nesse novo intervalo a função continue a apresentar sinais contrários, deve-se:
    - Escolher um ponto  $c$  dentro do intervalo original  $[a, b]$ ;
    - Redefinir o novo intervalo substituindo o extremo cujo sinal da função é o mesmo que no ponto escolhido.



### 2.3.3 Método da Bisseção

- A estimativa do zero da função  $Y=f(X)$  é feita a partir do ponto médio do intervalo analisado.
- Se o valor estimado não atender à tolerância estabelecida para o problema, ou seja,  $|f(ze)| > tol$ , redefine-se o intervalo de estudo, repetindo-se a estratégia até que a tolerância seja verificada.

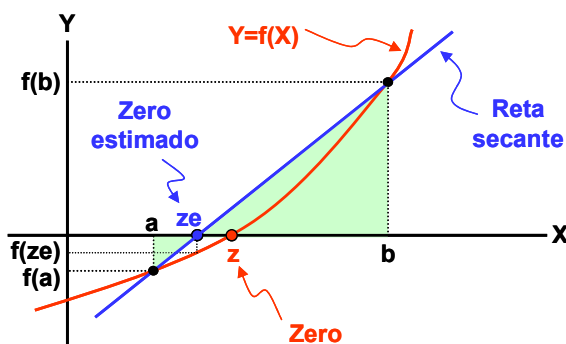


**Equação de recorrência:**

$$ze = \frac{a + b}{2}$$

### 2.3.4 Método das Cordas

- O método das cordas fundamenta-se no fato de que, geralmente, o zero da função vai estar localizado o mais próximo do extremo do intervalo onde a função apresenta o menor valor em módulo.
- A estimativa do zero da função  $Y=f(X)$  é feita a partir da **reta secante** que une os pares extremos  $(a, f(a))$  e  $(b, f(b))$  do intervalo analisado.
- O ponto em que essa reta secante intercepta o eixo das abscissas corresponde à **estimativa do zero da função**.
- Se o valor estimado não atender à tolerância estabelecida para o problema, ou seja,  $|f(ze)| > tol$ , redefine-se o intervalo de estudo, repetindo-se a estratégia até que a tolerância seja verificada.



**Equação de recorrência:**

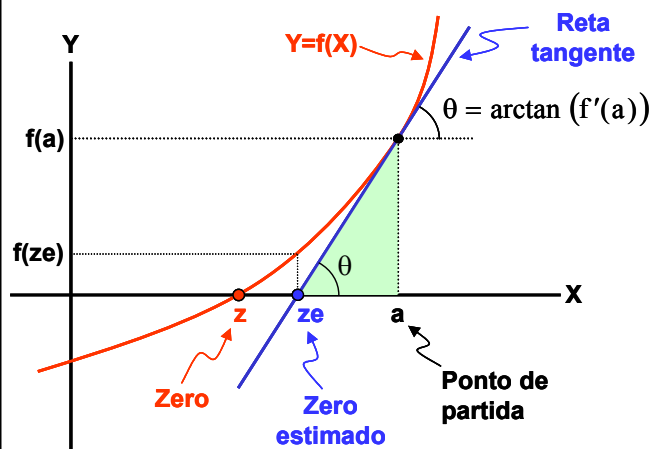
- Pela semelhança dos **triângulos retângulos** destacados na figura:

$$\frac{-f(a)}{ze - a} = \frac{f(b)}{b - ze}$$

$$\therefore ze = \frac{a \cdot f(b) - b \cdot f(a)}{f(b) - f(a)}$$

### 2.3.5 Método de Newton

- A estimativa do zero da função  $Y=f(X)$  é feita a partir da **reta tangente à função** em um **ponto de partida**.
- O ponto em que essa reta tangente intercepta o eixo das abscissas corresponde à **estimativa do zero da função**.
- Se o valor estimado não atender à tolerância estabelecida para o problema, ou seja,  $|f(ze)| > tol$ , repete-se o esquema até que a mesma seja verificada.



#### ***Equação de recorrência:***

Para o **triângulo retângulo** destacado na figura:

$$\tan(\theta) = f'(a) = \frac{f(a)}{a - ze}$$

$$\therefore ze = a - \frac{f(a)}{f'(a)}$$

## 3 SOLUÇÃO DE SISTEMAS DE EQUAÇÕES LINEARES

### 3.1 Caracterização Matemática

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

onde

- $a_{ij}$  são os coeficientes;
- $x_j$  são as variáveis;
- $b_i$  são as constantes, tal que  $1 \leq i \leq m$  e  $1 \leq j \leq n$ .

### 3.2 Notação Matricial

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}_{m \times n} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}_{m \times 1}$$

$$A x = b$$

onde

- $A$  é a matriz dos coeficientes;
- $x$  é o vetor de incógnitas;
- $b$  é o vetor de constantes.

### 3.3 Classificação quanto à solução

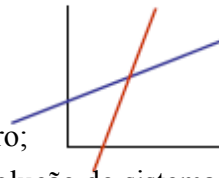


### 3.3.1 Sistema Possível ou Compatível

- Admite solução.

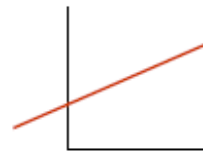
#### 3.3.1.1 Sistema Possível e Determinado

- Possui uma única solução;
- O determinante de A deve ser diferente de zero;
- Se b for um vetor nulo (constantes nulas), a solução do sistema será a solução trivial, ou seja, o vetor x também será nulo.



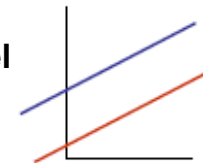
#### 3.3.1.2 Sistema Possível e Indeterminado

- Possui infinitas soluções;
- O determinante de A deve ser nulo;
- O vetor de constantes B deve ser nulo ou múltiplo de uma coluna de A.



#### 3.3.2 Sistema Impossível ou Incompatível

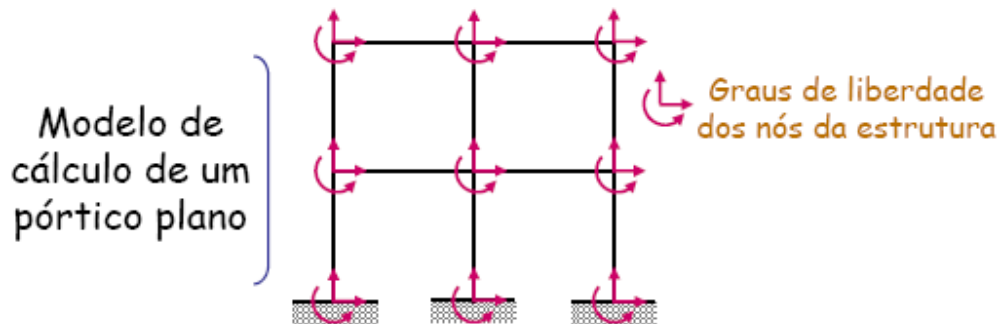
- Não possui solução;
- O determinante de A deve ser nulo;
- O vetor B não pode ser nulo ou múltiplo de alguma coluna de A.



### 3.4 Ilustração com Problemas de Engenharia

#### 3.4.1 Método da Rigidez

**Objetivo:** Aplicar a análise matricial de estruturas para calcular os deslocamentos, as reações de apoio e os esforços internos solicitantes em uma dada estrutura.



Incógnitas: Valores dos deslocamentos (translações e rotações) correspondentes aos graus de liberdade dos nós da estrutura.  
Sistema de equações lineares resultante do método da rigidez:

$$[\bar{K}]\{d\} = \{\bar{f}\}$$

Matriz dos coeficientes (Matriz de rigidez da estrutura):

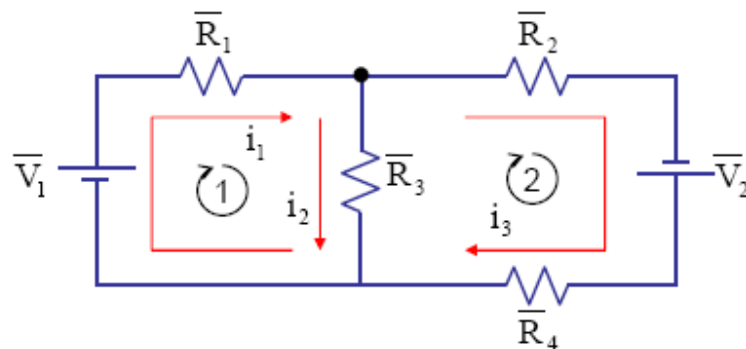
- Função de parâmetros geométricos e do material

Vetor das constantes (Vetor de forças nodais):

- Função de parâmetros geométricos, do material e das solicitações

### 3.4.2 Circuitos Elétricos

**Exemplo:** Aplicação das Leis de Kirchhoff das Tensões e das Correntes para a determinação das correntes nos circuitos elétricos CC.



Incógnitas: Correntes  $i_1$ ,  $i_2$  e  $i_3$  passando pelos vários trechos do circuito.

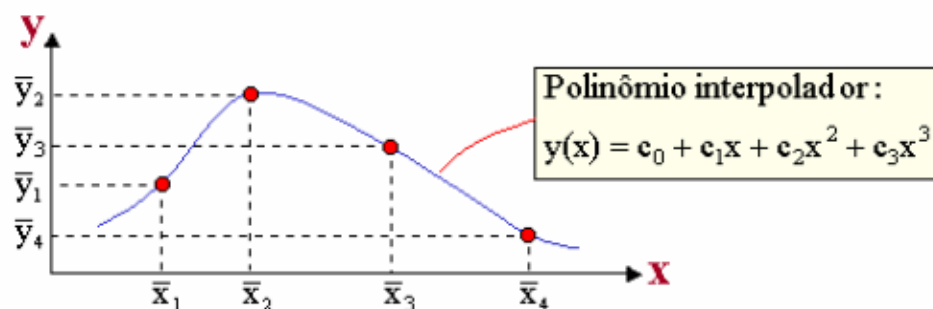
Equações resultantes da aplicação da Lei de Kirchhoff das Tensões:

$$\text{Ciclo 1: } \bar{V}_1 - \bar{R}_1 i_1 - \bar{R}_3 i_2 = 0 \text{ e Ciclo 2: } \bar{V}_2 - \bar{R}_4 i_3 + \bar{R}_3 i_2 - \bar{R}_2 i_3 = 0$$

Equação resultante da aplicação da Lei de Kirchhoff das Correntes:

### 3.4.3 Interpolação

**Objetivo:** A partir de um conjunto de pares de dados discretos (ex: tempo vs. intensidade de chuva, carga vs. deslocamento, etc), encontrar o polinômio interpolador que permita fazer estimativas para outros dados.



Incógnitas: Coeficientes  $c_0$ ,  $c_1$ ,  $c_2$  e  $c_3$  do polinômio interpolador.  
Sistema de equações lineares (notação matricial) resultante da imposição que os pares de dados pertençam ao polinômio em questão:

$$y(\bar{x}_i) = \bar{y}_i, \text{ para } i = 1..4 \Rightarrow \begin{bmatrix} 1 & \bar{x}_1 & \bar{x}_1^2 & \bar{x}_1^3 \\ 1 & \bar{x}_2 & \bar{x}_2^2 & \bar{x}_2^3 \\ 1 & \bar{x}_3 & \bar{x}_3^2 & \bar{x}_3^3 \\ 1 & \bar{x}_4 & \bar{x}_4^2 & \bar{x}_4^3 \end{bmatrix} \begin{Bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{Bmatrix} = \begin{Bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_3 \\ \bar{y}_4 \end{Bmatrix}$$

### 3.5 Classificação dos Métodos de Solução de Sistemas de Equações Lineares

### 3.5.1 Métodos Diretos

- São aqueles que conduzem à solução, exata a menos de erros de arredondamento introduzidos pela máquina, após um número finito de passos;

$$Ax = b \rightarrow x = A^{-1} b$$

- Pertencem a esta classe todos os métodos estudados no 1º e 2º graus. No entanto, esses métodos não são usados em problemas práticos quando o número de equações é elevado, pois apresentam problemas de desempenho;
- Surge então, a necessidade de utilizar técnicas mais avançadas e eficientes como: *Método de Eliminação de Gauss* e *Método de Gauss-Jordan*.

### 3.5.2 Métodos Indiretos (Iterativos)

- São aqueles que se baseiam na construção de seqüências de aproximações. A cada passo, os valores calculados anteriormente são utilizados para reforçar a aproximação.

$$Ax = b \rightarrow x_{k+1} = x_k + d$$

- O resultado obtido é aproximado;
- Geralmente são utilizados os seguintes critérios de parada para as iterações: *Limitação no número de iterações* e *Determinação de uma tolerância para a exatidão da solução*;
- Podem não convergir para a solução exata;
- Podem ser inviáveis quando o sistema é muito grande ou mal-condicionado;
- Exemplos de Métodos Iterativos: *Métodos de Gauss-Jacobi* e *de Gauss-Seidel*.

### 3.6 Métodos Diretos

$$Ax = b$$

- Para sistemas lineares possíveis e determinados de dimensão  $n \times n$ , o vetor solução,  $x$ , é dado por:

$$x = A^{-1} b$$

- No entanto, calcular explicitamente a inversa de uma matriz não é aconselhável devido à quantidade de operações envolvidas.



### 3.6.1 Método da Eliminação de Gauss

- Evita o cálculo da inversa de A;
- A solução usando o Método da Eliminação de Gauss consiste em duas etapas:
  - a) Transformação do sistema original num sistema equivalente usando uma matriz triangular superior (Escalonamento);
  - b) Resolução deste sistema equivalente.

Por questões didáticas, a resolução do sistema equivalente será mostrada antes do escalonamento do sistema.

#### 3.6.1.1 Resolução do Sistema Equivalente

Exemplo:

$$\begin{cases} 3x_1 + 2x_2 + 4x_3 = 1 \\ \phantom{3x_1} + \frac{1}{3}x_2 + \frac{2}{3}x_3 = \frac{5}{3} \\ \phantom{3x_1} \phantom{+ \frac{1}{3}x_2} - 8x_3 = 0 \end{cases} \longrightarrow x = \begin{Bmatrix} -3 \\ +5 \\ 0 \end{Bmatrix}$$

- Tendo o sistema escalonado  $n \times n$ , torna-se simples a obtenção da solução;
- Calcula-se inicialmente o  $x_3$  pela última equação;
- Depois, utiliza-se o valor de  $x_3$  na 2ª equação para obter o valor de  $x_2$ ;
- Em seguida, faz-se uso dos valores já conhecidos de  $x_2$  e  $x_3$  na 1ª equação para computar  $x_1$ .

De forma geral, temos:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1 \\ 0 \phantom{+ a_{12}x_2} + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2 \\ 0 \phantom{+ a_{12}x_2} \phantom{+ a_{22}x_2} + a_{33}x_3 + \dots + a_{3n}x_n = b_3 \\ \vdots \phantom{+ a_{12}x_2} \phantom{+ a_{22}x_2} \phantom{+ a_{33}x_3} \phantom{+ \dots} \phantom{+ a_{3n}x_n} \phantom{+ \dots} \phantom{+ a_{3n}x_n} \phantom{+ \dots} \phantom{+ a_{3n}x_n} \phantom{+ \dots} \\ 0 \phantom{+ a_{12}x_2} \phantom{+ a_{22}x_2} \phantom{+ a_{33}x_3} \phantom{+ \dots} + a_{nn}x_n = b_n \end{cases}$$

$$x_n = \frac{b_n}{a_{nn}} \quad \left| \quad x_i = \frac{b_i - \sum_{j=i+1}^n a_{ij} x_j}{a_{ii}} \right.$$

*Algoritmo para resolução do sistema equivalente*

```

 $x_n = b_n / a_{nn}$ 
Para i = (n-1), ..., 1
     $s = 0$ 
    Para j = (i+1), ..., n
         $s = s + a_{ij} x_j$ 
     $x_i = (b_i - s) / a_{ii}$ 
Fim
Fim

```

### 3.6.1.1 Escalonamento

Percorre-se os elementos abaixo da diagonal principal, transformando-os, através de operações elementares, em termos nulos, e garantindo que os elementos que já foram transformados anteriormente não mais sejam modificados.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}_{n \times n} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}_{n \times 1}$$

$$\rightarrow \begin{bmatrix} a_{11}^* & a_{12}^* & \dots & a_{1n}^* \\ 0 & a_{22}^* & \dots & a_{2n}^* \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn}^* \end{bmatrix}_{n \times n} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1} = \begin{bmatrix} b_1^* \\ b_2^* \\ \vdots \\ b_n^* \end{bmatrix}_{n \times 1}$$

*Operações Elementares*

- Permutar duas equações do sistema;
- Multiplicar uma das equações do sistema por um número real não nulo;

- c) Somar a uma das equações do sistema uma outra equação desse sistema multiplicada por um número real;

A aplicação de operações elementares ao sistema em questão garante que o novo sistema será equivalente ao original.

### 3.6.1.2 Escalonamento sem pivoteamento

Exemplo:

$$\begin{cases} 2x + y + 3z + 4w = 17 \\ x + 4y + 2z + w = 9 \\ 3x + 2y + z + 4w = 20 \\ 2x + 2y + 3z + w = 9 \end{cases}$$

$$\begin{array}{c} \text{L} \rightarrow \end{array} A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 1 & 4 & 2 & 1 \\ 3 & 2 & 1 & 4 \\ 2 & 2 & 3 & 1 \end{bmatrix} \quad b = \begin{bmatrix} 17 \\ 9 \\ 20 \\ 9 \end{bmatrix}$$

Etapas 1:

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 1 & 4 & 2 & 1 \\ 3 & 2 & 1 & 4 \\ 2 & 2 & 3 & 1 \end{bmatrix} \quad b = \begin{Bmatrix} 17 \\ 9 \\ 20 \\ 9 \end{Bmatrix}$$

Pivô:  $a_{11}$

$$m_{21} = \frac{a_{21}}{a_{11}} \quad m_{31} = \frac{a_{31}}{a_{11}} \quad m_{41} = \frac{a_{41}}{a_{11}}$$

$$L_2 = L_2 - L_1 m_{21}$$

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 0 & 3.5 & 0.5 & -1 \\ 3 & 2 & 1 & 4 \\ 2 & 2 & 3 & 1 \end{bmatrix} \quad b = \begin{Bmatrix} 17 \\ 0.5 \\ 20 \\ 9 \end{Bmatrix}$$

$$L_3 = L_3 - L_1 m_{31}$$

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 0 & 3.5 & 0.5 & -1 \\ 0 & 0.5 & -3.5 & -2 \\ 2 & 2 & 3 & 1 \end{bmatrix} \quad b = \begin{Bmatrix} 17 \\ 0.5 \\ -5.5 \\ 9 \end{Bmatrix}$$

$$L_4 = L_4 - L_1 m_{41}$$

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 0 & 3.5 & 0.5 & -1 \\ 0 & 0.5 & -3.5 & -2 \\ 0 & 1 & 0 & -3 \end{bmatrix} \quad b = \begin{Bmatrix} 17 \\ 0.5 \\ -5.5 \\ -8 \end{Bmatrix}$$

Etapla 2:

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 0 & 3.5 & 0.5 & -1 \\ 0 & 0.5 & -3.5 & -2 \\ 0 & 1 & 0 & -3 \end{bmatrix} \quad b = \begin{bmatrix} 17 \\ 0.5 \\ -5.5 \\ -8 \end{bmatrix} \quad \begin{array}{l} \text{Pivô : } a_{22} \\ m_{32} = \frac{a_{32}}{a_{22}} \\ m_{42} = \frac{a_{42}}{a_{22}} \end{array}$$

$$L_3 = L_3 - L_2 m_{32}$$

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 0 & 3.5 & 0.5 & -1 \\ 0 & 0 & -3.571 & -1.857 \\ 0 & 1 & 0 & -3 \end{bmatrix} \quad b = \begin{bmatrix} 17 \\ 0.5 \\ -5.571 \\ -8 \end{bmatrix}$$

$$L_4 = L_4 - L_2 m_{42}$$

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 0 & 3.5 & 0.5 & -1 \\ 0 & 0 & -3.571 & -1.857 \\ 0 & 0 & -0.143 & -2.714 \end{bmatrix} \quad b = \begin{bmatrix} 17 \\ 0.5 \\ -5.571 \\ -8.143 \end{bmatrix}$$

Etapla 3:

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 0 & 3.5 & 0.5 & -1 \\ 0 & 0 & -3.571 & -1.857 \\ 0 & 0 & -0.143 & -2.714 \end{bmatrix} \quad b = \begin{bmatrix} 17 \\ 0.5 \\ -5.571 \\ -8.143 \end{bmatrix} \quad \begin{array}{l} \text{Pivô : } a_{33} \\ m_{43} = \frac{a_{43}}{a_{33}} \end{array}$$

$$L_4 = L_4 - L_3 m_{43}$$

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 0 & 3.5 & 0.5 & -1 \\ 0 & 0 & -3.571 & -1.857 \\ 0 & 0 & 0 & -2.64 \end{bmatrix} \quad b = \begin{bmatrix} 17 \\ 0.5 \\ -5.571 \\ -7.92 \end{bmatrix}$$

### Algoritmo para escalonamento do sistema

```

Para j = 1,...,(n-1)
  Para i = (j+1),...,n
     $m = a_{ij}/a_{jj}$ 
    Para k = 1,...,n
       $a_{ik} = a_{ik} - m * a_{jk}$ 
    Fim
     $b_i = b_i - m * b_j$ 
  Fim
Fim

```

### 3.6.1.3 Escalonamento com pivoteamento

- Evitar que os pivôs usados no escalonamento sejam nulos.

Exemplo:

$$\begin{cases} 5x + 10y + z - 2w = -5 \\ 4x + 8y + 2z - w = 3 \\ 10x + 5y + 3z + w = 9 \\ 2x + y + z + 2w = 12 \end{cases}$$

$$\rightarrow A = \begin{bmatrix} 5 & 10 & 1 & -2 \\ 4 & 8 & 2 & -1 \\ 10 & 5 & 3 & 1 \\ 2 & 1 & 1 & 2 \end{bmatrix} \quad b = \begin{bmatrix} -5 \\ 3 \\ 9 \\ 12 \end{bmatrix}$$

Etapas 1:

$$A = \begin{bmatrix} 5 & 10 & 1 & -2 \\ 4 & 8 & 2 & -1 \\ 10 & 5 & 3 & 1 \\ 2 & 1 & 1 & 2 \end{bmatrix} \quad b = \begin{bmatrix} -5 \\ 3 \\ 9 \\ 12 \end{bmatrix}$$

Pivô:  $a_{11}$

$$m_{21} = \frac{a_{21}}{a_{11}}$$

$$m_{31} = \frac{a_{31}}{a_{11}}$$

$$m_{41} = \frac{a_{41}}{a_{11}}$$

$$L_2 = L_2 - L_1 m_{21}$$

$$A = \begin{bmatrix} 5 & 10 & 1 & -2 \\ 0 & 0 & 1.2 & 0.6 \\ 10 & 5 & 3 & 1 \\ 2 & -1 & 1 & 2 \end{bmatrix} \quad b = \begin{bmatrix} -5 \\ 7 \\ 9 \\ 12 \end{bmatrix}$$

$$L_3 = L_3 - L_1 m_{31}$$

$$A = \begin{bmatrix} 5 & 10 & 1 & -2 \\ 0 & 0 & 1.2 & 0.6 \\ 0 & -15 & 1 & 5 \\ \boxed{2} & -1 & 1 & 2 \end{bmatrix} \quad b = \begin{Bmatrix} -5 \\ 7 \\ 19 \\ 12 \end{Bmatrix}$$

$$L_4 = L_4 - L_1 m_{41}$$

$$A = \begin{bmatrix} 5 & 10 & 1 & -2 \\ 0 & \textcircled{0} & 1.2 & 0.6 \\ 0 & -15 & 1 & 5 \\ 0 & -5 & 0.6 & 2.8 \end{bmatrix} \quad b = \begin{Bmatrix} -5 \\ 7 \\ 19 \\ 14 \end{Bmatrix}$$

Critério: buscar linha com maior coeficiente em módulo.

**Trocar a segunda pela terceira linha:**

$$A = \begin{bmatrix} 5 & 10 & 1 & -2 \\ 0 & 0 & 1.2 & 0.6 \\ 0 & -15 & 1 & 5 \\ 0 & -5 & 0.6 & 2.8 \end{bmatrix} \quad b = \begin{Bmatrix} -5 \\ 7 \\ 19 \\ 14 \end{Bmatrix}$$

↓

$$A = \begin{bmatrix} 5 & 10 & 1 & -2 \\ 0 & -15 & 1 & 5 \\ 0 & 0 & 1.2 & 0.6 \\ 0 & -5 & 0.6 & 2.8 \end{bmatrix} \quad b = \begin{Bmatrix} -5 \\ 19 \\ 7 \\ 14 \end{Bmatrix}$$

**Continuar escalonando ...**

Observação:

O pivoteamento pode ser empregado com o objetivo de minimizar os erros de arredondamento e truncamento quando a matriz dos coeficientes  $A$  não for diagonalmente predominante. Antes do escalonamento de uma dada coluna, procura-se colocar como pivô o maior elemento em módulo de todos aqueles da diagonal principal **para baixo**.

Resumindo:

### Escalonamento sem pivoteamento

- Repetir da primeira até a penúltima coluna;
- Repetir para as linhas abaixo da diagonal principal;
- Aplicar operação elementar com o objetivo de zerar o elemento da linha corrente abaixo da diagonal principal;
- Alterar linha da matriz dos coeficientes;
- Alterar linha do vetor das constantes.

### Escalonamento com pivoteamento

- Repetir da primeira até a penúltima coluna;
- Verificar a necessidade de se fazer o pivoteamento;
- Procurar uma linha adequada;
- No caso de encontrar, fazer a permuta das linhas;
- Verificar a necessidade de se fazer o escalonamento da coluna corrente;
- Repetir para as linhas abaixo da diagonal principal;
- Aplicar operação elementar com o objetivo de zerar o elemento da linha corrente abaixo da diagonal principal;
- Alterar linha da matriz dos coeficientes;
- Alterar linha do vetor das constantes.

## 3.7 Métodos Iterativos

Geram uma sequência de vetores  $\{x\}^k$ , a partir de uma aproximação inicial  $\{x\}^0$ . Sob certas condições essa sequência converge para a solução, caso ela exista.

Seja o sistema linear  $Ax=b$ , onde:

- $A$ : matriz dos coeficientes,  $nxn$ ;
- $b$ : vetor de termos constantes,  $nx1$ ;
- $x$ : vetor de incógnitas,  $nx1$ .

Esse sistema é convertido, de alguma forma, em um sistema do tipo  $x = Cx + g$ , onde:

- $C$  é uma matriz  $nxn$ ;
- $g$  é um vetor  $nx1$ .

Partindo de uma aproximação inicial  $x^0$ , construímos consecutivamente os vetores:

$$x^1 = Cx^0 + g \quad x^2 = Cx^1 + g \quad x^3 = Cx^2 + g \quad x^{k+1} = Cx^k + g$$



Costuma-se adotar como critério de parada para os métodos iterativos os seguintes testes:

- $\mathbf{x}^k$  seja suficiente próximo de  $\mathbf{x}^{k-1}$   
(ou seja, distância entre  $\mathbf{x}^k$  e  $\mathbf{x}^{k-1}$  seja menor que uma dada tolerância);
- Número máximo de iterações.

### 3.7.1 Método de Gauss-Jacobi

Idéia principal:

Cada coordenada do vetor correspondente à nova aproximação é calculada a partir da respectiva equação do sistema, utilizando-se as demais coordenadas do vetor aproximação da iteração anterior.

De forma genérica tem-se o sistema  $n \times n$  abaixo:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n = b_3 \\ \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \ddots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots + a_{nn}x_n = b_n \end{cases}$$

onde  $a \neq 0$ ,  $l = 1, \dots, n$ .

Pode-se então, isolar os termos do vetor de incógnitas  $x$ , da seguinte forma:

$$\begin{aligned} x_1 &= \frac{b_1 - a_{12}x_2 - a_{13}x_3 - \dots - a_{1n}x_n}{a_{11}} \\ x_2 &= \frac{b_2 - a_{21}x_1 - a_{23}x_3 - \dots - a_{2n}x_n}{a_{22}} \\ &\vdots \\ x_n &= \frac{b_n - a_{n1}x_1 - a_{n2}x_2 - \dots - a_{n(n-1)}x_{n-1}}{a_{nn}} \end{aligned}$$

Desta forma, podemos montar a matriz C e o vetor g:

$$C = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \dots & -\frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \dots & 0 \end{bmatrix} \quad g = \left\{ \begin{array}{c} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \vdots \\ \frac{b_n}{a_{nn}} \end{array} \right\}$$

Então, pode-se calcular o vetor solução para cada iteração k, como sendo:

$$x^{(k)} = C x^{(k-1)} + g$$

Ou generalizando para os termos  $x_i$  do vetor solução de uma iteração k:

$$x_i^{(k)} = \frac{b_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} x_j^{(k-1)}}{a_{ii}}, \text{ para } i = 1, \dots, n$$

Exemplo:

$$\begin{bmatrix} 3 & 1 & 1 \\ 1 & 4 & 2 \\ 0 & 2 & 5 \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} = \begin{Bmatrix} 7 \\ 4 \\ 5 \end{Bmatrix}$$

$$\text{Chute inicial} \rightarrow \{x\}^{(0)} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$$

Calculando a matriz C e o vetor g, obtém-se:

$$C = \begin{bmatrix} 0 & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{4} & 0 & -\frac{2}{4} \\ 0 & -\frac{2}{5} & 0 \end{bmatrix} \quad g = \begin{Bmatrix} \frac{7}{3} \\ 1 \\ 1 \end{Bmatrix}$$

Pode-se então calcular o vetor x para as iterações:

$$\begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix}^{(1)} = \begin{bmatrix} 0 & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{4} & 0 & -\frac{2}{4} \\ 0 & -\frac{2}{5} & 0 \end{bmatrix} \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix} + \begin{Bmatrix} \frac{7}{3} \\ 1 \\ 1 \end{Bmatrix} \rightarrow \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix}^{(1)} = \begin{Bmatrix} \frac{7}{3} \\ 1 \\ 1 \end{Bmatrix}$$

$$\begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix}^{(2)} = \begin{bmatrix} 0 & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{4} & 0 & -\frac{2}{4} \\ 0 & -\frac{2}{5} & 0 \end{bmatrix} \begin{Bmatrix} \frac{7}{3} \\ 1 \\ 1 \end{Bmatrix} + \begin{Bmatrix} \frac{7}{3} \\ 1 \\ 1 \end{Bmatrix} \rightarrow \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix}^{(2)} = \begin{Bmatrix} \frac{5}{3} \\ 1 \\ \frac{12}{5} \end{Bmatrix}$$

Os resultados obtidos para as iterações estão dispostos na tabela a seguir:

Iteração	$\{x\}^{(k)}$		
1	2.3333	1.0000	1.0000
2	1.6667	-0.0833	0.6000
3	2.1611	0.2833	1.0333
4	1.8944	-0.0569	0.8867
5	2.0568	0.0831	1.0228
6	1.9647	-0.0256	0.9668
7	2.0196	0.0254	1.0102
8	1.9881	-0.0100	0.9898
...	...		
<b>Solução exata</b>	<b>2.0000</b>	<b>0.0000</b>	<b>1.0000</b>

Observa-se que quanto mais iterações forem realizadas, mais próximo estará o vetor  $x$  da solução exata do sistema linear.

### Algoritmo

Enquanto

- $\text{dist} > \text{tolerância}$
- $\text{nite} < \text{número máximo de iterações}$ .

então:

nite = nite + 1

Para  $i = 1, \dots, n$

$s = b_i$

Para  $j = 1, \dots, n$

Se  $i$  for diferente de  $j$

$s = s - a_{ij} * x_{0j}$

Fim

Fim

$x_i = s/a_{ii}$

Fim

$\text{dist} = \text{norma}(x - x_0)$

$x_0 = x$

Fim

$$x_i^{(k)} = \frac{b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^n a_{ij} x_j^{(k-1)}}{a_{ii}}, \text{ para } i = 1, \dots, n$$

Exemplo:

$$\begin{bmatrix} 3 & 1 & 1 \\ 1 & 4 & 2 \\ 0 & 2 & 5 \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} = \begin{Bmatrix} 7 \\ 4 \\ 5 \end{Bmatrix}$$

Chute inicial  $\rightarrow \{x\}^{(0)} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$

Estimativas para a primeira iteração:

$$\begin{cases} x_1^{(1)} = \frac{1}{3}(7 - 1x_2^0 - 1x_3^0) = \frac{7}{3} \\ x_2^{(1)} = \frac{1}{4}(4 - 1x_1^1 - 2x_3^0) = \frac{5}{12} \\ x_3^{(1)} = \frac{1}{5}(5 - 0x_1^1 - 2x_2^1) = \frac{5}{6} \end{cases}$$

Os resultados obtidos para as iterações estão dispostos na tabela a seguir.

Note que para o mesmo sistema linear e para um mesmo chute inicial, o método de Gauss-Seidel tende a convergir para a resposta exata do sistema numa quantidade menor de iterações que o método de Gauss-Jacobi. Isto ocorre porque como vimos, o método de Gauss-Seidel faz uso de elementos do próprio vetor solução da iteração corrente para atualizar sua estimativa.

<b>Iteração</b>	<b><math>\{x\}^{(k)}</math></b>		
<b>1</b>	2.3333	0.4167	0.8333
<b>2</b>	1.9167	0.1042	0.9583
<b>3</b>	1.9792	0.0260	0.9896
<b>4</b>	1.9948	0.0065	0.9974
<b>5</b>	1.9987	0.0016	0.9993
<b>6</b>	1.9997	0.0004	0.9998
<b>7</b>	<b>1.9999</b>	<b>0.0001</b>	<b>1.0000</b>
...	...		
<b>Solução exata</b>	<b>2.0000</b>	<b>0.0000</b>	<b>1.0000</b>

Observa-se também que quanto mais iterações forem realizadas, mais próximo estará o vetor  $x$  da solução exata do sistema linear em questão.

### Algoritmo

Enquanto

- $\text{dist} > \text{tolerância}$
- $\text{nite} < \text{número máximo de iterações}$ .

então:

$\text{nite} = \text{nite} + 1$

    Para  $i = 1, \dots, n$

$s_0 = b_i$

$s_1 = 0$ ;

        Para  $j = 1, \dots, (i-1)$

$s_0 = s_0 - a_{ij} * x_j$

        Fim

        Para  $j = (i+1), \dots, n$

$s_1 = s_1 - a_{ij} * x_0$

        Fim

$x_i = (s_0 + s_1) / a_{ii}$

    Fim

$\text{dist} = \text{norma}(x - x_0)$

$x_0 = x$

Fim

### **3.7.3 Condição de suficiência para a convergência dos métodos iterativos**

Ao se utilizar um método iterativo para solucionar um sistema de equações lineares deve tomar cuidado pois, dependendo do sistema em questão, e da estimativa inicial escolhida, o método pode não convergir para a solução do sistema.

Existem, porém, alguns critérios para verificar a convergência de um método iterativo. Basta atender a pelo menos um deles para que a convergência ocorra independentemente da aproximação inicial escolhida.

Nesses critérios são calculados valores  $\alpha_s$ , onde  $1 \leq s \leq n$ . A condição de convergência é que o valor máximo de todos os  $\alpha_s$  deve ser inferior a 1.

#### **3.7.3.1 Critério das linhas**

Os valores de  $\alpha_s$  são calculados conforme a equação abaixo:

$$\alpha_s = \frac{\sum_{j=1, j \neq s}^n |\bar{a}_{sj}|}{|\bar{a}_{ss}|}$$

### 3.7.3.2 Critério das colunas

Os valores de  $\alpha_s$  são calculados conforme a equação abaixo:

$$\alpha_s = \frac{\sum_{\substack{j=1 \\ j \neq s}}^n |\bar{a}_{js}|}{|\bar{a}_{ss}|}$$

### 3.7.3.3 Critério de Sassenfeld

Onde o valor de  $\alpha_1$  é calculado da mesma forma que o  $\alpha_1$  do critério das linhas:

$$\alpha_1 = \frac{|\bar{a}_{12}| + \dots + |\bar{a}_{1n}|}{|\bar{a}_{11}|}$$

E os demais  $\alpha_s$  são calculados utilizando os valores já calculados de  $\alpha_s$ :

$$\alpha_s = \frac{|\bar{a}_{s1}| \alpha_1 + \dots + |\bar{a}_{ss-1}| \alpha_{s-1} + |\bar{a}_{ss+1}| + \dots + |\bar{a}_{sn}|}{|\bar{a}_{ss}|}$$

Exemplo:

Utilizando o critério das linhas, verificar se o sistema com matriz dos coeficientes A abaixo garante condição de convergência para os métodos iterativos.

$$A = \begin{bmatrix} 10 & 2 & 1 \\ 1 & 5 & 1 \\ 2 & 3 & 10 \end{bmatrix}$$

$$\alpha_1 = \frac{2+1}{10} = \frac{3}{10} = 0.3 < 1$$

$$\alpha_2 = \frac{1+1}{5} = \frac{2}{5} = 0.4 < 1$$

$$\alpha_3 = \frac{2+3}{10} = \frac{5}{10} = 0.5 < 1$$

$$\max_{1 \leq s \leq 3} \alpha_s = 0.5 < 1 \rightarrow \text{Garantia de convergência}$$

Verifica-se então que independentemente do chute inicial para o vetor solução  $x^0$ , ao utilizar um método iterativo para resolver um sistema linear com matriz dos coeficientes igual a apresentada acima, irá se convergir para a solução exata do sistema.

## 4 INTERPOLAÇÃO

Interpolarm uma função  $f(x)$  consiste em aproximar essa função por uma outra função  $p(x)$  que satisfaça algumas propriedades. Em geral, a interpolação de funções é usada nas seguintes situações:

- São conhecidos valores numéricos da função  $f(x)$  em alguns pontos discretos de  $x$  e deseja-se obter valores de  $f(x)$  em pontos desconhecidos, mas dentro do limite avaliado;
- Quando uma determinada função  $f(x)$  possui os operadores de diferenciação e integração muito complexos;
- Na solução numérica de equações diferenciais usando o método das diferenças finitas e o método dos elementos finitos.

Considere  $n$  pontos distintos  $(x_1, f(x_1)), (x_2, f(x_2)), \dots, (x_n, f(x_n))$ .

O objetivo é encontrar uma função interpolante  $p(x)$ , tal que:

$$p(x_k) = f(x_k), \quad k = 1, \dots, n$$

$$p(x_1) = f(x_1)$$

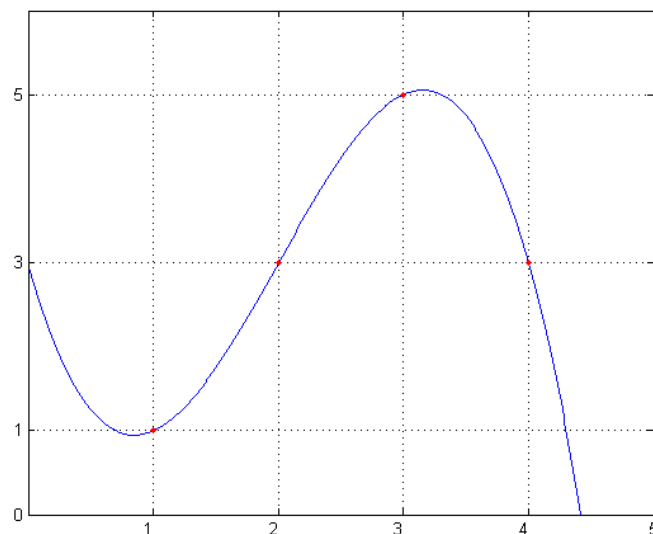
$$p(x_2) = f(x_2)$$

$$\vdots$$

$$p(x_n) = f(x_n)$$

As principais técnicas de interpolação utilizadas atualmente são baseadas em **polinômios** (ou seja,  $p(x)$  é uma função polinomial).

O gráfico abaixo representa uma função interpoladora para os pontos  $(1,1)$ ,  $(2,3)$ ,  $(3,5)$  e  $(4,3)$ . Note que a curva intercepta todos os pontos fornecidos.





## 4.1 Métodos Numéricos para Interpolação

Dados  $n$  pontos distintos  $(x_1, f(x_1)), (x_2, f(x_2)), \dots, (x_n, f(x_n))$ , deseja-se aproximar  $f(x)$  por um polinômio  $p(x)$  de grau menor ou igual a  $(n-1)$ .

Suponha que você tenha dois pontos distintos ( $n=2$ ), então, o melhor polinômio que interpola esses dois pontos será uma reta, ou seja, um polinômio de grau 1. Da mesma forma, dados 3 pontos distintos, o melhor polinômio será uma parábola. Caso você forneça, por exemplo, 3 pontos ( $n=3$ ) que pertençam a uma reta, o polinômio interpolador ainda sim será ter grau 1, ou seja, grau menor que  $(n-1)$ .

### 4.1.1 Método de Vandermonde

Considerando a condição básica para a interpolação:

$$f(x_k) = p(x_k), \quad k = 1, \dots, n$$

e o fato de que o polinômio interpolador terá, no máximo, grau  $(n-1)$ , temos que o polinômio interpolador assumirá a seguinte forma:

$$p(x) = a_1 + a_2x + a_3x^2 + \dots + a_nx^{n-1}$$

Então, obter o polinômio  $p(x)$ , significa encontrar os coeficientes  $a_1, \dots, a_n$  de forma que  $p(x_k) = f(x_k)$ , para  $k=1, \dots, n$ .

Desse modo:

$$\begin{cases} a_1 + a_2x_1 + a_3x_1^2 + \dots + a_nx_1^{n-1} = f(x_1) \\ a_1 + a_2x_2 + a_3x_2^2 + \dots + a_nx_2^{n-1} = f(x_2) \\ a_1 + a_2x_3 + a_3x_3^2 + \dots + a_nx_3^{n-1} = f(x_3) \\ \vdots \\ a_1 + a_2x_n + a_3x_n^2 + \dots + a_nx_n^{n-1} = f(x_n) \end{cases}$$

Obtém um sistema de equações lineares, com  $n$  equações e  $n$  incógnitas.

Escrevendo o sistema acima na notação matricial, tem-se:

$$A x = b$$

$$\begin{bmatrix} x_1^0 & x_1^1 & x_1^2 & x_1^3 & \dots & x_1^{n-1} \\ x_2^0 & x_2^1 & x_2^2 & x_2^3 & \dots & x_2^{n-1} \\ x_3^0 & x_3^1 & x_3^2 & x_3^3 & \dots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n^0 & x_n^1 & x_n^2 & x_n^3 & \dots & x_n^{n-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ f(x_4) \\ \vdots \\ f(x_n) \end{bmatrix}$$

A matriz A é a chamada matriz de Vandermonde e desde que os valores de  $x_1, x_2, \dots, x_n$  sejam distintos, o determinante de A é diferente de zero, e então, o sistema apresenta solução única. Então, para encontrar o polinômio interpolador de uma série de n pontos distintos conhecidos, basta encontrar a solução do sistema linear acima, assunto tratado no capítulo anterior.

Exemplo:

Encontrar o polinômio de grau 2 que interpola os pontos da tabela abaixo:

$x$	$f(x)$
-1	4
0	1
2	-1

Solução:

$$p(x) = a_1 + a_2x + a_3x^2$$

$$p(x_1) = f(x_1) \rightarrow a_1 - a_2 + a_3 = 4$$

$$p(x_2) = f(x_2) \rightarrow a_1 = 1$$

$$p(x_3) = f(x_3) \rightarrow a_1 + 2a_2 + 4a_3 = -1$$

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 2 & 4 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 1 \\ -1 \end{bmatrix}$$

Resolvendo o sistema:

$$x = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{7}{3} \\ \frac{2}{3} \end{bmatrix} \rightarrow p(x) = 1 - \frac{7}{3}x + \frac{2}{3}x^2$$

\*\* Obs: A matriz dos coeficientes A (Matriz de Vandermonde) pode estar mal condicionada, neste caso, tal método não é recomendado.

Algoritmo:

Para  $i = 1, \dots, n$

Para  $j = 1, \dots, n$

$$A_{ij} = x_i^{j-1}$$

Fim

Fim

$$a = A^{-1} \cdot \{y\}$$

### 4.1.2 Método de Lagrange

Seja  $p(x)$  um polinômio com grau  $(n-1)$  que interpola  $f$  em  $x_1, x_2, \dots, x_n$ .  
Então, podemos representar  $p(x)$  na forma:

$$p(x) = f(x_1)L_1(x) + f(x_2)L_2(x) + \dots + f(x_n)L_n(x)$$

ou

$$p(x) = \sum_{i=1}^n f(x_i)L_i(x)$$

A equação mostrada acima é o chamado *Polinômio de Lagrange*, onde

$$L_i(x) = \prod_{k=1, k \neq i}^n \frac{x - x_k}{x_i - x_k}$$

Vamos tomar como exemplo um polinômio quadrático ( $n=3$ ), então:

$$L_1(x) = \prod_{k=1, k \neq 1}^3 \frac{x - x_k}{x_1 - x_k} = \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)}$$

$$L_2(x) = \prod_{k=1, k \neq 2}^3 \frac{x - x_k}{x_2 - x_k} = \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)}$$

$$L_3(x) = \prod_{k=1, k \neq 3}^3 \frac{x - x_k}{x_3 - x_k} = \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}$$

E desse modo:

$$p(x) = f(x_1)L_1(x) + f(x_2)L_2(x) + f(x_3)L_3(x)$$

Exemplo:

Encontre o polinômio de grau 2 que interpole o seguinte conjunto de pontos

$x$	$f(x)$
-1	4
0	1
2	-1

Solução:

Polinômio adotado de grau 2, então  $n=3$ , logo:

$$p(x) = f(x_1)L_1(x) + f(x_2)L_2(x) + f(x_3)L_3(x)$$

$$L_1(x) = \frac{(x-0)(x-2)}{(-1-0)(-1-2)} = \frac{x^2 - 2x}{3}$$

$$L_2(x) = \frac{(x-(-1))(x-2)}{(0-(-1))(0-2)} = \frac{x^2 - x - 2}{-2}$$

$$L_3(x) = \frac{(x-(-1))(x-0)}{(2-(-1))(2-0)} = \frac{x^2 + x}{6}$$

Então, o polinômio interpolador de Lagrange é:

$$p(x) = 4 \frac{x^2 - 2x}{3} + 1 \frac{-x^2 + x + 2}{2} - 1 \frac{x^2 + x}{6}$$

$$p(x) = \frac{8x^2 - 16x - 3x^2 + 3x + 6 - x^2 - x}{6}$$

$$p(x) = \frac{2}{3}x^2 - \frac{7}{3}x + 1$$

Algoritmo:

$p = 0$

Para  $i = 1, \dots, n$

$s = 1$

    Para  $k = 1, \dots, n$

        Se  $k$  for diferente de  $i$

$s = s * (x - x_k) / (x_i - x_k)$

        Fim

    Fim

$p = p + f(x_i) * s$

Fim

### 4.1.3 Método de Newton

A fórmula de Newton é dada por:

$$p(x) = d_1 + d_2(x - x_1) + d_3(x - x_1)(x - x_2) + \dots + d_n(x - x_1)(x - x_2) \dots (x - x_{n-1})$$

onde  $d_k$  são os operadores diferenças divididas entre os pontos  $(x_1, f(x_1)), \dots, (x_n, f(x_n))$ .

Esses operadores são dados por:

$$d_1 = f[x_1] = f(x_1)$$

$$d_2 = f[x_1, x_2] = \frac{f[x_2] - f[x_1]}{x_2 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

$$d_3 = f[x_1, x_2, x_3] = \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1} = \frac{\frac{f[x_3] - f[x_2]}{x_3 - x_2} - \frac{f[x_2] - f[x_1]}{x_2 - x_1}}{x_3 - x_1}$$

;

$$d_n = f[x_1, x_2, \dots, x_{n-1}, x_n] = \frac{f[x_2, x_3, \dots, x_n] - f[x_1, x_2, \dots, x_{n-1}]}{x_n - x_1}$$

Desse modo:

$$p(x) = f(x_1) + f[x_1, x_2](x - x_1) + f[x_1, x_2, x_3](x - x_1)(x - x_2) + \dots + f[x_1, x_2, \dots, x_n](x - x_1)(x - x_2)(x - x_{n-1})$$

Exemplo:

Encontrar o polinômio de grau 2 que interpole o seguinte conjunto de pontos:

$x$	$f(x)$
-1	4
0	1
2	-1

Solução:

Grau do polinômio = 2, logo  $n=3$

Polinômio adotado:

$$p(x) = f[x1] + f[x1,x2](x - x1) + f[x1,x2,x3](x - x1)(x - x2)$$

Calculando os operadores diferenças divididas:

$$f[x1] = f(x1) = 4$$

$$f[x1,x2] = \frac{f[x2] - f[x1]}{x2 - x1} = \frac{f(x2) - f(x1)}{x2 - x1} = \frac{1 - 4}{0 - (-1)} = -3$$

$$f[x1,x2,x3] = \frac{f[x2,x3] - f[x1,x2]}{x3 - x1} = \frac{\frac{f(x3) - f(x2)}{x3 - x2} - f[x1,x2]}{x3 - x1}$$

$$\rightarrow f[x1,x2,x3] = \frac{\frac{-1 - 1}{2 - 0} - (-3)}{2 - (-1)} = \frac{2}{3}$$

Então, tem-se o Polinômio de Newton:

$$p(x) = 4 - 3(x - (-1)) + \frac{2}{3}(x - (-1))x = 4 - 3(x + 1) + \frac{2}{3}x(x + 1)$$

$$\rightarrow p(x) = \frac{2}{3}x^2 - \frac{7}{3}x + 1$$

Algoritmo:

D = matriz nula  $n \times n$

1ª coluna de D = {y}

Para  $j = 2, \dots, n$

Para  $i = j, \dots, n$

$$d_{i,j} = (d_{i,j-1} - d_{i-1,j-1}) / (x_i - x_{i-j+1})$$

Fim

Fim

p=0

Para  $i = 1, \dots, n$

s = 1

Para  $j = 1, \dots, (i-1)$

$$s = s * (x - x_j)$$

Fim

$$p = p + s * d_{i,i}$$

Fim

\*\* Obs: Note que para cada método numérico de interpolação apresentado utilizou-se o mesmo exemplo e como resposta para todos os casos foi obtido o mesmo polinômio interpolador.

## 5 AJUSTE

Dado um conjunto de pontos, no ajuste ou aproximação tenta-se encontrar uma função  $p(x)$  que melhor aproxime esses pontos. Aqui, não existe a necessidade da função passar pelos pontos conhecidos.

Em geral, usa-se aproximação de funções nas seguintes situações:

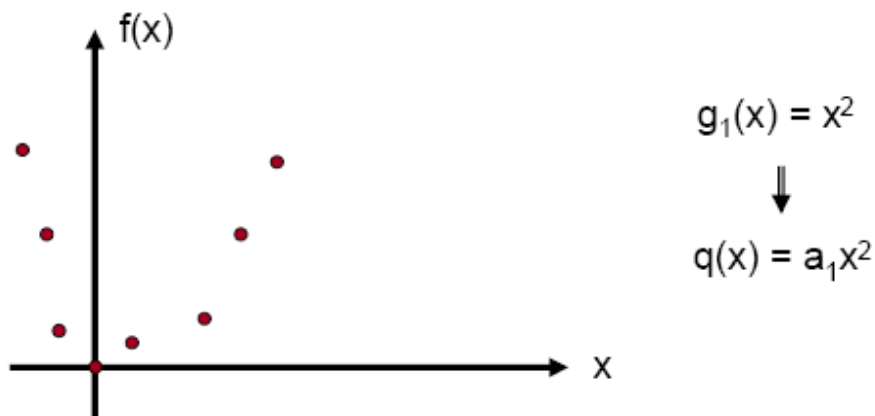
- Quando se desejar extrapolar ou fazer previsões em regiões fora do intervalo considerado;
- Quando os dados tabelados são resultados de experimentos, onde erros na obtenção destes resultados podem influenciar a sua qualidade.

Considere uma tabela de  $m$  pontos  $(x_1, f(x_1)), (x_2, f(x_2)), \dots, (x_m, f(x_m))$  com  $x_1, x_2, \dots, x_m$  pertencentes a um intervalo  $[a, b]$ . Deseja-se encontrar uma função  $q(x) = a_1g_1(x) + a_2g_2(x) + \dots + a_n g_n(x)$  que melhor ajuste esses pontos. Ou seja, determinar a função  $q(x)$  que mais se **aproxime** de  $f(x)$ .

Diz-se que este é um modelo matemático linear porque os coeficientes a determinar aparecem linearmente, embora as funções  $g_1(x), g_2(x), \dots, g_n(x)$  possam ser funções não lineares de  $x$  como, por exemplo,  $g_1(x) = ex$ ,  $g_2(x) = 1+x^2$ , etc.

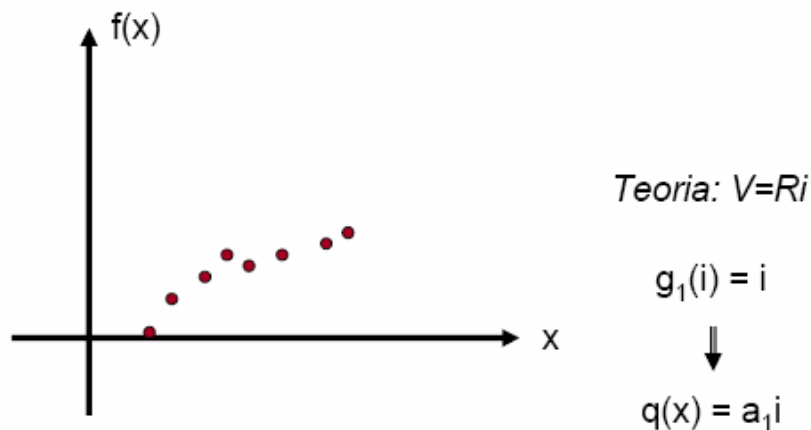
Problema: Como escolher as funções  $g_1(x), g_2(x), \dots, g_n(x)$ ?

A escolha das funções pode ser feita observando o gráfico dos pontos tabelados ou baseando-se em fundamentos teóricos dos experimentos que forneceu a tabela.



Exemplo:

Experiência onde foram medidos valores de corrente elétrica que passa por uma resistência submetida a várias tensões.



## 5.1 Método dos Mínimos Quadrados

O Método dos Mínimos Quadrados é um método bastante utilizado para ajustar uma determinada quantidade de pontos. Sua dedução será mostrada a seguir.

Considere dados  $m$  pontos  $(x_1, f(x_1)), (x_2, f(x_2)), \dots, (x_m, f(x_m))$  e as  $n$  funções  $g_1(x), g_2(x), \dots, g_n(x)$  escolhidas de alguma forma. Considere que o número de pontos tabelados  $m$  é sempre maior ou igual ao número de funções escolhidas  $n$  (ou ao número de coeficientes a determinar  $a_i$ ).

O objetivo é encontrar os coeficientes  $a_1, a_2, \dots, a_n$  tais que a função  $q(x) = a_1 g_1(x) + a_2 g_2(x) + \dots + a_n g_n(x)$  se **aproxime** ao máximo de  $f(x)$ . Seja  $d_k = f(x_k) - q(x_k)$  o desvio em  $x_k$ . Um conceito de proximidade é que  $d_k$  seja mínimo para todo  $k = 1, 2, \dots, m$ .

O Método dos Mínimos Quadrados consiste em escolher os  $a_i$ 's de tal forma que a soma dos quadrados dos desvios seja mínima.

$$S = \sum_{k=1}^m d_k^2 = \sum_{k=1}^m (f(x_k) - q(x_k))^2$$

Usando cálculo diferencial, sabe-se que para encontrar um ponto de mínimo de  $F(a_1, a_2, \dots, a_n)$ , é necessário achar inicialmente os pontos críticos (ou seja, todos os  $a_i$ 's).



$$\frac{\delta S}{\delta a_i} = 0$$

### 5.1.1 Ajuste Linear

Considere a função de ajuste dada por:

$$q(x) = a_1 + a_2 x$$

onde  $a_1$  e  $a_2$  são os coeficientes a serem determinados pelo método dos mínimos quadrados.

$$S = \sum_{k=1}^m d_k^2 = \sum_{k=1}^m (f(x_k) - q(x_k))^2$$

$$S = (f(x_1) - q(x_1))^2 + (f(x_2) - q(x_2))^2 + \dots + (f(x_m) - q(x_m))^2$$

$$S = (f(x_1) - (a_1 + a_2 x_1))^2 + \dots + (f(x_m) - (a_1 + a_2 x_m))^2$$

A condição de minimização é satisfeita se:

$$\frac{\delta S}{\delta a_1} = \frac{\delta S}{\delta a_2} = 0$$

$$\text{Para } \frac{\delta S}{\delta a_1} = 0:$$

$$2(f(x_1) - a_1 - a_2 x_1)(-1) + 2(f(x_2) - a_1 - a_2 x_2)(-1) + \dots + 2(f(x_m) - a_1 - a_2 x_m)(-1) = 0$$

$$\rightarrow m a_1 + a_2 \sum_{k=1}^m x_k = \sum_{k=1}^m f(x_k)$$

$$\text{Para } \frac{\delta S}{\delta a_2} = 0:$$

$$2(f(x_1) - a_1 - a_2 x_1)(-x_1) + 2(f(x_2) - a_1 - a_2 x_2)(-x_2) + \dots + 2(f(x_m) - a_1 - a_2 x_m)(-x_m) = 0$$

$$\rightarrow a_1 \sum_{k=1}^m x_k + a_2 \sum_{k=1}^m x_k^2 = \sum_{k=1}^m x_k f(x_k)$$

Com isso, obtém um sistema de equações lineares:

$$\begin{cases} ma_1 + a_2 \sum_{k=1}^m xk = \sum_{k=1}^m f(xk) \\ a_1 \sum_{k=1}^m xk + a_2 \sum_{k=1}^m xk^2 = \sum_{k=1}^m xk f(xk) \end{cases}$$

Também podendo ser representado na forma matricial:

$$\begin{bmatrix} m & \sum_{k=1}^m xk \\ \sum_{k=1}^m xk & \sum_{k=1}^m xk^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^m f(xk) \\ \sum_{k=1}^m xk f(xk) \end{bmatrix}$$

Exemplo:

Encontrar a melhor reta que ajusta os valores da tabela abaixo:

x	0,00	0,25	0,5	0,75	1,00
f(x)	1,0000	1,2840	1,6487	2,1170	2,7183

$$\sum_{k=1}^5 xk = 0 + 0,25 + 0,5 + 0,75 + 1 = 2,5$$

$$\sum_{k=1}^5 xk^2 = 0^2 + 0,25^2 + 0,5^2 + 0,75^2 + 1^2 = 1,875$$

$$\sum_{k=1}^5 f(xk) = 1 + 1,284 + 1,6487 + 2,117 + 2,7183 = 8,768$$

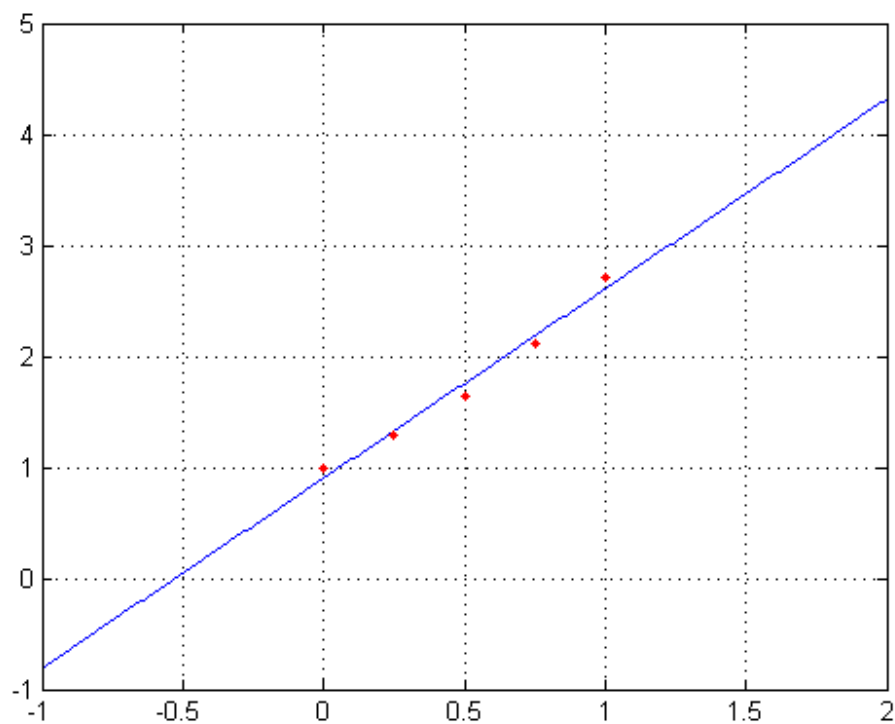
$$\begin{aligned} \sum_{k=1}^5 xk f(xk) &= 0 * 1 + 0,25 * 1,284 + 0,5 * 1,6487 + 0,75 * 2,117 + 1 * 2,7183 \\ &= 5,4514 \end{aligned}$$

$$\begin{bmatrix} 5 & 2,5 \\ 2,5 & 1,875 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 8,768 \\ 5,4514 \end{bmatrix} \rightarrow \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0,8997 \\ 1,7078 \end{bmatrix}$$

Logo, a função de ajuste é dada por:

$$q(x) = 0,8997 + 1,7078x$$

e seu gráfico é mostrado abaixo.



### 5.1.2 Ajuste Polinomial

O processo usado acima para cálculo da função para ajuste linear pode ser estendido para ajuste polinomial. Assim, uma função polinomial de grau  $(n-1)$  é dada por:

$$q(x) = a_1 + a_2x + a_3x^2 + \dots + a_n^{n-1}$$

onde os coeficientes  $a_i$  podem ser obtidos através da expansão do sistema utilizado no ajuste linear através do cálculo de

$$\frac{\partial S}{\partial a_i} = 0$$

agora para  $i = 1, \dots, n$ .

Essa expansão irá resultar no seguinte sistema de equações (n equações, n incógnitas):

$$\left\{ \begin{array}{l} ma_1 + a_2 \sum_{k=1}^m xk + a_3 \sum_{k=1}^m xk^2 + \dots + a_n \sum_{k=1}^m xk^{n-1} = \sum_{k=1}^m f(xk) \\ a_1 \sum_{k=1}^m xk + a_2 \sum_{k=1}^m xk^2 + a_3 \sum_{k=1}^m xk^3 + \dots + a_n \sum_{k=1}^m xk^n = \sum_{k=1}^m xk f(xk) \\ a_1 \sum_{k=1}^m xk^2 + a_2 \sum_{k=1}^m xk^3 + a_3 \sum_{k=1}^m xk^4 + \dots + a_n \sum_{k=1}^m xk^{n+1} = \sum_{k=1}^m xk^2 f(xk) \\ \vdots \\ a_1 \sum_{k=1}^m xk^{n-1} + a_2 \sum_{k=1}^m xk^n + a_3 \sum_{k=1}^m xk^{n+1} + \dots + a_n \sum_{k=1}^m xk^{2(n-1)} = \sum_{k=1}^m xk^{n-1} f(xk) \end{array} \right.$$

ou em notação matricial:

$$A a = b$$

$$\left[ \begin{array}{ccccc} m & \sum_{k=1}^m xk & \sum_{k=1}^m xk^2 & \dots & \sum_{k=1}^m xk^{n-1} \\ \sum_{k=1}^m xk & \sum_{k=1}^m xk^2 & \sum_{k=1}^m xk^3 & \dots & \sum_{k=1}^m xk^n \\ \sum_{k=1}^m xk^2 & \sum_{k=1}^m xk^3 & \sum_{k=1}^m xk^4 & \dots & \sum_{k=1}^m xk^{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^m xk^{n-1} & \sum_{k=1}^m xk^n & \sum_{k=1}^m xk^{n+1} & \dots & \sum_{k=1}^m xk^{2(n-1)} \end{array} \right] \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^m f(xk) \\ \sum_{k=1}^m xk f(xk) \\ \sum_{k=1}^m xk^2 f(xk) \\ \vdots \\ \sum_{k=1}^m xk^{n-1} f(xk) \end{bmatrix}$$

Note que a matriz A é simétrica, ou seja,  $A = A^T$ .

Exemplo:

Encontrar a melhor parábola que ajusta os valores da tabela abaixo:

x	0,00	0,25	0,5	0,75	1,00
f(x)	1,0000	1,2840	1,6487	2,1170	2,7183

Polinômio adotado: (n=3)

$$q(x) = a_1 + a_2x + a_3x^2$$

Calculando os termos da matriz A e do vetor b:

$$\sum_{k=1}^5 xk = 0 + 0,25 + 0,5 + 0,75 + 1 = 2,5$$

$$\sum_{k=1}^5 xk^2 = 0^2 + 0,25^2 + 0,5^2 + 0,75^2 + 1^2 = 1,875$$

$$\sum_{k=1}^5 xk^3 = 0^3 + 0,25^3 + 0,5^3 + 0,75^3 + 1^3 = 1,5625$$

$$\sum_{k=1}^5 xk^4 = 0^4 + 0,25^4 + 0,5^4 + 0,75^4 + 1^4 = 1,3828$$

$$\sum_{k=1}^5 f(xk) = 1 + 1,284 + 1,6487 + 2,117 + 2,7183 = 8,768$$

$$\sum_{k=1}^5 xk f(xk) = 0 * 1 + 0,25 * 1,284 + 0,5 * 1,6487 + 0,75 * 2,117 + 1 * 2,7183$$

$$= 5,4514$$

$$\sum_{k=1}^5 xk^2 f(xk) = 0^2 * 1 + 0,25^2 * 1,284 + 0,5^2 * 1,6487 + 0,75^2 * 2,117 + 1^2$$

$$* 2,7183 = 4,4015$$

Montando o sistema linear, encontra-se o seguinte sistema matricial:

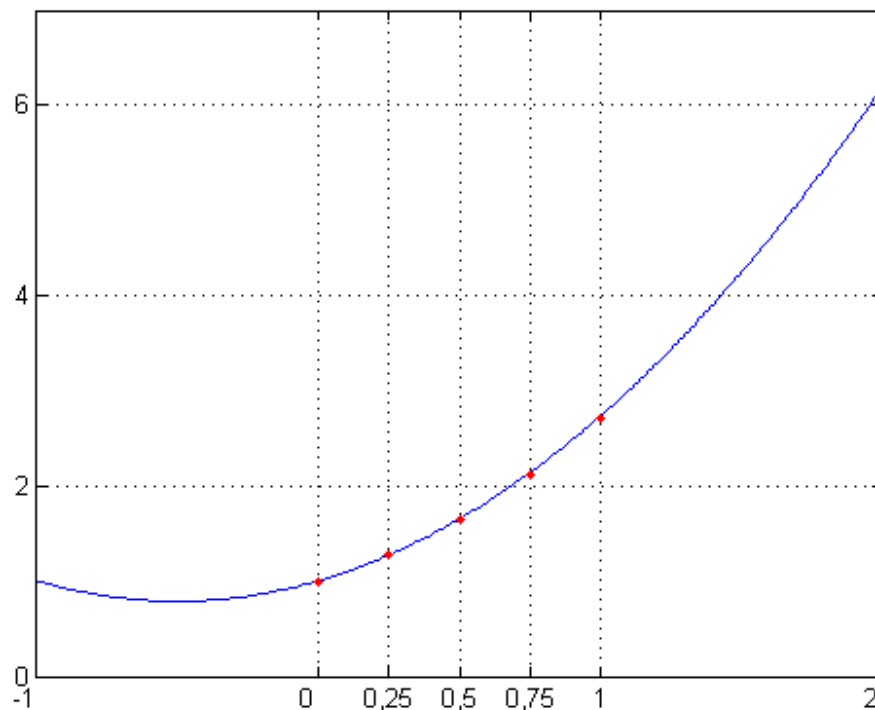
$$\begin{bmatrix} 5 & 2,5 & 1,875 \\ 2,5 & 1,875 & 1,5625 \\ 1,875 & 1,5625 & 1,3828 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 8,768 \\ 5,4514 \\ 4,4015 \end{bmatrix}$$

Resolvendo o sistema acima, encontra-se a seguinte solução para o problema:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1,0051 \\ 0,8647 \\ 0,8432 \end{bmatrix}$$

Ou seja, a parábola que melhor ajusta os pontos fornecidos tem equação:

$$q(x) = 1,0051 + 0,8647x + 0,8432x^2$$



### 5.1.3 Linearização

Algumas funções de duas constantes podem ser linearizadas antes da aplicação do método dos mínimos quadrados, com o objetivo de obter o sistema de equações visto anteriormente. O procedimento varia de acordo com o tipo de função:

#### 5.1.3.1 Função Exponencial

$$y = a e^{bx}$$

Aplicando logaritmo em ambos os lados, tem-se:

$$\ln y = \ln(a e^{bx}) = \ln(a) + bx$$

Então, se fizermos:

$$y' = \ln y \quad ; \quad a_1 = \ln(a) \quad ; \quad a_2 = b$$

Encontra-se a seguinte expressão:

$$y' = a_1 + a_2 x$$

que nada mais é senão uma reta. Daí o nome linearização.

### 5.1.3.2 Função Logarítmica

$$y = a \ln(bx)$$

A função pode ser expandida para:

$$y = a \ln(b) + a \ln(x)$$

Logo, se fizermos:

$$y' = y \quad ; \quad x' = \ln(x) \quad ; \quad a_1 = a \ln(b) \quad ; \quad a_2 = a$$

encontramos a linearização:

$$y' = a_1 + a_2 x'$$

### 5.1.3.3 Função Potencial

$$y = a x^b$$

Aplicando logaritmo em ambos os lados:

$$\ln y = \ln(a x^b) = \ln(a) + \ln(x^b) = \ln(a) + b \ln(x)$$

Com as seguintes hipóteses:

$$y' = \ln(y) \quad ; \quad x' = \ln(x) \quad ; \quad a_1 = \ln(a) \quad ; \quad a_2 = b$$

encontra-se a expressão:

$$y' = a_1 + a_2 x'$$

### 5.1.3.4 Função Hiperbólica

$$y = a + \frac{b}{x}$$

Fazendo:

$$y' = y \quad ; \quad x' = \frac{1}{x} \quad ; \quad a_1 = a \quad ; \quad a_2 = b$$

Tem-se também:

$$y' = a_1 + a_2 x'$$

Exemplo:

Encontrar a melhor função que ajusta os valores da tabela abaixo:

x	-1	-0,7	-0,4	-0,1	0,2	0,5	0,8	1,0
y	36,547	17,264	8,155	3,852	1,82	0,86	0,406	0,246

Sugestão: Utilizar uma função exponencial.

Solução:

Como vamos ajustar os pontos por uma função exponencial precisamos fazer a seguinte adaptação:

$$y' = \ln y$$

Ou seja, a coordenada y de cada ponto deverá ser substituída por seu logaritmo, logo:

y'	3,5986	2,8486	2,0986	1,3486	0,5988	-0,1508	-0,9014	-1,4024
----	--------	--------	--------	--------	--------	---------	---------	---------

Então faz-se um ajuste linear dos pontos de abscissa x e ordenada y', obtendo-se os seguintes valores para os coeficientes da reta:



$$\begin{cases} a_1 = 1,0986 \\ a_2 = -2,5002 \end{cases}$$

Para adaptar esses valores, coeficientes da reta, para a função exponencial, ainda basta fazer as seguintes adaptações:

$$a_1 = \ln(a) \quad ; \quad a_2 = b$$

Logo,

$$a = e^{a_1} \quad ; \quad b = a_2$$

E então, calcula-se os valores de a e b:

$$a = e^{1,0986} = 3,00$$

$$b = -2,5002$$

Então, a função exponencial que melhor ajusta os pontos fornecidos no exemplo é:

$$q(x) = 3,00 e^{-2,5002 x}$$

#### 5.1.4 Qualidade do Ajuste

Uma forma de avaliar a qualidade do ajuste é através do coeficiente de correlação de Pearson r. Este coeficiente pode ser calculado pela seguinte expressão:

$$r = \frac{\sum_{i=1}^m [(y_i - \bar{y}) \cdot (q_i - \bar{q})]}{\sqrt{\sum_{i=1}^m (y_i - \bar{y})^2} \cdot \sqrt{\sum_{i=1}^m (q_i - \bar{q})^2}}$$

onde,

$$\bar{y} = \frac{1}{m} \sum_{i=1}^m y_i \quad ; \quad \bar{q} = \frac{1}{m} \sum_{i=1}^m q_i$$

Este coeficiente, assume apenas valores entre -1 e 1.

- $r = 1$ , significa uma correlação perfeita positiva entre as duas variáveis;
- $r = -1$ , significa uma correlação negativa perfeita entre as duas variáveis, isto é, se uma aumenta, a outra sempre diminui;

- $r=0$ , significa que as duas variáveis não dependem linearmente uma da outra. No entanto, pode existir uma outra dependência que seja "não linear". Assim, o resultado  $r=0$  deve ser investigado por outros meios.

## Algoritmo

Verifica tipo\_de\_ajuste;

Caso tipo\_de\_ajuste seja Exponencial

Para  $i = 1, \dots, m$

$y_i = \ln Y_i$

Fim

Fazer Ajuste Linear com  $x, y$  retornando coeficientes  $s_1$  e  $s_2$

$a = e^{s_1}$

$b = s_2$

Caso tipo\_de\_ajuste seja Logarítmico

Para  $i = 1, \dots, m$

$x_i = \ln x_i$

Fim

Fazer Ajuste Linear com  $x, y$  retornando coeficientes  $s_1$  e  $s_2$

$a = s_2$

$b = e^{s_1} / a$

Caso tipo\_de\_ajuste seja Potencial

Para  $i = 1, \dots, m$

$y_i = \ln y_i$

$x_i = \ln x_i$

Fim

Fazer Ajuste Linear com  $x, y$  retornando coeficientes  $s_1$  e  $s_2$

$a = e^{s_1}$

$b = s_2$

Caso tipo\_de\_ajuste seja Hiperbólico

Para  $i = 1, \dots, m$

$x_i = 1/x_i$

Fim

Fazer Ajuste Linear com  $x, y$  retornando coeficientes  $s_1$  e  $s_2$

$a = s_1$

$b = s_2$

Caso tipo\_de\_ajuste seja Polinomial (polinômio de grau  $n-1$ )

Para  $i = 1, \dots, n$

Para  $j = i, \dots, n$

$A_{ij} = 0$

Para  $k = 1, \dots, m$

$A_{ij} = A_{ij} + x_k^{(i+j-2)}$

Fim

$A_{ji} = A_{ij}$

```

Fim
bi = 0
Para k = 1,...,m
    bi = bi + yk*xk^(i-1)
Fim
Fim
s = (A^-1) . b
Fim

```

## 6 SISTEMA DE EQUAÇÕES NÃO LINEARES

Dada uma função não linear

$$F: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad F = (f_1, f_2, \dots, f_n)^T$$

o objetivo é encontrar as soluções para

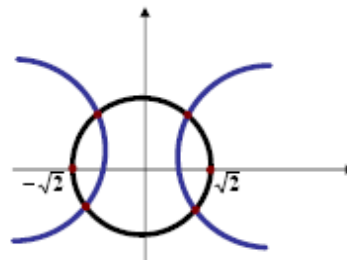
$$F(X) = 0$$

ou seja,

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, \dots, x_n) = 0 \end{cases}$$

Por exemplo:

$$\begin{cases} f_1(x_1, x_2) = x_1^2 + x_2^2 - 2 = 0 \\ f_2(x_1, x_2) = x_1^2 + \frac{x_2^2}{9} - 1 = 0 \end{cases}$$



Este sistema não linear admite quatro soluções, representadas pelos pontos onde as curvas se interceptam.

### 6.1 Notação Utilizada

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad e \quad F(X) = \begin{bmatrix} f_1(X) \\ f_2(X) \\ \vdots \\ f_n(X) \end{bmatrix}$$

Cada função  $f_i(X)$  é uma função não linear em  $X$  e portanto  $F(X)$  também é uma função não linear em  $X$ .

Para sistemas lineares, tínhamos:

$$F(X) = AX - b, \quad A \in \mathbb{R}^{m \times n}$$

## 6.2 Considerações

- $F(X)$  tem derivadas contínuas no domínio;
- Existe pelo menos um ponto  $X^* \in D$ , tal que  $F(X^*) = 0$ .

O vetor das derivadas parciais da função  $f_i(X)$  é denominado *vetor gradiente* de  $f_i(X)$  e é denotado por:

$$\nabla f_i(X) = \left( \frac{\partial f_i(X)}{\partial x_1} \quad \frac{\partial f_i(X)}{\partial x_2} \quad \dots \quad \frac{\partial f_i(X)}{\partial x_n} \right)^T$$

A matriz das derivadas parciais de  $F(X)$  é chamada *matriz Jacobiana*  $J(X)$ :

$$J(X) = \begin{bmatrix} \nabla f_1(X)^T \\ \nabla f_2(X)^T \\ \vdots \\ \nabla f_n(X)^T \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1(X)}{\partial x_1} & \frac{\partial f_1(X)}{\partial x_2} & \dots & \frac{\partial f_1(X)}{\partial x_n} \\ \frac{\partial f_2(X)}{\partial x_1} & \frac{\partial f_2(X)}{\partial x_2} & \dots & \frac{\partial f_2(X)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(X)}{\partial x_1} & \frac{\partial f_n(X)}{\partial x_2} & \dots & \frac{\partial f_n(X)}{\partial x_n} \end{bmatrix}$$

Exemplo:

Determinar a matriz Jacobiana do sistema não linear abaixo:

$$F(X) = \begin{cases} x_1^3 - 3x_1x_2^2 + 1 = 0 \\ 3x_1^2x_2 - x_2^3 = 0 \end{cases}$$

$$J(X) = \begin{bmatrix} 3x_1^2 - 3x_2^2 & -6x_1x_2 \\ 6x_1x_2 & 3x_1^2 - 3x_2^2 \end{bmatrix}$$

### 6.3 Características dos Métodos para Resolução dos Sistemas de Equações não Lineares

- Iteratividade

A partir de um ponto inicial  $X^0$ , geram sequências  $X^K$ . Na situação de convergência,  $X^*$  é uma das soluções do sistema quando:

$$\lim_{K \rightarrow \infty} X^K = X^*$$

- Existência de critérios de convergência

- Verificar se  $F(X^K)$  tem módulo pequeno. Ou seja:

$$\|F(X^K)\| < \varepsilon$$

- Verificar se  $\|X^{K+1} - X^K\|$  está próximo de zero. Ou seja:

$$\|X^{K+1} - X^K\| < \varepsilon$$

- Limitar o número de iterações K por um número máximo de iterações.

### 6.4 Métodos Numéricos

Veremos aqui basicamente três tipos de métodos numéricos para a resolução de sistemas de equações não lineares. Os métodos serão descritos e caracterizados a seguir.

#### 6.4.1 Método de Newton-Raphson

Este é o método mais amplamente utilizado para resolver sistemas de equações lineares. O método combina duas idéias básicas comuns nas aproximações numéricas:

- Linearização

Procura-se substituir, numa certa vizinhança, um problema complicado por sua aproximação linear. Essa aproximação pode ser obtida, por exemplo, tomando-se os primeiros termos de uma expansão usando Série de Taylor.

- Iteração

Devido à repetição do procedimento, até que se garanta a convergência para a solução do sistema ou o fim desejado.

### 6.4.1.1 Caso Escalar

Para ilustrar mais facilmente o uso do método de Newton para a solução de sistemas de equações não lineares, considere um sistema com uma incógnita e uma única equação:

$$f(x) = 0$$

Expandindo essa equação usando série de Taylor próximo a um ponto inicial  $(x_1, f(x_1))$  e tomando-se apenas os primeiros termos desta expansão (linearização), tem-se:

$$f(x) = f(x_1) + f'(x_1) \cdot (x - x_1)$$

onde  $f'(x_1)$  é a primeira derivada de  $f$  em  $x_1$ .

Igualando a equação anterior a zero e desenvolvendo-a, tem-se:

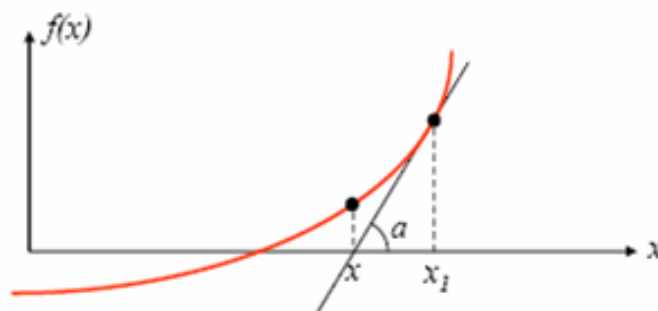
$$f(x) = 0 \rightarrow f(x_1) + f'(x_1) \cdot (x - x_1) = 0$$

$$x - x_1 = -\frac{f(x_1)}{f'(x_1)} \rightarrow x = x_1 - \frac{f(x_1)}{f'(x_1)}$$

Pensando no processo iterativo:

$$x^{K+1} = x^K - \frac{f(x^K)}{f'(x^K)}$$

Graficamente, temos:



Tomando a tangente à curva em  $x_1$ , tem-se que:

$$\tan \alpha = \frac{f(x_1)}{x_1 - x} \rightarrow \tan \alpha = f'(x_1) \rightarrow f'(x_1) = \frac{f(x_1)}{x_1 - x} \rightarrow x = x_1 - \frac{f(x_1)}{f'(x_1)}$$

E para uma iteração  $k$  qualquer:

$$x^{K+1} = x^K - \frac{f(x^K)}{f'(x^K)}$$

#### 6.4.1.2 Caso Vetorial

Considere agora o sistema mostrado inicialmente.

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, \dots, x_n) = 0 \end{cases} \quad \text{ou} \quad \begin{cases} f_1(X) = 0 \\ f_2(X) = 0 \\ \vdots \\ f_n(X) = 0 \end{cases}$$

Usando o mesmo raciocínio do caso escalar, tem-se que:

$$\begin{cases} f_1(X) = f_1(X^1) + (x_1 - x_1^1) \cdot \frac{\partial f_1(X^1)}{\partial x_1} + \dots + (x_n - x_n^1) \cdot \frac{\partial f_1(X^1)}{\partial x_n} = 0 \\ f_2(X) = f_2(X^1) + (x_1 - x_1^1) \cdot \frac{\partial f_2(X^1)}{\partial x_1} + \dots + (x_n - x_n^1) \cdot \frac{\partial f_2(X^1)}{\partial x_n} = 0 \\ \vdots \\ f_n(X) = f_n(X^1) + (x_1 - x_1^1) \cdot \frac{\partial f_n(X^1)}{\partial x_1} + \dots + (x_n - x_n^1) \cdot \frac{\partial f_n(X^1)}{\partial x_n} = 0 \end{cases}$$

onde o índice superior 1 no vetor  $X$  indica a iteração:

$$X^1 = [x_1^1 \quad x_2^1 \quad \dots \quad x_n^1]^T$$

Rearranjando o sistema, colocando-o na forma matricial, tem-se:

$$\begin{bmatrix} f_1(X) \\ f_2(X) \\ \vdots \\ f_n(X) \end{bmatrix} = \begin{bmatrix} f_1(X^1) \\ f_2(X^1) \\ \vdots \\ f_n(X^1) \end{bmatrix} + \underbrace{\begin{bmatrix} \frac{\partial f_1(X^1)}{\partial x_1} & \dots & \frac{\partial f_1(X^1)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(X^1)}{\partial x_1} & \dots & \frac{\partial f_n(X^1)}{\partial x_n} \end{bmatrix}}_{\text{Matriz Jacobiana}} \begin{bmatrix} x_1 - x_1^1 \\ x_2 - x_2^1 \\ \vdots \\ x_n - x_n^1 \end{bmatrix} = 0$$

Reescrevendo,

$$[J] \cdot \{X - X^1\} + \{f(X^1)\} = 0 \quad \rightarrow \quad [J] \cdot \{X - X^1\} = -\{f(X^1)\}$$

Multiplicando a equação acima pelo inverso da matriz Jacobiana, tem-se:

$$[J]^{-1} [J] \{X - X^1\} = -[J]^{-1} \{f(X^1)\} \quad \rightarrow \quad \{X - X^1\} = -[J]^{-1} \{f(X^1)\}$$

Generalizando para uma iteração k qualquer, temos:

$$\{X^{k+1}\} = \{X^k\} - [J]^{-1} \cdot \{f(X^k)\}$$

Porém, como o processo de inversão é muito caro computacionalmente, opta-se por resolver o sistema de equações lineares abaixo para obter a sua solução.

$$[J]\{X^{k+1} - X^k\} = -\{f(X^k)\}$$

Exemplo:

Aplicar o método de Newton à resolução do sistema não linear  $F(X) = 0$ , onde:

$$F(X) = \begin{bmatrix} x_1 + x_2 - 3 \\ x_1^2 + x_2^2 - 9 \end{bmatrix}$$

considerando tolerância  $\varepsilon = 10^{-4}$ , número máximo de iterações  $k(\max) = 2$  e chute inicial  $X^1 = [1 \ 5]^T$ .

Solução:

Para  $k = 1$  (Primeira iteração)

$$F(X^1) = [3 \ 17]^T \quad \rightarrow \quad \|F(X^1)\| = 17.2627 > \varepsilon \quad ; \quad J(X^1) = \begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix} \Delta X^1 = -\begin{bmatrix} 3 \\ 17 \end{bmatrix} \quad \rightarrow \quad \Delta X^1 = \begin{bmatrix} -1.625 \\ -1.375 \end{bmatrix} \quad \therefore \quad X^2 = X^1 + \Delta X^1$$

$$X^2 = \begin{bmatrix} -0.625 \\ 3.625 \end{bmatrix}$$



Para  $k = 2$  (Segunda iteração)

$$F(X^2) = [0 \quad 4,5313]^T \rightarrow \|F(X^2)\| = 4,5313 > \varepsilon \quad ; \quad J(X^2) = \begin{bmatrix} 1 & 1 \\ -1,25 & 7,25 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ -1,25 & 7,25 \end{bmatrix} \Delta X^2 = - \begin{bmatrix} 0 \\ 4,5313 \end{bmatrix} \rightarrow \Delta X^2 = \begin{bmatrix} 0,5331 \\ -0,5331 \end{bmatrix} \quad \therefore \quad X^3 = X^2 + \Delta X^2$$

$$X^2 = \begin{bmatrix} -0,0919 \\ 3,0919 \end{bmatrix}$$

**Solução exata:**  $X^* = [3 \quad 0]^T$  ou  $X^* = [0 \quad 3]^T$

Algoritmo:

Enquanto ( $\|F(X^k)\| < \varepsilon$ ) e ( $k < k(\max)$ )

$k = k + 1$

Calcular  $F(X^k)$  e  $J(X^k)$

$Dx = -[J]^{-1} \cdot \{f(X^k)\}$

$X^{k+1} = X^k + Dx$

Fim

## 6.4.2 Métodos Quasi-Newton

O método de Newton apresenta três características importantes que influenciam na velocidade de convergência:

- Escolha do ponto inicial (chute inicial);
- Cálculo do Jacobiano (derivadas);
- Solução do Sistema Linear.

Vários métodos encontrados na literatura apresentam alternativas para o cálculo do Jacobiano, tornando-os úteis para a solução dos sistemas de equações não lineares. Esses métodos são conhecidos como Métodos *Quasi-Newton*.

Entre esses métodos estão o Método de *Newton-Raphson modificado* e o Método da *Secante*, que serão descritos a seguir.

### 6.4.2.1 Método de Newton-Raphson Modificado

Este método consiste em tomar, em cada iteração  $k$ , a mesma matriz Jacobiana computada no passo inicial. Desse modo,

$$\{X^{k+1} - X^k\} = -[J(X^1)]^{-1} \cdot \{f(X^k)\}$$

Apesar da redução do custo computacional, este método pode ser mais sensível à convergência, ou seja, o número de iterações necessárias geralmente é maior que quando se usa o método de Newton-Raphson.

Exemplo:

Aplicar o método de Newton Modificado à resolução do sistema não linear  $F(X) = 0$ , onde:

$$F(X) = \begin{bmatrix} x_1 + x_2 - 3 \\ x_1^2 + x_2^2 - 9 \end{bmatrix}$$

considerando tolerância  $\varepsilon = 10^{-4}$ , número máximo de iterações  $k(\max) = 2$  e chute inicial  $X^1 = [1 \ 5]^T$ .

$$J(X^1) = \begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix}$$

Para  $k = 1$  (Primeira iteração)

$$F(X^1) = [3 \ 17]^T \rightarrow \|F(X^1)\| = 17.2627 > \varepsilon$$

$$\begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix} \Delta X^1 = - \begin{bmatrix} 3 \\ 17 \end{bmatrix} \rightarrow \Delta X^1 = \begin{bmatrix} -1,625 \\ -1,375 \end{bmatrix} \therefore X^2 = X^1 + \Delta X^1$$

$$X^2 = \begin{bmatrix} -0,625 \\ 3,625 \end{bmatrix}$$

Para  $k = 2$  (Segunda iteração)

$$F(X^2) = [0 \ 4,5313]^T \rightarrow \|F(X^2)\| = 4,5313 > \varepsilon$$

$$\begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix} \Delta X^2 = - \begin{bmatrix} 0 \\ 4,5313 \end{bmatrix} \rightarrow \Delta X^2 = \begin{bmatrix} 0,5664 \\ -0,5664 \end{bmatrix} \therefore X^3 = X^2 + \Delta X^2$$

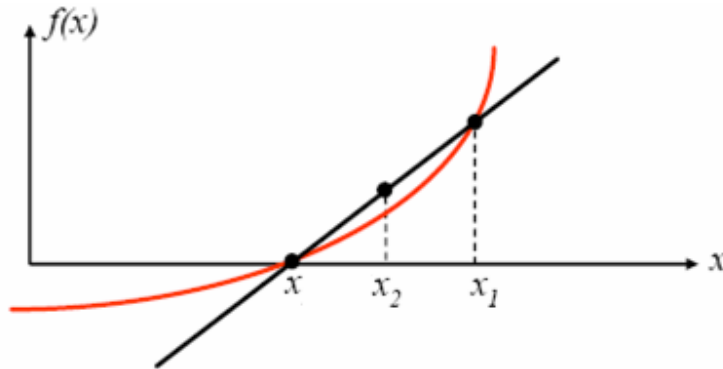
$$X^2 = \begin{bmatrix} -0,0586 \\ 3,0586 \end{bmatrix}$$

### 6.4.2.2 Método Secante

Este método consiste em calcular as derivadas da matriz Jacobiana de forma aproximada:

$$\frac{\partial f_i}{\partial x_j} \sim \frac{f_i(x_1, x_2, \dots, x_j + h, \dots, x_n) - f_i(x_1, x_2, \dots, x_j, \dots, x_n)}{h}$$

Para o caso escalar, tem-se graficamente:



Por semelhança de triângulos:

$$\frac{x_2 - x}{f(x_2)} = \frac{x_1 - x}{f(x_1)} \quad \rightarrow \quad x = \frac{f(x_2) \cdot x_1 - f(x_1) \cdot x_2}{f(x_2) - f(x_1)}$$

Estendendo para uma iteração qualquer k:

$$x^{k+1} = \frac{f(x^k) \cdot x^{k-1} - f(x^{k-1}) \cdot x^k}{f(x^k) - f(x^{k-1})}$$

ou ainda, utilizando expansão por série de Taylor:

$$x^{k+1} = x^k - \frac{f(x^k)}{\frac{f(x^k) - f(x^{k-1})}{x^k - x^{k-1}}}$$

Note que a equação acima é bastante semelhante a de Newton Raphson:

$$X^{K+1} = X^K - [J]^{-1} \cdot F(X^K)$$

se fizermos a aproximação da secante para a matriz Jacobiana.

### 6.4.3 Outros Métodos

Outros métodos bastante conhecidos para solução numérica de sistemas de equações não lineares são: *BFGS*, *DFT*, *Gradiente Conjugado*, *Máximo Declive* e *Fletcher-Rivers*.