



Universidade de Brasília  
Instituto de Ciências Exatas  
Departamento de Matemática  
Programa de Mestrado Profissional em  
Matemática em Rede Nacional



# Métodos de Resolução de Sistemas Lineares e Custo Computacional

Felipe Torres Vital

Brasília  
2014

Felipe Torres Vital

# Métodos de Resolução de Sistemas Lineares e Custo Computacional

Trabalho de Conclusão de Curso apresentado ao  
Programa de Pós-Graduação do Departamento de  
Matemática da Universidade de Brasília, como  
parte dos requisitos para obtenção do grau de  
Mestre em Matemática.

**Área de Concentração:** Matemática do Ensino Básico  
**Orientador:** Prof. Dr. Adail de Castro Cavalheiro

Brasília  
2014

Todos os direitos reservados. É proibida a reprodução total ou parcial deste trabalho sem a autorização da universidade, do autor e do orientador.

**Felipe Torres Vital** graduou-se em Matemática pela UnB - Universidade de Brasília.

## Agradecimentos

Agradeço à minha família - meus pais, meu irmão e minha querida esposa, pelo amor, incentivo, apoio incondicional e força para superar as dificuldades.

Aos meus professores, do Ensino Fundamental ao Ensino Superior, que me concederam a chance de estar onde estou agora.

Aos meus amigos, fonte de força e descontração, durante toda esta estrada.

Ao meu orientador, pelo suporte no pouco tempo que lhe coube, pelas suas correções e incentivos.

A CAPES, pelo suporte financeiro.

Enfim, a todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

## **Resumo**

O presente trabalho tem como objetivo indicar o método mais eficaz para resolução de um determinado sistema linear. Os sistemas lineares são utilizados em cálculos nas mais diversas áreas, tais como a Física, a Engenharia, a Geometria e a Probabilidade. Apresentamos, através de exemplos contextualizados, seis métodos distintos, quatro diretos e dois iterativos, a fim de estudar qual possui o menor custo computacional. Demonstramos que não existe o melhor método de forma absoluta. Assim, deve ser feita uma análise prévia do que se busca com a solução para a escolha do método de resolução. Veremos que existem métodos mais rápidos e outros mais precisos.

### **Palavras-chaves**

Sistemas lineares, métodos de resolução, custo computacional, métodos diretos, métodos iterativos.

## **Abstract**

The following paper has as its purpose to indicate the most effective method when solving a determined linear system. Linear systems are used in calculations on many different areas, such as Physics, Engineering, Geometry and Probability. We introduce, through contextualized examples, six distinct methods, four direct and two iterative, in order to study which one has the lowest operational cost. We demonstrate that there is not an absolute best method. Therefore, there must be a previous analysis on what you seek with the solution before you choose your solving method. We will see that there are faster methods and more precise ones.

### **Keywords**

Linear systems, solving methods, operational cost, direct methods, iterative methods.

# Sumário

<b>1</b>	<b>Preliminares</b>	<b>11</b>
1.1	Sistemas Lineares . . . . .	11
1.2	Matrizes e Determinantes . . . . .	11
1.2.1	Matriz Transposta . . . . .	13
1.3	Sequências Convergentes . . . . .	13
1.4	Conjunto-Solução e Equivalência de Sistemas . . . . .	14
<b>2</b>	<b>Métodos Diretos de Resolução</b>	<b>18</b>
2.1	Substituição . . . . .	19
2.2	Eliminação Gaussiana . . . . .	24
2.3	Fatoração LU . . . . .	30
2.4	Regra de Cramer . . . . .	35
2.5	Comparando os Métodos Diretos . . . . .	38
2.5.1	Substituição x Eliminação Gaussiana . . . . .	39
2.5.2	Substituição x Fatoração LU . . . . .	39
2.5.3	Fatoração LU x Eliminação Gaussiana . . . . .	40
<b>3</b>	<b>Métodos Iterativos de Resolução</b>	<b>41</b>
3.1	Método de Gauss-Jacobi . . . . .	44
3.2	Método de Gauss-Seidel . . . . .	49
3.3	Cálculo de Iterações através de uma Planilha Eletrônica . . . . .	57
3.4	Comparando os Métodos Iterativos . . . . .	59
3.5	Limitação dos Métodos Iterativos . . . . .	60
<b>4</b>	<b>Custo Computacional Comparado</b>	<b>64</b>
4.1	Temperatura em um Ponto de uma Placa . . . . .	64
4.1.1	Resolvendo por Eliminação Gaussiana . . . . .	66
4.1.2	Resolvendo por Gauss-Seidel . . . . .	68
4.2	Eliminação Gaussiana x Gauss-Seidel . . . . .	70

## Introdução

Vejam a seguinte situação, comumente exposta a alunos do Ensino Fundamental, e a seguir, tentemos resolver o problema:

*Em um parque de diversões há dois preços de ingressos: crianças pagam 18 reais e adultos pagam 25 reais. Joana e seu marido levaram os filhos e alguns amigos a esse parque e compraram 7 ingressos. No total, gastaram 147 reais. Quantos ingressos de cada tipo foram comprados?*

Denominemos por  $x$  o número de ingressos comprados para crianças e por  $y$  o número de ingressos comprados para adultos.

Se tomarmos como base somente a informação de que foram comprados um total de 7 ingressos, podemos concluir que  $x + y = 7$ . Entretanto, não há como termos certeza de quantos de cada tipo foram comprados. Teríamos como respostas possíveis, na forma  $(x, y)$ , os pares ordenados  $(0, 7)$ ,  $(1, 6)$ ,  $(2, 5)$ ,  $(3, 4)$ ,  $(4, 3)$ ,  $(5, 2)$ ,  $(6, 1)$  e  $(7, 0)$ . Assim, concluímos que a primeira informação não basta para resolvermos o problema.

Entretanto, sabemos também que o casal gastou 147 reais no total. Como foram comprados  $x$  ingressos para criança, concluímos que o casal gastou  $18x$  reais com esse tipo de ingresso. Analogamente, foram gastos  $25y$  reais com ingressos para adultos. Logo, temos que

$$18x + 25y = 147.$$

Vamos aplicar as possíveis respostas obtidas com a primeira informação, para ver quais dariam certo:

$$\begin{aligned} 18(0) + 25(7) &= 0 + 175 = 175; \\ 18(1) + 25(6) &= 18 + 150 = 168; \\ 18(2) + 25(5) &= 36 + 125 = 161; \\ 18(3) + 25(4) &= 54 + 100 = 154; \\ 18(4) + 25(3) &= 72 + 75 = 147; \\ 18(5) + 25(2) &= 90 + 50 = 140; \\ 18(6) + 25(1) &= 108 + 25 = 133; \\ 18(7) + 25(0) &= 126 + 0 = 126. \end{aligned}$$

Somente o par ordenado  $(4, 3)$  satisfaz as duas equações obtidas simultaneamente. Podemos concluir assim que foram comprados 4 ingressos para crianças e 3 ingressos para adultos.

Para resolver o problema acima, resolvemos simultaneamente as equações

$$x + y = 7 \quad \text{e} \quad 18x + 25y = 147.$$

A esse conjunto de equações envolvendo um número  $n$  de variáveis, damos o nome de **sistema de equações**. Como todas as variáveis das equações são termos lineares

(expoente 1), e não existe multiplicação entre variáveis, chamamos o sistema de **sistema de equações lineares**, ou, simplesmente, **sistema linear**.

A presença de uma equação aos moldes de  $xy + x^2 = 3$  no sistema, por exemplo, o caracterizaria como um sistema não-linear, por ter uma multiplicação de  $x$  e  $y$  no primeiro termo e um expoente 2 na incógnita  $x$  no segundo termo. Trabalharemos, nesta dissertação, apenas com sistemas lineares.

Foi apresentada uma forma simples de resolver um sistema linear com 2 equações e 2 incógnitas. Não houve dificuldade pois, no exemplo acima, trabalhou-se com inteiros positivos. Agora imaginemos como seria resolver um sistema com 10 equações e 10 incógnitas. Seria simples assim obter uma solução, se estivéssemos trabalhando com números reais?

O intuito deste trabalho é mostrar diferentes formas com que podemos solucionar um sistema linear. Além disso, queremos analisar os prós e contras de cada método, na tentativa de determinar o mais vantajoso.

No **primeiro capítulo**, veremos alguns *conceitos matemáticos* que serão importantes nesta caminhada. No **segundo**, analisaremos os *métodos diretos* de resolução de um sistema linear, contabilizando o custo operacional de cada um dos métodos, e no **terceiro**, faremos o mesmo para os *métodos iterativos*. Por fim, no **quarto capítulo**, será feita a *comparação* entre os dois tipos de métodos.

Durante todo o texto, há exemplos de sistemas em contextos de diferentes áreas, como por exemplo a Física, a Engenharia, a Geometria e a Probabilidade.

Ao final, queremos que o leitor possa ter a condição de, ao se deparar com um problema possível de se equacionar na forma de um sistema linear, solucioná-lo da melhor forma possível para a sua intenção.

# 1 Preliminares

Antes de começar a análise dos métodos de resolução de sistemas lineares, devemos abordar alguns assuntos que serão importantes no processo de entendimento do conteúdo deste trabalho.

Veremos um pouco da teoria por trás dos sistemas lineares, além de assuntos correlatos de outras áreas da Matemática.

## 1.1 Sistemas Lineares

Um sistema linear com  $m$  equações e  $n$  variáveis ( $m \times n$ ) pode ser escrito da seguinte forma:

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

Na notação acima,  $x_k$ ,  $k \in \{1, \dots, n\}$ , corresponde as variáveis, ou incógnitas, do sistema, enquanto  $a_{ij}$ ,  $i \in \{1, \dots, m\}$ ,  $j \in \{1, \dots, n\}$ , corresponde ao coeficiente da variável  $x_j$  na  $i$ -ésima equação. Por fim,  $b_t$ ,  $t \in \{1, \dots, m\}$ , corresponde ao termo independente da  $t$ -ésima equação.

**Observação 1.1.** *Equações onde, para um  $i$  fixo,  $a_{ij} = 0$ ,  $j \in \{1, \dots, n\}$ , não serão analisadas neste trabalho, pois:*

*Se  $b_i = 0$ , então qualquer ponto satisfaria tal equação. De fato,  $a_{ij}x_j = 0$ ,  $j \in \{1, \dots, n\}$ , logo  $a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = 0 = b_i$ , qualquer que seja a lista ordenada  $(x_1, x_2, \dots, x_n)$ ;*

*Se  $b_i \neq 0$ , então tal equação não tem solução. De fato,  $a_{ij}x_j = 0$ ,  $j \in \{1, \dots, n\}$ , logo  $a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = 0 \neq b_i$ , qualquer que seja a lista ordenada  $(x_1, x_2, \dots, x_n)$ .*

## 1.2 Matrizes e Determinantes

Além da forma usual de um sistema linear, vista na introdução, podemos escrevê-lo na sua forma matricial, separando seus coeficientes das incógnitas e dos termos independentes.

Assim, temos a representação

$$A \times x = b,$$

onde

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

**Observação 1.2.** De fato, realizando a multiplicação das matrizes  $A$  e  $x$  e igualando a  $b$ , obteremos o formato original do sistema.

A definição que daremos em seguida tem a sua validade apenas para as matrizes as quais chamamos de **quadradas**. Essas matrizes possuem o mesmo número de linhas e colunas, ou seja, são da forma  $n \times n$ .

Assim, a cada matriz quadrada, é associado um número real, chamado **determinante**. Uma das mais importantes aplicações dos determinantes é a *Regra de Cramer*, método comumente utilizado no Ensino Médio para a resolução de sistemas lineares, e que analisaremos posteriormente neste trabalho. Vejamos a seguir como é feito o cálculo do determinante de uma matriz. Para isso, vamos partir do seguinte:

**Definição 1.1.** Uma **permutação** de um conjunto finito  $X$  é um rearranjo na ordem dos elementos desse conjunto, sem repetições e não omitindo nenhum dos elementos.

**Exemplo 1.1.**  $I_3 = \{1, 2, 3\}$  tem, no seu conjunto de permutações, os elementos  $(1, 2, 3)$  (permutação identidade: elementos em sua ordem inicial),  $(1, 3, 2)$ ,  $(2, 1, 3)$ ,  $(2, 3, 1)$ ,  $(3, 1, 2)$  e  $(3, 2, 1)$ , configurando um total de 6 elementos.

O número de permutações do conjunto  $I_n = \{1, 2, \dots, n\}$  é dado por

$$n! = n \times (n - 1) \times \dots \times 1.$$

**Demonstração 1.1.** Como temos  $n$  espaços a serem preenchidos, temos  $n$  opções de números para o 1º espaço. Sobram assim  $n - 1$  escolhas diferentes para o 2º espaço, e assim por diante. Para o último espaço, teremos 1 opção.

**Definição 1.2.** Uma **transposição** é uma troca de posição entre dois elementos do conjunto.

Vamos definir  $p$  como sendo uma permutação do conjunto  $I_n$  e  $p_i$  como sendo o  $i$ -ésimo termo dessa permutação. A permutação  $p$  pode ser escrita como uma composição (não única) de transposições.

A demonstração para este resultado pode ser encontrada em [6], Capítulo V.

**Exemplo 1.2.** Podemos considerar a permutação  $(2, 3, 1)$ , do conjunto  $I_3 = \{1, 2, 3\}$ , como a composição das transposições:

- $t_1$ , que troca o elemento 1 com o elemento 3, em  $(1, 2, 3)$ , obtendo assim a terna  $(3, 2, 1)$ ;
- $t_2$ , que troca o elemento 3 com o elemento 2, obtendo assim a permutação desejada.

Além disso, vamos definir  $\sigma(p)$  como a função sinal da permutação, determinada pela paridade do número de transposições necessárias para que  $p$  retorne a forma original. Assim,

$$\sigma(p) = \begin{cases} +1, & \text{se } p \text{ retorna à forma original após um número par de transposições} \\ -1, & \text{se } p \text{ retorna à forma original após um número ímpar de transposições} \end{cases}.$$

**Observação 1.3.** A função  $\sigma(p)$  está bem-definida, ou seja, se  $p$  é escrita como composição de  $n$  transposições e, por outro modo, de  $m$  transposições, então  $n$  e  $m$  são ambos pares ou ambos ímpares.

Desta forma, definimos o determinante de uma matriz quadrada

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

por

$$\det A = \sum_p [\sigma(p) a_{1,p_1} a_{2,p_2} \cdots a_{n,p_n}].$$

### 1.2.1 Matriz Transposta

A matriz transposta de uma matriz  $P$   $m \times n$ , denotada por  $P^T$ , é resultado da troca de linhas por colunas na matriz  $P$ . Assim,

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mn} \end{bmatrix} \Rightarrow P^T = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{m1} \\ p_{12} & p_{22} & \cdots & p_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1n} & p_{2n} & \cdots & p_{mn} \end{bmatrix}.$$

$P^T$  será uma matriz  $n \times m$ .

## 1.3 Sequências Convergentes

Veremos posteriormente que, ao utilizarmos os métodos iterativos para encontrar a solução de um sistema linear, a cada iteração realizada é calculado um termo de uma sequência de listas ordenadas que, cumprindo alguns requisitos, converge para a solução.

A sequência  $\{x^{(k)}\}$ ,  $k \in \mathbb{N}$ , onde  $x^{(t)}$  é o termo calculado na  $t$ -ésima iteração, converge para a solução  $\bar{x}$  se, para toda margem de erro  $\epsilon > 0$ , existe  $K \in \mathbb{N}$  tal que,

$$\forall k > K \Rightarrow |x^{(k)} - \bar{x}| < \epsilon.$$

$$\text{Notação: } \lim_{k \rightarrow \infty} x^{(k)} = \bar{x} \text{ ou } x^{(k)} \rightarrow \bar{x}$$

Em outras palavras, estipulado um erro  $\epsilon$ , podemos encontrar  $K \in \mathbb{N}$  tal que, ao calcularmos a  $K$ -ésima lista ordenada, o erro para a solução exata é menor que  $\epsilon$ .

## 1.4 Conjunto-Solução e Equivalência de Sistemas

O conjunto-solução  $S$  de um sistema de equações lineares consiste das listas ordenadas  $(x_1, x_2, \dots, x_n)$  que satisfazem todas as  $m$  equações presentes no sistema. De acordo com esse conjunto  $S$ , dividimos os sistemas em três tipos:

- **Possível determinado:** o conjunto  $S$  é unitário, ou seja, a lista ordenada que satisfaz o sistema linear é única;
- **Possível indeterminado:** o conjunto  $S$  é composto de um número infinito de listas ordenadas que satisfazem o sistema linear;
- **Impossível:** o conjunto  $S$  é vazio, ou seja, não existem listas ordenadas que satisfaçam o sistema linear.

**Observação 1.4.** *Quando um sistema linear possível tem um número finito de soluções, obrigatoriamente esse número é igual a 1. Caso sejam encontradas 2 ou mais soluções, é certo que, na verdade, podem ser encontradas infinitas soluções para esse sistema.*

A demonstração deste fato pode ser encontrada em [4], Capítulo 1.

Como o objetivo deste trabalho é analisar qual o melhor método de obtenção da solução do sistema, nosso objeto de estudo será o sistema **possível determinado**.

Os sistemas lineares possíveis determinados podem ser facilmente identificados através do cálculo do determinante da matriz  $A$ , dos coeficientes.

**Teorema 1.1.** *Se  $\det A \neq 0$ , então o sistema linear tem solução única, ou seja, é possível determinado.*

A demonstração deste teorema pode ser encontrada em [1], Capítulo 2.

No intuito de otimizar o processo de obtenção da solução de um sistema, pode ser necessário modificar o formato original de suas equações (modificando assim as matrizes  $A$  e  $b$ ) sem que a solução final se altere.

**Definição 1.3.** *Dois sistemas lineares que possuem o mesmo conjunto-solução são ditos equivalentes.*

Vejam algumas formas de modificação do sistema que não influenciam na sua solução (tornando assim o novo sistema equivalente ao anterior):

### 1. Multiplicar uma das equações por uma constante $k \neq 0$ :

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array} \right. \Rightarrow \left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ ka_{i1}x_1 + ka_{i2}x_2 + \dots + ka_{in}x_n = kb_i \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array} \right.$$

**Demonstração 1.2.** ( $\Rightarrow$ ) De fato, se  $(\bar{x}_1; \bar{x}_2; \dots; \bar{x}_n)$  é a solução do sistema, temos que  $a_{i1}\bar{x}_1 + a_{i2}\bar{x}_2 + \dots + a_{in}\bar{x}_n = b_i$ .

Substituindo a solução na nova  $i$ -ésima equação, temos

$$\begin{aligned} ka_{i1}\bar{x}_1 + ka_{i2}\bar{x}_2 + \dots + ka_{in}\bar{x}_n &= \\ k(a_{i1}\bar{x}_1 + a_{i2}\bar{x}_2 + \dots + a_{in}\bar{x}_n) &= kb_i. \end{aligned}$$

Logo, a solução também satisfaz a nova equação  $i$ , determinando a solução do novo sistema.

( $\Leftarrow$ ) Se  $(\bar{x}_1; \bar{x}_2; \dots; \bar{x}_n)$  é solução do novo sistema, então  $ka_{i1}\bar{x}_1 + ka_{i2}\bar{x}_2 + \dots + ka_{in}\bar{x}_n = kb_i$ .

$$\begin{aligned} &\Rightarrow k(a_{i1}\bar{x}_1 + a_{i2}\bar{x}_2 + \dots + a_{in}\bar{x}_n) = kb_i \\ \Rightarrow a_{i1}\bar{x}_1 + a_{i2}\bar{x}_2 + \dots + a_{in}\bar{x}_n &= b_i, \text{ pois } k \neq 0. \end{aligned}$$

Assim, a solução também satisfaz a  $i$ -ésima equação original, e, por consequência, o sistema linear original.

Como ambos os sistemas têm a mesma solução  $(\bar{x}_1; \bar{x}_2; \dots; \bar{x}_n)$ , eles são sistemas equivalentes.

## 2. Permutar linhas:

Considerando  $i < j$ :

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i \\ \vdots \\ a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n = b_j \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array} \right. \Rightarrow \left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n = b_j \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array} \right.$$

Aqui, não há o que demonstrar. A solução satisfaz cada equação do sistema, independente da ordem em que elas forem tratadas.

**Observação 1.5.** O fato de permutar colunas matematicamente modifica o sistema, pois a solução é uma lista ordenada, e assim,

$$(x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_n) \neq (x_1, x_2, \dots, x_j, \dots, x_i, \dots, x_n), \text{ com } i < j.$$

No entanto, na resolução de problemas, não haverá grandes consequências, basta tomar cuidado com a ordem das incógnitas na solução.

## 3. Somar a uma linha uma combinação linear de outras linhas:

Vejamos aqui um caso simplificado, onde somamos a uma linha uma outra linha multiplicada por uma constante. Assim:

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i \\ \vdots \\ a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n = b_j \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array} \right.$$

$$\Rightarrow \left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i \\ \vdots \\ (a_{j1} + ka_{i1})x_1 + (a_{j2} + ka_{i2})x_2 + \dots + (a_{jn} + ka_{in})x_n = (b_j + kb_i) \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array} \right.$$

**Demonstração 1.3.** ( $\Rightarrow$ ) De fato, se  $(\bar{x}_1; \bar{x}_2; \dots; \bar{x}_n)$  é a solução do sistema inicial, temos que  $a_{i1}\bar{x}_1 + a_{i2}\bar{x}_2 + \dots + a_{in}\bar{x}_n = b_i$  e  $a_{j1}\bar{x}_1 + a_{j2}\bar{x}_2 + \dots + a_{jn}\bar{x}_n = b_j$ .

Substituindo a solução na nova  $j$ -ésima equação, temos

$$(a_{j1} + ka_{i1})\bar{x}_1 + (a_{j2} + ka_{i2})\bar{x}_2 + \dots + (a_{jn} + ka_{in})\bar{x}_n = (a_{j1}\bar{x}_1 + a_{j2}\bar{x}_2 + \dots + a_{jn}\bar{x}_n) + k(a_{i1}\bar{x}_1 + a_{i2}\bar{x}_2 + \dots + a_{in}\bar{x}_n) = b_j + kb_i.$$

Logo, a solução também satisfaz a nova equação  $j$ , e assim, resolve-se o novo sistema linear.

( $\Leftarrow$ ) Se  $(\bar{x}_1; \bar{x}_2; \dots; \bar{x}_n)$  é solução do novo sistema, temos que  $(a_{j1} + ka_{i1})\bar{x}_1 + (a_{j2} + ka_{i2})\bar{x}_2 + \dots + (a_{jn} + ka_{in})\bar{x}_n = (b_j + kb_i)$ .

$$\begin{aligned} &\Rightarrow (a_{j1}\bar{x}_1 + a_{j2}\bar{x}_2 + \dots + a_{jn}\bar{x}_n) + k(a_{i1}\bar{x}_1 + a_{i2}\bar{x}_2 + \dots + a_{in}\bar{x}_n) = b_j + kb_i \\ &\Rightarrow (a_{j1}\bar{x}_1 + a_{j2}\bar{x}_2 + \dots + a_{jn}\bar{x}_n) + kb_i = b_j + kb_i, \text{ pois a solução satisfaz a } i\text{-ésima} \\ &\hspace{10em} \text{equação} \\ &\Rightarrow a_{j1}\bar{x}_1 + a_{j2}\bar{x}_2 + \dots + a_{jn}\bar{x}_n = b_j. \end{aligned}$$

Assim, a solução satisfaz a antiga  $j$ -ésima equação, e, conseqüentemente, o sistema linear original.

Assim, temos sistemas lineares equivalentes, pois ambos possuem a mesma solução.

A demonstração é análoga para combinações lineares envolvendo mais de uma linha.

**Observação 1.6.** As técnicas apresentadas também podem ser utilizadas em matrizes, com a mesma eficácia.

Utilizaremos essas técnicas para simplificar os cálculos, nos diferentes métodos que veremos.

Analisaremos, neste trabalho, sistemas lineares com  $n$  equações e  $n$  incógnitas ( $n \times n$ ), da forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases} .$$

O caso  $n = 1$  é trivial, sendo equivalente a uma equação simples. Dessa forma, os resultados serão analisados para  $n \geq 2, n \in \mathbb{N}$ .

Os métodos de resolução são divididos em dois tipos, os métodos **diretos** e **iterativos**. Analisaremos, em um primeiro momento, os tipos de método separadamente. Veremos a contagem de passos de cada método, as limitações, bem como as vantagens de um sobre o outro.

## 2 Métodos Diretos de Resolução

Os métodos diretos fornecem uma solução exata do sistema linear, a menos de arredondamentos, após um número finito de operações, caso tal solução exista.

Analisaremos os seguintes métodos diretos:

1. Substituição;
2. Eliminação Gaussiana;
3. Fatoração LU;
4. Regra de Cramer.

Após demonstrar a resolução de um problema proposto através desses métodos, será calculado o custo computacional necessário para a obtenção da solução em cada um deles. Além disso, será feita uma generalização (para o caso de um sistema linear  $n \times n$ ) desse cálculo. Esta generalização será feita considerando o pior caso possível, no qual o maior número de operações são necessárias, ignorando quaisquer facilidades que possam ser encontradas no sistema.

**Observação 2.1.** *Para determinar o custo computacional do algoritmo, serão contabilizadas as multiplicações (divisões incluídas) e adições (subtrações incluídas) realizadas durante o processo. As substituições de uma incógnita pelo seu valor específico não influenciam nesta conta, visto que o computador, ao encontrar esse valor, associa ambos automaticamente.*

Ao final, analisaremos as limitações de cada algoritmo, bem como as vantagens de um sobre o outro.

Apresentamos, para ilustrar a resolução, uma aplicação na área da Geometria:

*Sejam  $\alpha_1$ ,  $\alpha_2$  e  $\alpha_3$  três planos em  $\mathbb{R}^3$ , tais que*  
 $\alpha_1 : x_1 + 2x_2 - 3x_3 = 4$ ,  $\alpha_2 : 2x_1 + 3x_2 + 4x_3 = 5$  e  $\alpha_3 : 4x_1 + 7x_2 - x_3 = 13$ .  
*Queremos determinar a interseção entre esses planos.*

**Observação 2.2.** *Um ponto  $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$  pertence ao plano  $\alpha_1$  quando a terna  $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$  é solução da equação  $x_1 + 2x_2 - 3x_3 = 4$  referente ao plano. O processo é análogo para os demais planos.*

Assim, o problema pode ser reduzido ao seguinte sistema  $3 \times 3$  de equações lineares:

$$\begin{cases} x_1 + 2x_2 - 3x_3 = 4 & (I) \\ 2x_1 + 3x_2 + 4x_3 = 5 & (II) \\ 4x_1 + 7x_2 - x_3 = 13 & (III) \end{cases} .$$

Ao calcularmos o determinante da matriz dos coeficientes

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 3 & 4 \\ 4 & 7 & -1 \end{bmatrix},$$

através da definição, vemos que  $\det A = -1 \neq 0$ , logo, o sistema é possível determinado.

De fato,

$$\begin{aligned} \det A &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} + a_{13}a_{21}a_{32} \\ &= 1 \times 3 \times (-1) - 1 \times 4 \times 7 - 2 \times 2 \times (-1) \\ &\quad + 2 \times 4 \times 4 - (-3) \times 3 \times 4 + (-3) \times 2 \times 7 \\ &= -3 - 28 + 4 + 32 + 36 - 42 \\ &= -73 + 72 \\ &= -1. \end{aligned}$$

A partir de então, encontraremos o único ponto  $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$  que satisfaz as três equações simultaneamente, ou seja, está na interseção entre os três planos. Para isso, utilizaremos os quatro métodos diretos propostos.

## 2.1 Substituição

O método da Substituição (Subst.) consiste em reescrever as equações do sistema, para assim isolar certa incógnita e substituí-la em outra equação.

Primeiramente, reescrevemos a equação (III), escolha sem perda de generalidade, de forma a isolar a variável  $x_3$ :

$$\begin{aligned} (III) : \quad 4x_1 + 7x_2 - x_3 &= 13 \\ \Rightarrow -x_3 &= 13 - 4x_1 - 7x_2 \\ \Rightarrow x_3 &= 4x_1 + 7x_2 - 13. \end{aligned} \tag{1}$$

*Na segunda linha, foram realizadas 2 adições, enquanto na terceira foram realizadas 3 multiplicações, totalizando 5 operações nesta etapa.*

Agora, substituímos (1) na equação (II), eliminando assim a variável  $x_3$  na equação:

$$\begin{aligned} (II) : \quad 2x_1 + 3x_2 + 4x_3 &= 5 \\ \Rightarrow 2x_1 + 3x_2 + 4 \times (4x_1 + 7x_2 - 13) &= 5 \\ \Rightarrow 2x_1 + 3x_2 + 16x_1 + 28x_2 - 52 &= 5 \\ \Rightarrow 18x_1 + 31x_2 &= 57. \end{aligned} \tag{2}$$

*Foram realizadas 3 multiplicações na segunda linha e 3 adições na terceira linha, para um total de 6 operações neste passo.*

Da mesma forma, substituímos (1) na equação (I):

$$\begin{aligned} (I) : \quad x_1 + 2x_2 - 3x_3 &= 4 \\ \Rightarrow x_1 + 2x_2 - 3 \times (4x_1 + 7x_2 - 13) &= 4 \\ \Rightarrow x_1 + 2x_2 - 12x_1 - 21x_2 + 39 &= 4 \\ \Rightarrow -11x_1 - 19x_2 &= -35. \end{aligned} \tag{3}$$

Como este passo é análogo ao anterior, foram realizadas **6** operações para sua conclusão.

Em seguida, reescreveremos (2), de forma a isolar a variável  $x_2$ , tal como no primeiro passo:

$$\begin{aligned}(2) : \quad & 18x_1 + 31x_2 = 57 \\ \Rightarrow \quad & 31x_2 = 57 - 18x_1 \\ \Rightarrow \quad & x_2 = \frac{57 - 18x_1}{31}.\end{aligned}\tag{4}$$

Agora, realizamos 1 adição na segunda linha e 2 multiplicações na terceira linha: **3** operações.

Substituiremos então (4) na equação (3), de forma a termos somente a incógnita  $x_1$  na equação (I) original:

$$\begin{aligned}(3) : \quad & -11x_1 - 19x_2 = -35 \\ \Rightarrow \quad & -11x_1 - 19 \times \frac{57 - 18x_1}{31} = -35 \\ \Rightarrow \quad & -11x_1 - \frac{1083}{31} + \frac{342x_1}{31} = -35 \\ \Rightarrow \quad & \frac{x_1}{31} = -\frac{2}{31}.\end{aligned}\tag{5}$$

Nesta etapa, realizamos 2 multiplicações da segunda para a terceira linha e 2 duas adições da terceira para a quarta linha, fechando **4** operações.

A partir de agora, podemos resolver as equações para obter os valores das incógnitas diretamente. Primeiramente, vamos resolver (5) para obter o valor de  $x_1$ :

$$\begin{aligned}(5) : \quad & \frac{x_1}{31} = -\frac{2}{31} \\ \Rightarrow \quad & x_1 = -2.\end{aligned}$$

Aqui, foi realizada 1 multiplicação, logo, apenas **1** operação.

Substituindo  $x_1$  em (4), temos:

$$\begin{aligned}(4) : \quad & x_2 = \frac{57 - 18x_1}{31} \\ \Rightarrow \quad & x_2 = \frac{57 - 18 \times (-2)}{31} \\ \Rightarrow \quad & x_2 = \frac{57 + 36}{31} \\ \Rightarrow \quad & x_2 = \frac{93}{31} \\ \Rightarrow \quad & x_2 = 3.\end{aligned}$$

Para encontrar  $x_2$  a partir de  $x_1$ , fizemos 1 multiplicação na terceira linha e 1 adição na quarta linha. Não convém contar a passagem  $\frac{93}{31}$  para 3, pois o computador já havia efetuado estas multiplicações ao obter a equação (4). Logo, são **2** operações.

Por fim, substituindo  $x_1$  e  $x_2$  em (1), obteremos:

$$\begin{aligned}(1) : \quad & x_3 = 4x_1 + 7x_2 - 13 \\ \Rightarrow & x_3 = 4 \times (-2) + 7 \times 3 - 13 \\ \Rightarrow & x_3 = -8 + 21 - 13 \\ \Rightarrow & x_3 = 0.\end{aligned}$$

*Aqui, foram 2 multiplicações na terceira linha e 2 adições na terceira linha, finalizando com 4 operações.*

Assim, obtivemos como solução para o sistema linear o conjunto unitário cujo elemento é a terna ordenada  $(-2, 3, 0)$ . Para isso, tiveram de ser realizadas  $5 + 6 + 6 + 3 + 4 + 1 + 2 + 4 = \mathbf{31}$  operações.

Podemos generalizar essa contagem para um sistema linear  $n \times n$ :

- Para isolar  $x_n$  na  $n$ -ésima equação, é necessário fazer  $n - 1$  adições (deixando assim  $a_{nn}x_n$  isolado em um dos lados da equação) e  $n$  multiplicações (fazendo assim cada termo do lado oposto da equação ser dividido por  $a_{nn}$ );
- Para substituir  $x_n$  na  $k$ -ésima equação,  $k \in \{1, \dots, n - 1\}$  são necessárias  $n$  multiplicações (cada termo da fórmula de  $x_n$  multiplicado pelo coeficiente  $a_{kn}$ ), e em seguida  $n$  adições, para juntar os monômios semelhantes. Este processo se repete para cada  $k \in \{1, \dots, n - 1\}$ , logo,  $n - 1$  vezes. Assim, são realizadas  $n \times (n - 1)$  multiplicações e  $n \times (n - 1)$  adições nesta etapa;
- Para isolar  $x_{n-1}$  na  $(n - 1)$ -ésima equação, é necessário fazer  $n - 2$  adições (deixando assim a multiplicação de  $x_{n-1}$  com seu coeficiente isolada em um dos lados da equação) e  $n - 1$  multiplicações (fazendo assim cada termo do lado oposto da equação ser dividido pelo coeficiente);
- Para substituir  $x_{n-1}$  na  $k$ -ésima equação,  $k \in \{1, \dots, n - 2\}$  são necessárias  $n - 1$  multiplicações (cada termo da fórmula de  $x_{n-1}$  multiplicado pelo seu respectivo coeficiente, e em seguida  $n - 1$  adições, para juntar os monômios semelhantes. Este processo se repete para cada  $k \in \{1, \dots, n - 2\}$ , logo,  $n - 2$  vezes. Assim, são realizadas  $(n - 1) \times (n - 2)$  multiplicações e  $(n - 1) \times (n - 2)$  adições nesta etapa;

⋮

- Para isolar  $x_2$  na segunda equação, é necessário fazer 1 adição (deixando assim  $x_2$ , juntamente com seu coeficiente, isolados em um dos lados da equação) e 1 multiplicação (fazendo assim cada termo do lado oposto da equação ser dividido pelo coeficiente);
- Para substituir  $x_2$  na primeira equação, são necessárias 2 multiplicações (cada termo da fórmula de  $x_2$  multiplicado pelo coeficiente, e em seguida 2 adições, para juntar os monômios semelhantes. Assim, são realizadas  $2 \times 1$  multiplicações e  $2 \times 1$  adições nesta etapa;

- Isolando  $x_1$  na primeira equação, basta 1 multiplicação e, desta forma, encontramos o valor de  $x_1$  por meio de 1 multiplicação e 0 adição;
- Para encontrar o valor de  $x_2$ , substitui-se  $x_1$  pelo seu valor. Em seguida, é realizada 1 multiplicação (de  $x_1$  pelo seu coeficiente) e 1 soma (com o novo termo independente da equação);

⋮

- Para obter  $x_n$ , substituímos as variáveis  $x_k$ ,  $k \in \{1, \dots, n-1\}$  pelos seus respectivos valores, e realizamos  $n-1$  multiplicações (cada variável com seu coeficiente) e  $n-1$  adições.

Somando as contagens das etapas necessárias para isolar as variáveis, temos:

$$\begin{aligned} \#(+\text{Isolar Subst.}) &= (n-1) + \dots + 1 + 0 \\ &= \frac{(n-1) \times n}{2} \\ &= \frac{n^2 - n}{2}; \end{aligned}$$

$$\begin{aligned} \#(\times\text{Isolar Subst.}) &= n + \dots + 1 \\ &= \frac{(n+1) \times n}{2} \\ &= \frac{n^2 + n}{2}. \end{aligned}$$

Assim, o total de operações necessárias para isolar todas as  $n$  variáveis do sistema é dado pela fórmula

$$\begin{aligned} \#(\text{Isolar Subst.}) &= \frac{n^2 - n}{2} + \frac{n^2 + n}{2} \\ &= \frac{n^2 - n + n^2 + n}{2} \\ &= \frac{2n^2}{2} \\ &= n^2. \end{aligned} \tag{6}$$

Para as etapas de substituição, temos:

$$\begin{aligned}
\#(+\text{Substituir Subst.}) &= n \times (n - 1) + (n - 1) \times (n - 2) + \dots + 2 \times 1 \\
&= \sum_{i=1}^{n-1} [i \times (i + 1)] \\
&= \sum_{i=1}^{n-1} [i^2 + i] \\
&= \sum_{i=1}^{n-1} i^2 + \sum_{i=1}^{n-1} i \\
&= \frac{n \times (n + 1) \times (2n + 1)}{6} - n^2 + \frac{n \times (n - 1)}{2} \\
&= \frac{2n^3 + 3n^2 + n}{6} - n^2 + \frac{n^2 - n}{2} \\
&= \frac{2n^3 + 3n^2 + n - 6n^2 + 3n^2 - 3n}{6} \\
&= \frac{2n^3 - 2n}{6} \\
&= \frac{n^3 - n}{3};
\end{aligned}$$

$$\begin{aligned}
\#(\times\text{Substituir Subst.}) &= \#(+\text{Substituir Subst.}) \\
&= \frac{n^3 - n}{3}.
\end{aligned}$$

Assim, o total de operações necessárias no processo de substituição de uma variável em outra equação é dado por

$$\begin{aligned}
\#(\text{Substituir Subst.}) &= \frac{n^3 - n}{3} + \frac{n^3 - n}{3} \\
&= \frac{2(n^3 - n)}{3} \\
&= \frac{2n^3 - 2n}{3}. \tag{7}
\end{aligned}$$

Por fim, no processo de obtenção do valor das variáveis, temos:

$$\begin{aligned}
\#(+\text{Obter Subst.}) &= 1 + 2 + \dots + (n - 1) \\
&= \frac{(n - 1) \times n}{2} \\
&= \frac{n^2 - n}{2};
\end{aligned}$$

$$\begin{aligned}
\#(\times \text{Obter Subst.}) &= 1 + 2 + \dots + (n - 1) \\
&= \frac{(n - 1) \times n}{2} \\
&= \frac{n^2 - n}{2}.
\end{aligned}$$

Assim, o total de operações necessárias para obter todas as demais variáveis do sistema é dado pela fórmula

$$\begin{aligned}
\#(\text{Obter Subst.}) &= \frac{n^2 - n}{2} + \frac{n^2 - n}{2} \\
&= \frac{2n^2 - 2n}{2} \\
&= n^2 - n.
\end{aligned} \tag{8}$$

Juntando as equações (6), (7) e (8), encontramos a fórmula para o cálculo do total de operações necessárias na resolução de um sistema linear  $n \times n$  através do método da Substituição:

$$\begin{aligned}
\#(\text{Subst.}) &= n^2 + \frac{2n^3 - 2n}{3} + n^2 - n \\
&= \frac{3n^2 + 2n^3 - 2n + 3n^2 - 3n}{3} \\
&= \frac{2n^3 + 6n^2 - 5n}{3}.
\end{aligned}$$

## 2.2 Eliminação Gaussiana

O método da Eliminação Gaussiana (Elim.) transforma o sistema linear a ser analisado em um sistema linear equivalente. No entanto, para que este método seja vantajoso, o novo sistema linear deve ser mais fácil de resolver do que o anterior.

Para que o processo seja feito, podemos trabalhar com a *matriz ampliada* correspondente ao sistema, da forma

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & b_n \end{bmatrix}.$$

A matriz ampliada é uma junção das matrizes  $A$  (coeficientes das variáveis) e  $b$  (termos independentes) já vistas anteriormente.

O processo de **escalonamento** da matriz consiste em anularmos os termos  $a_{jk}$ ,  $j \in \{k + 1, \dots, n\}$ ,  $k \in \{1, \dots, n - 1\}$ , coluna por coluna. Para isso, basta substituir a linha  $L_j$  (representação que será utilizada para representar a  $j$ -ésima linha da matriz) pela combinação linear  $L_j - \frac{a_{jk}}{a_{kk}}L_k$ . Sabemos que tais mudanças não alteram o determinante da matriz nem a solução do sistema proposto.

**Observação 2.3.** Ao final da anulação dos termos de uma determinada coluna, os termos das colunas posteriores terão sofrido modificações, e não necessariamente terão o mesmo valor que tinham na matriz original. Para representarmos tais modificações, será utilizada a notação  $a_{ij}^{(t)}$  para representar o termo  $a_{ij}$  após a eliminação da  $t$ -ésima coluna.

**Observação 2.4.** Durante o processo de anulação dos termos desejados, percebe-se que o coeficiente  $a_{kk}$ ,  $k \in \{1, \dots, n-1\}$ , pertencente à diagonal principal da matriz original  $A$ , corresponde ao denominador da fração à qual multiplicamos a linha  $L_k$ . Assim, caso exista algum termo nulo na diagonal principal, não será possível escrever a combinação linear correspondente.

Para isso, podemos trocar a posição de algumas linhas ou algumas colunas, a fim de retirarmos o termo nulo da diagonal principal. Vimos, anteriormente, que tal troca não altera o resultado final, pois encontraríamos assim um sistema linear equivalente ao original.

Além disso, existe a possibilidade de um coeficiente localizado na diagonal principal, que originalmente não era igual a zero, tornar-se nulo, ou vice-versa, após a anulação de alguma coluna. O leitor deve atentar-se de que tais trocas de linhas ou colunas também podem ser feitas no decorrer do processo de escalonamento da matriz ampliada.

**Observação 2.5.** A troca de linhas ou colunas pode ser utilizada também para simplificar as contas durante o processo de eliminação. Essa técnica é chamada de **pivoteamento**, e é dividida em dois tipos: o pivoteamento parcial, no qual apenas linhas são trocadas; e o pivoteamento total, onde pode-se trocar linhas e colunas. Assim, o número de operações necessárias para a conclusão do processo de escalonamento pode diminuir. No entanto, a forma como ela é utilizada depende da matriz apresentada. Para generalizar as contas, e assim fixar uma equação para o cálculo das operações, essa técnica **não** será utilizada na resolução de sistemas neste trabalho.

No final do escalonamento, a parte  $A$  da matriz ampliada torna-se uma matriz triangular superior, da seguinte forma:

$$\begin{bmatrix} a_{11}^{(n-1)} & a_{12}^{(n-1)} & a_{13}^{(n-1)} & \cdots & a_{1n}^{(n-1)} & b_1^{(n-1)} \\ 0 & a_{22}^{(n-1)} & a_{23}^{(n-1)} & \cdots & a_{2n}^{(n-1)} & b_2^{(n-1)} \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{n-1,n-1}^{(n-1)} & a_{n,n-1}^{(n-1)} & b_{n-1}^{(n-1)} \\ 0 & \cdots & 0 & 0 & a_{nn}^{(n-1)} & b_n^{(n-1)} \end{bmatrix}.$$

Veremos que uma matriz nestes moldes facilita a obtenção dos valores de cada incógnita.

Agora, resolveremos novamente o problema proposto neste capítulo. A matriz ampliada do sistema é dada por

$$\begin{bmatrix} 1 & 2 & -3 & 4 \\ 2 & 3 & 4 & 5 \\ 4 & 7 & -1 & 13 \end{bmatrix}.$$

Primeiramente, queremos eliminar os termos  $a_{21}$  e  $a_{31}$ , da primeira coluna, tornando-os nulos. Para isso, substituiremos  $L_2$  por uma combinação entre as linhas  $L_1$  e  $L_2$ , mais

precisamente  $L_2 - \frac{a_{21}}{a_{11}}L_1 = L_2 - 2L_1$ , e  $L_3$  por  $L_3 - \frac{a_{31}}{a_{11}}L_1 = L_3 - 4L_1$ . Assim,

$$\begin{bmatrix} 1 & 2 & -3 & 4 \\ 2 & 3 & 4 & 5 \\ 4 & 7 & -1 & 13 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & -3 & 4 \\ 0 & -1 & 10 & -3 \\ 0 & -1 & 11 & -3 \end{bmatrix}.$$

Para a anulação de  $a_{21}$ , foi utilizada 1 multiplicação para o cálculo do coeficiente ao qual multiplicamos  $L_1$  e 3 multiplicações, do coeficiente com  $a_{22}$ ,  $a_{23}$  e  $b_2$ . Em seguida, são feitas 3 adições, novamente uma para cada termo citado anteriormente, totalizando 7 operações. Como o processo é análogo para a eliminação de  $a_{31}$ , são necessárias 7 operações para isso, totalizando **14** operações nesta etapa.

**Observação 2.6.** Para o computador não é necessária a multiplicação e a divisão que envolvem o termo a ser anulado, dado que o algoritmo da Eliminação Gaussiana dá a ordem direta para que tal termo seja substituído por 0 (como mencionado anteriormente, substituições de termos por valores específicos não tem custo computacional). Da mesma forma, o aluno que resolve o sistema pelo método da eliminação Gaussiana não tem a necessidade de fazer estas contas, bastando apenas escrever 0 onde é devido.

Agora que a primeira coluna foi finalizada, faremos o mesmo com a segunda coluna. Para isso, definiremos a nova linha 2 como  $L_2^{(1)}$  e a nova linha 3 como  $L_3^{(1)}$ . Substituiremos  $L_3^{(1)}$  por uma combinação entre  $L_2^{(1)}$  e  $L_3^{(1)}$ , da forma  $L_3^{(1)} - \frac{a_{32}^{(1)}}{a_{22}^{(1)}}L_2^{(1)} = L_3^{(1)} - L_2^{(1)}$ . Assim,

$$\begin{bmatrix} 1 & 2 & -3 & 4 \\ 0 & -1 & 10 & -3 \\ 0 & -1 & 11 & -3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & -3 & 4 \\ 0 & -1 & 10 & -3 \\ 0 & 0 & -1 & 0 \end{bmatrix}.$$

Para anular  $a_{32}^{(1)}$ , utilizamos 1 multiplicação para o coeficiente e multiplicações do coeficiente com os termos  $a_{33}^{(1)}$  e  $b_3^{(1)}$ , separadamente (logo, 2). Além disso, 2 adições, para cada um dos termos citados. Assim, nessa etapa foram realizadas **5** operações.

Finalizado o processo, transformamos a parte relativa a matriz  $A$  da matriz ampliada em uma matriz triangular superior, logo escalonada. Para isso, foram necessárias  $14 + 5 = 19$  operações. Agora, podemos transformá-la novamente em um sistema diferente, mas equivalente ao sistema original:

$$\begin{bmatrix} 1 & 2 & -3 & 4 \\ 0 & -1 & 10 & -3 \\ 0 & 0 & -1 & 0 \end{bmatrix} \Rightarrow \begin{cases} x_1 + 2x_2 - 3x_3 = 4 & (I') \\ -x_2 + 10x_3 = -3 & (II') \\ -x_3 = 0 & (III') \end{cases}.$$

A partir de agora, assim como no método da Substituição, podemos resolver as equações para obter os valores das incógnitas diretamente. Primeiramente, vamos resolver  $(III')$  para obter o valor de  $x_3$ :

$$\begin{aligned} (III') : \quad & -x_3 = 0 \\ \Rightarrow \quad & x_3 = 0. \end{aligned}$$

Para obter o valor de  $x_3$ , foi necessária 1 multiplicação e 0 adição, logo, **1** operação.

Substituindo  $x_3$  em  $(II')$ , temos:

$$\begin{aligned}
 (II') : \quad & -x_2 + 10x_3 = -3 \\
 \Rightarrow \quad & -x_2 + 10 \times 0 = -3 \\
 \Rightarrow \quad & -x_2 + 0 = -3 \\
 \Rightarrow \quad & -x_2 = -3 \\
 \Rightarrow \quad & x_2 = 3.
 \end{aligned}$$

Nesta etapa, na terceira linha foi realizada 1 multiplicação, 1 adição na quarta linha, finalizando com 1 multiplicação na quinta linha. Total: **3** operações.

Por fim, substituindo  $x_2$  e  $x_3$  em  $(I')$ , obteremos:

$$\begin{aligned}
 (I') : \quad & x_1 + 2x_2 - 3x_3 = 4 \\
 \Rightarrow \quad & x_1 + 2 \times 3 - 3 \times 0 = 4 \\
 \Rightarrow \quad & x_1 + 6 - 0 = 4 \\
 \Rightarrow \quad & x_1 = -2.
 \end{aligned}$$

Para encerrar o algoritmo, foram realizadas 2 multiplicações na terceira linha, 1 adição na quarta, 1 adição na quinta, e 1 multiplicação implícita acima, mas necessária ao computador, de dividir o termo independente pelo coeficiente de  $x_1$  ( $\frac{2}{1}$ ), totalizando então **5** operações.

Assim, para a etapa de resolução do sistema, foram necessárias  $1 + 3 + 5 = \mathbf{9}$  operações.

Encontramos a mesma solução, única, com a terna ordenada  $(-2, 3, 0)$ . Porém, desta vez nos utilizamos de apenas **28 operações** (19 na fase de escalonamento e 9 na fase de resolução).

Generalizando para um sistema linear  $n \times n$ , obtemos:

- Para eliminar os termos  $a_{k1}$ ,  $k \in \{2, \dots, n\}$ , temos que calcular o coeficiente  $\frac{a_{k1}}{a_{11}}$ , multiplicá-lo com os termos  $a_{12}, \dots, a_{1n}$  e  $b_1$ , para em seguida somar cada um com o seu termo respectivo da  $k$ -ésima linha. Totalizam-se assim  $n + 1$  multiplicações e  $n$  adições. Como o processo deve ser repetido para todas as linhas, temos que, nesta etapa, são realizadas  $(n + 1) \times (n - 1)$  multiplicações e  $n \times (n - 1)$  adições;
- Para os termos  $a_{k2}^{(1)}$ ,  $k \in \{3, \dots, n\}$ , o processo é análogo. Assim, realizamos  $n$  multiplicações e  $n - 1$  adições. Repetindo o processo para todas as  $n - 2$  linhas, temos  $(n) \times (n - 2)$  multiplicações e  $(n - 1) \times (n - 2)$  adições;

⋮

- Para o termo  $a_{n,n-1}^{(n-2)}$ , único termo da  $(n - 1)$ -ésima coluna a ser anulado, e último termo do processo, são necessárias 3 multiplicações  $(\frac{a_{n,n-1}^{(n-2)}}{a_{n-1,n-1}^{(n-2)}}$ , e suas multiplicações

com os termos  $a_{n-1,n}^{(n-2)}$  e  $b_{n-1}^{(n-2)}$  e 2 adições (cada um dos dois termos citados anteriormente, somados com os seus termos respectivos da  $n$ -ésima linha). Assim, temos  $3 \times 1$  multiplicações e  $2 \times 1$  adições;

- Para encontrar o valor de  $x_n$ , basta realizar 1 multiplicação (divisão de  $b_n^{(n-1)}$  por  $a_{nn}^{(n-1)}$ );
- Para obter  $x_{n-1}$ , substituímos o valor de  $x_n$  na  $(n-1)$ -ésima equação, multiplicamos com o seu coeficiente, fazemos 1 soma com o termo independente da equação, e finalizamos com 1 multiplicação, dividindo o novo termo independente pelo coeficiente de  $x_{n-1}$ . Assim, realizamos 2 multiplicações e 1 adição;

⋮

- Para obter  $x_1$ , substituímos o valor das demais incógnitas na primeira equação. Multiplicamos então cada uma pelo seu respectivo coeficiente ( $n-1$  multiplicações). Em seguida, realizamos  $n-1$  adições para encontrar o novo termo independente. Para terminar, fazemos 1 nova multiplicação, dividindo esse termo pelo coeficiente de  $x_1$ . No total, são realizadas  $n$  multiplicações e  $n-1$  adições nesta etapa.

Somando as contagens das etapas necessárias para o escalonamento da matriz, temos:

$$\begin{aligned}
 \#(+\text{Escalar Elim.}) &= n \times (n-1) + (n-1) \times (n-2) + \dots + 2 \times 1 \\
 &= \sum_{i=1}^{n-1} [i \times (i+1)] \\
 &= \sum_{i=1}^{n-1} [i^2 + i] \\
 &= \sum_{i=1}^{n-1} i^2 + \sum_{i=1}^{n-1} i \\
 &= \frac{n \times (n+1) \times (2n+1)}{6} - n^2 + \frac{n \times (n-1)}{2} \\
 &= \frac{2n^3 + 3n^2 + n}{6} - n^2 + \frac{n^2 - n}{2} \\
 &= \frac{2n^3 + 3n^2 + n - 6n^2 + 3n^2 - 3n}{6} \\
 &= \frac{2n^3 - 2n}{6} \\
 &= \frac{n^3 - n}{3};
 \end{aligned}$$

$$\begin{aligned}
\#(\times \text{Escalonar Elim.}) &= (n+1) \times (n-1) + n \times (n-2) + \dots + 3 \times 1 \\
&= \sum_{i=1}^{n-1} [i \times (i+2)] \\
&= \sum_{i=1}^{n-1} [i^2 + 2i] \\
&= \sum_{i=1}^{n-1} i^2 + 2 \sum_{i=1}^{n-1} i \\
&= \frac{n \times (n+1) \times (2n+1)}{6} - n^2 + 2 \frac{n \times (n-1)}{2} \\
&= \frac{2n^3 + 3n^2 + n}{6} - n^2 + n^2 - n \\
&= \frac{2n^3 + 3n^2 + n - 6n}{6} \\
&= \frac{2n^3 + 3n^2 - 5n}{6}.
\end{aligned}$$

Assim, o total de operações necessárias no processo de escalonamento é dado por

$$\begin{aligned}
\#(\text{Escalonar Elim.}) &= \frac{n^3 - n}{3} + \frac{2n^3 + 3n^2 - 5n}{6} \\
&= \frac{2n^3 - 2n + 2n^3 + 3n^2 - 5n}{6} \\
&= \frac{4n^3 + 3n^2 - 7n}{6}.
\end{aligned} \tag{9}$$

No processo de obtenção do valor das variáveis, temos:

$$\begin{aligned}
\#(+\text{Obter Elim.}) &= 0 + 1 + 2 + \dots + (n-1) \\
&= \frac{(n-1) \times n}{2} \\
&= \frac{n^2 - n}{2};
\end{aligned}$$

$$\begin{aligned}
\#(\times \text{Obter Elim.}) &= 1 + 2 + \dots + n \\
&= \frac{(n+1) \times n}{2} \\
&= \frac{n^2 + n}{2}.
\end{aligned}$$

Assim, o total de operações necessárias para obter as variáveis do sistema é dado pela

fórmula

$$\begin{aligned}
 \#(\text{Obter Elim.}) &= \frac{n^2 - n}{2} + \frac{n^2 + n}{2} \\
 &= \frac{2n^2}{2} \\
 &= n^2.
 \end{aligned} \tag{10}$$

Juntando as equações (9) e (10), encontramos a fórmula para o cálculo do total de operações necessárias na resolução de um sistema linear  $n \times n$  através da Eliminação Gaussiana:

$$\begin{aligned}
 \#(\text{Elim.}) &= \frac{4n^3 + 3n^2 - 7n}{6} + n^2 \\
 &= \frac{4n^3 + 3n^2 - 7n + 6n^2}{6} \\
 &= \frac{4n^3 + 9n^2 - 7n}{6}.
 \end{aligned}$$

### 2.3 Fatoração LU

A Fatoração LU (FatLU) consiste em transformarmos a matriz  $A$  no produto de duas matrizes  $L$  e  $U$ , onde  $L$  é uma matriz triangular inferior (do inglês *lower*), com diagonal unitária, e  $U$  é triangular superior (do inglês *upper*), da seguinte forma:

$$L = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ l_{n-1,1} & \cdots & l_{n-1,n-2} & 1 & 0 \\ l_{n1} & l_{n2} & \cdots & l_{n,n-1} & 1 \end{bmatrix} \quad \text{e} \quad U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & u_{n-1,n-1} & u_{n-1,n} \\ 0 & 0 & \cdots & 0 & u_{nn} \end{bmatrix}.$$

Assim, temos que

$$A \times x = b \Rightarrow LU \times x = b.$$

Considerando  $U \times x = y$ , temos que  $L \times y = b$ . Como as matrizes  $L$  e  $U$  são escalonadas, é simples resolver ambas as equações, resolvendo primeiramente o sistema  $L \times y = b$ , para obter a matriz vertical  $y$ , e em seguida  $U \times x = y$ , para finalmente resolver o sistema linear proposto.

Resolveremos agora o problema proposto no capítulo através da Fatoração LU. Porém, antes veremos uma maneira fácil de obter os coeficientes das matrizes  $L$  e  $U$ . Reescrevendo  $L \times U = A$ , temos:

$$\begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \times \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Aplicando a regra da multiplicação de matrizes, concluímos que:

- $u_{11} = a_{11}$ ;
- $u_{12} = a_{12}$ ;
- $u_{13} = a_{13}$ ;
- $l_{21} \times u_{11} = a_{21} \Rightarrow l_{21} \times a_{11} = a_{21} \Rightarrow l_{21} = \frac{a_{21}}{a_{11}}$ ;
- $l_{21} \times u_{12} + u_{22} = a_{22} \Rightarrow u_{22} = a_{22} - \frac{a_{21}}{a_{11}}a_{12} = a_{22}^{(1)}$ ;
- $l_{21} \times u_{13} + u_{23} = a_{23} \Rightarrow u_{23} = a_{23} - \frac{a_{21}}{a_{11}}a_{13} = a_{23}^{(1)}$ ;
- $l_{31} \times u_{11} = a_{31} \Rightarrow l_{31} \times a_{11} = a_{31} \Rightarrow l_{31} = \frac{a_{31}}{a_{11}}$ ;
- $l_{31} \times u_{12} + l_{32} \times u_{22} = a_{32} \Rightarrow \frac{a_{31}}{a_{11}}a_{12} + l_{32}(a_{22} - \frac{a_{21}}{a_{11}}a_{12}) = a_{32}$   
 $\Rightarrow l_{32} = \frac{a_{32} - \frac{a_{31}}{a_{11}}a_{12}}{a_{22} - \frac{a_{21}}{a_{11}}a_{12}} = \frac{a_{32}^{(1)}}{a_{22}^{(1)}}$ ;
- $l_{31} \times u_{13} + l_{32} \times u_{23} + u_{33} = a_{33} \Rightarrow \frac{a_{31}}{a_{11}}a_{13} + \frac{a_{32}^{(1)}}{a_{11}}a_{23}^{(1)} + u_{33} = a_{33}$   
 $\Rightarrow \frac{a_{32}^{(1)}}{a_{11}}a_{23}^{(1)} + u_{33} = a_{33} - \frac{a_{31}}{a_{11}}a_{13} \Rightarrow \frac{a_{32}^{(1)}}{a_{11}}a_{23}^{(1)} + u_{33} = a_{33}^{(1)}$   
 $\Rightarrow u_{33} = a_{33}^{(1)} - \frac{a_{32}^{(1)}}{a_{22}^{(1)}}a_{23}^{(1)} = a_{33}^{(2)}$ .

Vemos, então, que os coeficientes necessários para a composição das matrizes  $L$  e  $U$  são obtidos através do escalonamento da matriz  $A$ : a matriz  $U$  é a própria matriz  $A$  escalonada, enquanto a matriz  $L$  é composta pelas frações utilizadas nas combinações entre as linhas, calculadas também durante a eliminação.

A principal diferença entre a Fatoração LU e a Eliminação Gaussiana é a exclusão da matriz coluna  $b$  do escalonamento, neste caso. Assim, a cada combinação de linhas feita para a anulação de um termo, poupamos uma multiplicação e uma adição. Em contrapartida, temos que resolver dois sistemas triangulares após o escalonamento. Vejamos posteriormente como estas modificações afetarão nossas contas.

Utilizando-se do resultado anterior, obtido na Eliminação Gaussiana, as matrizes  $L$  e  $U$  são dadas por:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 1 & 1 \end{bmatrix} \quad \text{e} \quad U = \begin{bmatrix} 1 & 2 & -3 \\ 0 & -1 & 10 \\ 0 & 0 & 1 \end{bmatrix}.$$

*Para anularmos o termo  $a_{21}$ , foi necessário calcular a fração (agora conhecida como  $l_{21}$ ) a qual multiplicamos  $L_1$  para combinar com  $L_2$ , além de 2 multiplicações e 2 adições, para obtermos  $u_{22} = a_{22}^{(1)}$  e  $u_{23} = a_{23}^{(1)}$ .*

*Analogamente, para  $a_{31}$ , calculamos  $l_{31}$ ,  $u_{32} = a_{32}^{(1)}$  e  $u_{33} = a_{33}^{(1)}$ , utilizando-se de 3 multiplicações e 2 divisões.*

*Para  $a_{32}^{(1)}$ , calculamos  $l_{32}$  e  $u_{33} = a_{33}^{(2)}$ , através de 2 multiplicações e 1 adição.*

Assim, para o processo de obtenção das matrizes  $L$  e  $U$ , foram necessárias 8 multiplicações e 5 adições, totalizando **13** operações.

Temos então a seguinte situação:

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & -3 \\ 0 & -1 & 10 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 13 \end{bmatrix}.$$

Fazendo

$$\begin{bmatrix} 1 & 2 & -3 \\ 0 & -1 & 10 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix},$$

teremos

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 13 \end{bmatrix}.$$

Escrevendo a equação matricial acima na forma de um sistema de equações, temos:

$$\begin{cases} y_1 = 4 & (I') \\ 2y_1 + y_2 = 5 & (II') \\ 4y_1 + y_2 + y_3 = 13 & (III') \end{cases}.$$

Através da equação  $(I')$ , obtivemos  $y_1$ . Substituímos então esse valor na equação  $(II')$ , temos:

$$\begin{aligned} (II') : \quad & 2y_1 + y_2 = 5 \\ \Rightarrow & 2 \times 4 + y_2 = 5 \\ \Rightarrow & 8 + y_2 = 5 \\ \Rightarrow & y_2 = -3. \end{aligned}$$

Por fim, substituímos os valores de  $y_1$  e  $y_2$  na equação  $(III')$ :

$$\begin{aligned} (III') : \quad & 4y_1 + y_2 + y_3 = 13 \\ \Rightarrow & 4 \times 4 + (-3) + y_3 = 13 \\ \Rightarrow & 16 - 3 + y_3 = 13 \\ \Rightarrow & 13 + y_3 = 13 \\ \Rightarrow & y_3 = 0. \end{aligned}$$

Concluimos então que a matriz vertical  $y$  é dada por

$$\begin{bmatrix} 4 \\ -3 \\ 0 \end{bmatrix}.$$

Como este sistema triangular é análogo ao visto no método anterior, concluimos que são necessárias **9** operações para resolvê-lo.

Para finalizar a resolução do sistema linear através da fatoração LU, basta resolver a equação  $U \times x = y$ , cuja forma, como vimos agora, é dada por

$$\begin{bmatrix} 1 & 2 & -3 \\ 0 & -1 & 10 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \\ 0 \end{bmatrix}.$$

Escrevendo agora a nova equação matricial na forma de um sistema de equações, temos:

$$\begin{cases} x_1 + 2x_2 - 3x_3 = 4 & (I'') \\ -x_2 + 10x_3 = -3 & (II'') \\ x_3 = 0 & (III'') \end{cases}.$$

Através da equação  $(III'')$ , encontramos  $x_3$ . Substituímos então esse valor na equação  $(II'')$ , temos:

$$\begin{aligned} (II'') : \quad & -x_2 + 10x_3 = -3 \\ \Rightarrow \quad & -x_2 + 10 \times 0 = -3 \\ \Rightarrow \quad & -x_2 + 0 = -3 \\ \Rightarrow \quad & -x_2 = -3 \\ \Rightarrow \quad & x_2 = 3. \end{aligned}$$

Finalmente, substituímos os valores de  $x_2$  e  $x_3$  na equação  $(I'')$ :

$$\begin{aligned} (I'') : \quad & x_1 + 2x_2 - 3x_3 = 4 \\ \Rightarrow \quad & x_1 + 2 \times 3 - 3 \times 0 = 4 \\ \Rightarrow \quad & x_1 + 6 - 0 = 4 \\ \Rightarrow \quad & x_1 + 6 = 4 \\ \Rightarrow \quad & x_1 = -2. \end{aligned}$$

Nossa solução, como esperado, é novamente a terna ordenada  $(-2, 3, 0)$ .

Novamente, **9** operações são necessárias para a resolução do sistema triangular.

Totalizando a fase de eliminação e ambas as fases de resolução, concluímos que o sistema linear, através da fatoração LU, necessita de  $13 + 9 + 9 = \mathbf{31}$  operações para ser solucionado.

No caso de um sistema  $n \times n$ , temos:

- Para eliminar os termos  $a_{k1}$ ,  $k \in \{2, \dots, n\}$ , temos que calcular o coeficiente  $\frac{a_{k1}}{a_{11}}$ , multiplicá-lo com os termos  $a_{12}, \dots, a_{1n}$ , para em seguida somar cada um com o seu termo respectivo da  $k$ -ésima linha. Totalizam-se assim  $n$  multiplicações e  $n - 1$  adições. Como o processo deve ser repetido para todas as linhas, temos que, nesta etapa, são realizadas  $(n) \times (n - 1)$  multiplicações e  $(n - 1) \times (n - 1) = (n - 1)^2$  adições;
- Para os termos  $a_{k2}^{(1)}$ ,  $k \in \{3, \dots, n\}$ , o processo é análogo. Assim, realizamos  $n - 1$  multiplicações e  $n - 2$  adições. Repetindo o processo para todas as  $n - 2$  linhas, temos  $(n - 1) \times (n - 2)$  multiplicações e  $(n - 2) \times (n - 2) = (n - 2)^2$  adições;

⋮

- Para o termo  $a_{n,n-1}^{(n-2)}$ , único termo da  $(n-1)$ -ésima coluna a ser anulado, e último termo do processo, são necessárias 2 multiplicações ( $\frac{a_{n,n-1}^{(n-2)}}{a_{n-1,n-1}^{(n-2)}}$ , e sua multiplicação com os termos  $a_{n-1,n}^{(n-2)}$ ) e 1 adição (o termo citado anteriormente, somado com seu termo respectivo da  $n$ -ésima linha). Assim, temos  $2 \times 1$  multiplicações e  $1 \times 1 = 1^2$  adição;
- Agora com os termos das matrizes  $L$  e  $U$  calculados, precisamos resolver dois sistemas triangulares  $n \times n$ . Já vimos que são necessárias  $n^2$  operações para sua resolução de cada. Logo,  $2n^2$  operações no total.

Somando as contagens das etapas necessárias para a obtenção dos termos das matrizes  $L$  e  $U$ , temos:

$$\begin{aligned} \#(+\text{Termos LU}) &= (n-1)^2 + (n-2)^2 + \dots + 1^2 \\ &= \sum_{i=1}^{n-1} i^2 \\ &= \sum_{i=1}^n i^2 - n^2 \\ &= \frac{n \times (n+1) \times (2n+1)}{6} - n^2 \\ &= \frac{n \times (n+1) \times (2n+1)}{6} - n^2 \\ &= \frac{2n^3 + 3n^2 + n}{6} - n^2 \\ &= \frac{2n^3 + 3n^2 + n - 6n^2}{6} \\ &= \frac{2n^3 - 3n^2 + n}{6}; \end{aligned}$$

$$\begin{aligned}
\#(\times \text{Termos LU}) &= n \times (n-1) + (n-1) \times (n-2) + \dots + 2 \times 1 \\
&= \sum_{i=1}^{n-1} [i \times (i+1)] \\
&= \sum_{i=1}^{n-1} [i^2 + i] \\
&= \sum_{i=1}^{n-1} i^2 + \sum_{i=1}^{n-1} i \\
&= \frac{n \times (n+1) \times (2n+1)}{6} - n^2 + \frac{n \times (n-1)}{2} \\
&= \frac{2n^3 + 3n^2 + n}{6} - n^2 + \frac{n^2 - n}{2} \\
&= \frac{2n^3 + 3n^2 + n - 6n^2 + 3n^2 - 3n}{6} \\
&= \frac{2n^3 - 2n}{6} \\
&= \frac{n^3 - n}{3}.
\end{aligned}$$

Assim, o total de operações necessárias para se obter  $L$  e  $U$  é dado por:

$$\begin{aligned}
\#(\text{Termos LU}) &= \frac{2n^3 - 3n^2 + n}{6} + \frac{n^3 - n}{3} \\
&= \frac{2n^3 - 3n^2 + n + 2n^3 - 2n}{6} \\
&= \frac{4n^3 - 3n^2 - n}{6}.
\end{aligned} \tag{11}$$

No processo de obtenção do valor das variáveis, temos:

$$\begin{aligned}
\#(\text{Obter LU}) &= n^2 + n^2 \\
&= 2n^2.
\end{aligned} \tag{12}$$

Encontramos assim a fórmula geral do número de operações na resolução de um sistema através da Fatoração LU, somando as equações (11) e (12):

$$\begin{aligned}
\#(\text{FatLU}) &= \frac{4n^3 - 3n^2 - n}{6} + 2n^2 \\
&= \frac{4n^3 - 3n^2 - n + 12n^2}{6} \\
&= \frac{4n^3 + 9n^2 - n}{6}.
\end{aligned}$$

## 2.4 Regra de Cramer

Utilizando-se da Regra de Cramer, é possível solucionar um sistema de equações lineares apenas através do cálculo do determinante de algumas matrizes. É comumente

apresentada aos alunos do Ensino Médio, como uma das aplicações dos determinantes.

O valor da variável  $x_k$ ,  $\forall k \in \{1, \dots, n\}$ , será dado pela fórmula

$$x_k = \frac{\det A_{x_k}}{\det A},$$

onde  $A$  é a matriz dos coeficientes, e  $A_{x_n}$  é a matriz  $A$  com a coluna dos coeficientes relativos a  $x_k$  substituída pelo conteúdo da matriz coluna  $b$ , correspondente aos termos independentes do sistema. Assim,

$$A_{x_k} = \begin{bmatrix} a_{11} & \cdots & a_{1,k-1} & b_1 & a_{1,k+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,k-1} & b_2 & a_{2,k+1} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,k-1} & b_n & a_{n,k+1} & \cdots & a_{nn} \end{bmatrix}.$$

Começaremos, então, com a resolução do problema proposto, agora utilizando-se da Regra de Cramer. Já calculamos no início do capítulo o determinante da matriz  $A$ . Assim, temos  $\det A = -1$ .

Calcularemos, agora, os determinantes das matrizes  $A_{x_1}$ ,  $A_{x_2}$  e  $A_{x_3}$ . Fazendo a substituição do vetor coluna  $b$  na devida coluna, temos que

$$A_{x_1} = \begin{bmatrix} 4 & 2 & -3 \\ 5 & 3 & 4 \\ 13 & 7 & -1 \end{bmatrix}, \quad A_{x_2} = \begin{bmatrix} 1 & 4 & -3 \\ 2 & 5 & 4 \\ 4 & 13 & -1 \end{bmatrix} \quad \text{e} \quad A_{x_3} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 3 & 5 \\ 4 & 7 & 13 \end{bmatrix}.$$

Calculando  $\det A_{x_1}$ , temos

$$\begin{aligned} \det A_{x_1} &= 4 \times 3 \times (-1) - 4 \times 4 \times 7 - 2 \times 5 \times (-1) \\ &\quad + 2 \times 4 \times 13 - (-3) \times 3 \times 13 + (-3) \times 5 \times 7 \\ &= -12 - 112 + 10 + 104 + 117 - 105 \\ &= -229 + 231 \\ &= 2. \end{aligned}$$

*Em cada parcela do somatório para o cálculo do determinante, temos um produto de 3 números, sendo realizadas assim 2 multiplicações. Temos um total de 6 parcelas, o que nos dá um total de 12 multiplicações. Para juntar as 6 parcelas, precisamos realizar 5 adições. Assim, o total de operações para obter o determinante da matriz é igual a 17.*

Para o determinante da matriz  $A_{x_2}$ , temos

$$\begin{aligned} \det A_{x_2} &= 1 \times 5 \times (-1) - 1 \times 4 \times 13 - 4 \times 2 \times (-1) \\ &\quad + 4 \times 4 \times 4 - (-3) \times 5 \times 4 + (-3) \times 2 \times 13 \\ &= -5 - 52 + 8 + 64 + 60 - 78 \\ &= -135 + 132 \\ &= -3. \end{aligned}$$

Finalizando, com o determinante de  $A_{x_3}$ , temos

$$\begin{aligned} \det A_{x_3} &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} + a_{13}a_{21}a_{32} \\ &= 1 \times 3 \times 13 - 1 \times 5 \times 7 - 2 \times 2 \times 13 \\ &\quad + 2 \times 5 \times 4 - 4 \times 3 \times 4 + 4 \times 2 \times 7 \\ &= 39 - 35 - 52 + 40 - 48 + 56 \\ &= 135 - 135 \\ &= 0. \end{aligned}$$

Como os processos são análogos, e tivemos que calcular o determinante de 4 matrizes ( $A$ ,  $A_{x_1}$ ,  $A_{x_2}$  e  $A_{x_3}$ ), foram necessárias  $4 \times 17 = \mathbf{68}$  operações nesta primeira parte do método.

Utilizando a fórmula para o cálculo das variáveis, temos:

$$x_1 = \frac{\det A_{x_1}}{\det A} = \frac{2}{-1} = -2;$$

$$x_2 = \frac{\det A_{x_2}}{\det A} = \frac{-3}{-1} = 3;$$

$$x_3 = \frac{\det A_{x_3}}{\det A} = \frac{0}{-1} = 0.$$

*Para a obtenção de cada variável, é necessária 1 multiplicação. Assim, fizemos 3 multiplicações no total, nessa fase.*

No processo de obtenção das variáveis, foram necessárias **3** operações.

A terna ordenada  $(-2, 3, 0)$  é mais uma vez obtida como solução do problema. Desta vez, foram necessárias  $68 + 3 = \mathbf{71}$  operações.

Generalizando para um sistema  $n \times n$ , temos que em cada parcela do somatório para o cálculo de um determinante, temos um produto de  $n$  números, sendo realizadas assim  $n - 1$  multiplicações. Temos um total de  $n!$  parcelas, o que nos dá um total de  $n!(n - 1)$  multiplicações. Para juntar as  $n!$  parcelas,  $n! - 1$  adições. Assim, o total de operações para obter o determinante de uma matriz é igual a

$$\#(\text{Determinante}) = n!(n - 1) + n! - 1 = n \times n! - 1.$$

Como é necessário o cálculo do determinante de  $n+1$  matrizes ( $A, A_{x_1}, \dots, A_{x_n}$ ), temos que realizar, para a fase de obtenção dos determinantes no método de Cramer, um total de

$$\#(\text{Determin. Cramer}) = (n + 1)(n \times n! - 1) = n(n + 1)! - (n + 1) \text{ operações.}$$

Ao final, são realizadas  $n$  divisões, para finalmente obter o valor de cada uma das  $n$  variáveis. Assim, ao utilizar a regra de Cramer na resolução de um sistema linear, são necessárias

$$\#(\text{Cramer}) = n(n+1)! - (n+1) + n = n(n+1)! - n + n - 1 = n(n+1)! - 1 \text{ operações.}$$

## 2.5 Comparando os Métodos Diretos

Vimos que, para um sistema linear  $3 \times 3$ , foram necessários, para cada um dos métodos:

- **Substituição:** 31 operações;
- **Eliminação Gaussiana:** 28 operações;
- **Fatoração LU:** 31 operações;
- **Regra de Cramer:** 71 operações.

Assim, o método mais vantajoso, a contar pelo seu custo computacional, para um sistema linear  $3 \times 3$  é a Eliminação Gaussiana. Vemos que a Regra de Cramer tem uma contagem muito superior às demais, mesmo para um sistema linear pequeno.

Ao analisar o caso de um sistema linear  $n \times n$ , observa-se que são necessários:

- **Substituição:**  $\frac{2n^3 + 6n^2 - 5n}{3}$  operações;
- **Eliminação Gaussiana:**  $\frac{4n^3 + 9n^2 - 7n}{6}$  operações;
- **Fatoração LU:**  $\frac{4n^3 + 9n^2 - n}{6}$  operações;
- **Regra de Cramer:**  $n(n+1)! - 1$  operações.

A Regra de Cramer tem a sua contagem na ordem de  $n(n+1)!$ , enquanto as demais estão na ordem de  $\frac{2n^3}{3}$ . Como  $n!$  tende ao infinito mais rapidamente que  $n^3$ , para valores de  $n$  suficientemente grandes, a contagem de operações da Regra de Cramer terá valores maiores do que as demais, sendo assim o método direto menos vantajoso.

Para determinar o de menor custo computacional, faremos a comparação dos demais métodos 2 a 2:

### 2.5.1 Substituição x Eliminação Gaussiana

Seja  $n_1 \in \mathbb{N}$ . Assim, temos que  $n_1 - 1 \geq 0$  e  $3n_1 > 0$ .

$$\begin{aligned} &\iff 3n_1(n_1 - 1) \geq 0 \\ &\iff 3n_1^2 - 3n_1 \geq 0 \\ &\iff 12n_1^2 - 10n_1 \geq 9n_1^2 - 7n_1 \\ \iff &\frac{4n_1^3 + 12n_1^2 - 10n_1}{6} \geq \frac{4n_1^3 + 9n_1^2 - 7n_1}{6} \\ \iff &\frac{2n_1^3 + 6n_1^2 - 5n_1}{3} \geq \frac{4n_1^3 + 9n_1^2 - 7n_1}{6}. \end{aligned}$$

Assim, temos que, para qualquer número natural  $n \geq 2$ , o número de operações necessárias para a resolução de um sistema linear através do método da Substituição é maior do que ou igual ao número necessário para a Eliminação Gaussiana.

Conclui-se que a Eliminação de Gauss é mais vantajosa que a Substituição para qualquer sistema linear  $k \times k$ ,  $k \in \mathbb{N}$ ,  $k \geq 2$ .

### 2.5.2 Substituição x Fatoração LU

Agora, vamos supor  $n_2 \in \mathbb{N}$ ,  $n_2 \geq 3$ . Assim,  $n_2 - 3 \geq 0$  e  $3n_2 > 0$ .

$$\begin{aligned} &\iff 3n_2(n_2 - 3) \geq 0 \\ &\iff 3n_2^2 - 9n_2 \geq 0 \\ &\iff 12n_2^2 - 10n_2 \geq 9n_2^2 - n_2 \\ \iff &\frac{4n_2^3 + 12n_2^2 - 10n_2}{6} \geq \frac{4n_2^3 + 9n_2^2 - n_2}{6} \\ \iff &\frac{2n_2^3 + 6n_2^2 - 5n_2}{3} \geq \frac{4n_2^3 + 9n_2^2 - n_2}{6}. \end{aligned}$$

Assim, temos que, para qualquer número natural  $n \geq 3$ , o número de operações necessárias para a resolução de um sistema linear através do método da Substituição é maior do que ou igual ao número necessário para a Fatoração LU. Os métodos são igualmente vantajosos para  $n = 3$ .

Logo, a Fatoração LU é mais vantajosa que a Substituição para qualquer sistema linear  $k \times k$ ,  $k \in \mathbb{N}$ ,  $k \geq 4$ .

**Observação 2.7.** No caso de um sistema  $2 \times 2$ , temos que a Fatoração LU requer

$$\frac{4 \times 2^3 + 9 \times 2^2 - 2}{6} = \frac{32 + 36 - 2}{6} = \frac{66}{6} = 11 \text{ operações,}$$

enquanto a Substituição requer

$$\frac{2 \times 2^3 + 6 \times 2^2 - 5 \times 2}{3} = \frac{16 + 24 - 10}{3} = \frac{30}{3} = 10 \text{ operações.}$$

Logo, a Substituição é mais vantajosa.

### 2.5.3 Fatoração LU x Eliminação Gaussiana

Finalmente, tomemos  $n_3 \in \mathbb{N}$ . Assim,

$$\begin{aligned} & \iff n_3 > 0 \\ & \iff 6n_3 > 0 \\ & \iff -n_3 > -7n_3 \\ \iff & \frac{4n_3^3 + 9n_3^2 - n_3}{6} > \frac{4n_3^3 + 9n_3^2 - 7n_3}{6}. \end{aligned}$$

Assim, a Eliminação Gaussiana é mais vantajosa que a Fatoração LU para qualquer sistema linear  $k \times k$ ,  $k \in \mathbb{N}$ ,  $k \geq 2$ .

**Observação 2.8.** *O sistema  $1 \times 1$  é trivial, constituindo uma equação simples, e portanto não é objeto do nosso estudo.*

Temos, assim, que para qualquer sistema linear  $n \times n$ , a Eliminação Gaussiana é o método direto com o menor custo computacional, logo, mais vantajoso. Entretanto, os outros dois métodos têm as suas vantagens.

A Fatoração LU, por excluir o vetor coluna  $b$  do escalonamento, é vantajosa quando queremos repetir a resolução do mesmo sistema com termos independentes diferentes. A modificação no sistema será percebida apenas quando estivermos resolvendo os dois sistemas triangulares, pois a matriz a ser escalonada é a mesma.

A Substituição não tem vantagem alguma do ponto de vista computacional, mas pode ser vantajoso ao aluno que tenha certa dificuldade com matrizes. Sua resolução envolve apenas um conhecimento matemático de resolução de equações.

### 3 Métodos Iterativos de Resolução

Os métodos iterativos, diferentemente dos métodos diretos, não encontram a solução diretamente, e sim gerando uma sequência de vetores  $\{\vec{x}^{(k)}\}$ , através de uma fórmula recursiva, dada uma aproximação inicial  $\vec{x}^{(0)}$ , que pode ou não convergir para a solução, caso exista. Quando convergente, a sequência tenderá à solução quando  $k \rightarrow \infty$ .

Devemos transformar o sistema  $A \times x = b$  em uma nova forma matricial, de forma a termos  $x = f(x)$ ,  $f$  uma função matricial. Assim, faremos de  $x$  variável independente o termo  $x^{(k)}$  e de  $x$  variável dependente o termo  $x^{(k+1)}$ .

Para isso, poderíamos definir a matriz  $B = I - A \Rightarrow A = I - B$ ,  $I$  matriz identidade. Assim, teríamos

$$(B - I)x = b \Rightarrow Bx - x = b \Rightarrow x = Bx - b,$$

e nossa fórmula recursiva poderia ser dada por

$$x^{(k+1)} = B \times x^{(k)} - b.$$

No entanto, não há garantia de convergência para a sequência de vetores, quando tomamos a função matricial  $f(x) = Bx - b$ . Devemos, então, encontrar duas matrizes  $C$  e  $g$ , a primeira da forma  $n \times n$ , e a outra da forma  $n \times 1$ , de tal forma que, fazendo a fórmula recursiva

$$x^{(k+1)} = C \times x^{(k)} + g,$$

e cumprindo um determinado critério de convergência, teremos a garantia que obteremos a solução quando  $k \rightarrow \infty$ .

Veremos posteriormente como deve ser o formato das matrizes  $C$  e  $g$ .

Por se tratar de uma fórmula recursiva, onde o termo seguinte depende do termo anterior, devemos determinar o vetor  $\vec{x}^{(0)}$  que iniciará esta sequência. Novamente, veremos como é feita essa escolha mais adiante.

Analisaremos dois métodos iterativos:

1. Método de Gauss-Jacobi;
2. Método de Gauss-Seidel.

Uma condição necessária e suficiente para a garantia de convergência destes métodos, assegurando assim que a sequência alcançará a solução exata (quando esta existir), é dada por:

**Teorema 3.1.** *Sejam  $\lambda_i$  os autovalores da matriz  $C$ . Se  $\max |\lambda_i| < 1$ , então a convergência está garantida, independente da escolha feita para  $\vec{x}^{(0)}$ .*

Maiores informações sobre este teorema podem ser encontradas em [5].

Como o processo de convergência à solução pode ser demorado, são estipulados critérios de parada para os métodos iterativos. Podem ser adotados três critérios:

**a. Teste do erro absoluto:** Podemos repetir o processo de iteração até que o vetor  $x^{(k)}$  esteja suficientemente próximo do vetor  $x^{(k-1)}$ . Para isso, calculamos a distância entre os termos respectivos de  $x^{(k)}$  e  $x^{(k-1)}$ , fazendo  $d_i^{(k)} = |x_i^{(k)} - x_i^{(k-1)}|$ . A distância será dada por

$$d^{(k)} = \max_{1 \leq i \leq n} \{d_i^{(k)}\}.$$

Assim, dado  $\epsilon > 0$ , o processo é interrompido quando  $d^{(k)} < \epsilon$ .

**Observação 3.1.** *O teste do erro absoluto garante uma precisão satisfatória apenas quando estamos trabalhando com valores grandes. De fato, se o resultado obtido para uma variável  $x_k$  na  $i$ -ésima iteração é  $x_k^{(i)} = 1.123.479,6$ , e na iteração seguinte foi obtido o valor  $x_k^{(i+1)} = 1.123.478,9$ , o erro absoluto será dado por  $|1.123.478,9 - 1.123.479,6| = 0,8$ . Pelas grandezas que estão sendo trabalhadas, o erro de 0,8 dá uma boa precisão para o resultado exato, visto que a oscilação entre as iterações está pequena.*

*Para uma variável  $x_j$ , temos  $x_j^{(i)} = 2,1$  e  $x_j^{(i+1)} = 2,9$ . Assim, o erro é  $|2,9 - 2,1| = 0,8$ , mesmo valor encontrado no erro absoluto do primeiro exemplo. No entanto, como trata-se de grandezas menores, o erro caracteriza uma maior imprecisão.*

**b. Teste do erro relativo:** De forma análoga ao teste acima, é calculado  $d^{(k)}$ . Cessamos as iterações quando, dado  $\epsilon > 0$ , temos que a distância relativa, dada por

$$d_r^{(k)} = \frac{d^{(k)}}{\max_{1 \leq i \leq n} \{|x_i^{(k)}|\}},$$

é menor que a precisão  $\epsilon$  estipulada.

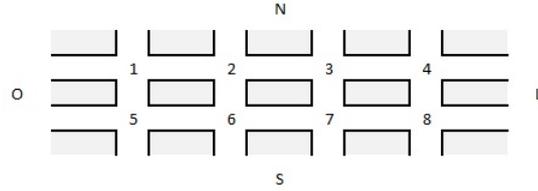
**Observação 3.2.** *Ao dividir o erro absoluto pelo módulo da maior das variáveis da última aproximação obtida, estamos calculando este erro como uma porcentagem do valor da variável. Assim, analisa-se quão incorreto é o valor calculado se comparado ao valor considerado correto naquele instante. Dessa forma, aumenta-se a precisão nos cálculos, comparado com o erro absoluto.*

**c. Número máximo de iterações:** É estipulado um número máximo de iterações, cessadas quando este número é alcançado. No entanto, definir o número máximo de iterações como critério de parada não garante uma aproximação satisfatória, pois a rapidez em que a convergência acontece varia de um sistema para outro.

Dessa forma, na resolução de problemas através dos métodos iterativos, utilizaremos o **teste do erro relativo** como critério de parada.

Consideremos então a seguinte situação-problema, que corresponde a uma aplicação dos sistemas lineares na Probabilidade:

Um rato é colocado em um labirinto retangular de corredores, como na figura abaixo. Em cada interseção, ele escolhe uma direção ao acaso e segue até a interseção seguinte, onde escolhe novamente ao acaso nova direção e assim por diante. Qual a probabilidade de o rato, estando na interseção  $i$ , sair eventualmente pelo sul?



Denotemos por  $P_i$  a probabilidade de o rato sair do labirinto pelo sul, partindo da interseção  $i$ ,  $i \in \{1, \dots, 8\}$ .

Estando na interseção 1, o rato pode:

- Tomar a direção norte e sair do labirinto (*I*);
- Tomar a direção oeste e sair do labirinto (*II*);
- Tomar a direção sul e chegar na interseção 5 (*III*);
- Tomar a direção leste e chegar na interseção 2 (*IV*).

Em (*I*), o rato saiu do labirinto pelo norte, logo sua chance de sair pelo sul é igual a 0. (*II*) é análogo a (*I*), logo a probabilidade também é igual a 0. Em (*III*), o rato tem uma nova localização, a interseção 5. Assim, a chance do rato sair do labirinto pelo sul é igual à probabilidade dele sair pelo sul a partir da interseção 5, logo,  $P_5$ . Em (*IV*), analogamente a (*III*), a probabilidade é de  $P_2$ . Como a chance do rato pegar um dos caminhos é dada por  $\frac{1}{4}$ , temos que:

$$P_1 = \frac{0 + 0 + P_2 + P_5}{4} \Rightarrow 4P_1 - P_2 - P_5 = 0.$$

Repetindo o processo para cada uma das interseções, temos:

$$P_2 = \frac{0 + P_1 + P_3 + P_6}{4} \Rightarrow -P_1 + 4P_2 - P_3 - P_6 = 0;$$

$$P_3 = \frac{0 + P_2 + P_4 + P_7}{4} \Rightarrow -P_2 + 4P_3 - P_4 - P_7 = 0;$$

$$P_4 = \frac{0 + 0 + P_3 + P_8}{4} \Rightarrow -P_3 + 4P_4 - P_8 = 0;$$

$$P_5 = \frac{0 + 1 + P_1 + P_6}{4} \Rightarrow -P_1 + 4P_5 - P_6 = 1;$$

$$P_6 = \frac{1 + P_2 + P_5 + P_7}{4} \Rightarrow -P_2 - P_5 + 4P_6 - P_7 = 1;$$

$$P_7 = \frac{1+P_3+P_6+P_8}{4} \Rightarrow -P_3 - P_6 + 4P_7 - P_8 = 1;$$

$$P_8 = \frac{1+0+P_4+P_7}{4} \Rightarrow -P_4 - P_7 + 4P_8 = 1.$$

Assim, temos um sistema linear com 8 incógnitas e 8 equações, da forma

$$\left\{ \begin{array}{l} 4P_1 - P_2 - P_5 = 0 \\ -P_1 + 4P_2 - P_3 - P_6 = 0 \\ -P_2 + 4P_3 - P_4 - P_7 = 0 \\ -P_3 + 4P_4 - P_8 = 0 \\ -P_1 + 4P_5 - P_6 = 1 \\ -P_2 - P_5 + 4P_6 - P_7 = 1 \\ -P_3 - P_6 + 4P_7 - P_8 = 1 \\ -P_4 - P_7 + 4P_8 = 1 \end{array} \right. ,$$

com

$$A = \begin{bmatrix} 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 4 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 4 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & 4 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 4 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

**Observação 3.3.** Podemos ver, através desse exemplo, que, na Probabilidade, quando a chance de um evento se concretizar depende da concretização de outros eventos do mesmo espaço amostral, uma saída para a resolução deste problema é um sistema linear, onde a probabilidade de cada evento corresponde a uma incógnita do sistema.

Veremos agora como cada um dos dois métodos funciona na sua resolução, além de entender as minúcias de seu algoritmo.

### 3.1 Método de Gauss-Jacobi

Para montarmos a fórmula de recorrência no método de Gauss-Jacobi, basta isolar a incógnita  $x_i$  na  $i$ -ésima linha,  $i \in \{1, \dots, n\}$ . Assim, reescrevemos o sistema linear  $n \times n$  da seguinte forma:

$$\left\{ \begin{array}{l} x_1 = \frac{b_1 - a_{12}x_2 - a_{13}x_3 - \dots - a_{1n}x_n}{a_{11}} \\ x_2 = \frac{b_2 - a_{21}x_1 - a_{23}x_3 - \dots - a_{2n}x_n}{a_{22}} \\ \vdots \\ x_{n-1} = \frac{b_{n-1} - a_{n-1,1}x_1 - \dots - a_{n-1,n-2}x_{n-2} + a_{n-1,n}x_n}{a_{n-1,n-1}} \\ x_n = \frac{b_n - a_{n1}x_1 - a_{n2}x_2 - \dots - a_{n,n-1}x_{n-1}}{a_{nn}} \end{array} \right. .$$

A fórmula de recursão é dada por

$$\left\{ \begin{array}{l} x_1^{(k+1)} = \frac{b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \dots - a_{1n}x_n^{(k)}}{a_{11}} \\ x_2^{(k+1)} = \frac{b_2 - a_{21}x_1^{(k)} - a_{23}x_3^{(k)} - \dots - a_{2n}x_n^{(k)}}{a_{22}} \\ \vdots \\ x_{n-1}^{(k+1)} = \frac{b_{n-1} - a_{n-1,1}x_1^{(k)} - \dots - a_{n-1,n-2}x_{n-2}^{(k)} + a_{n-1,n}x_n^{(k)}}{a_{n-1,n-1}} \\ x_n^{(k+1)} = \frac{b_n - a_{n1}^{(k)}x_1 - a_{n2}^{(k)}x_2 - \dots - a_{n,n-1}x_{n-1}^{(k)}}{a_{nn}} \end{array} \right.$$

Vemos que, a cada linha, é necessária a divisão pelo termo da diagonal principal correspondente a respectiva linha. Assim, é necessário que nenhum termo da diagonal principal seja nulo. Para que isso seja possível, já vimos que a troca entre linhas ou colunas é permitida, para adequar a matriz  $A$  à forma desejada.

Calculando uma determinada incógnita na  $k$ -ésima iteração, são necessárias  $n - 1$  multiplicações (das demais incógnitas com seus coeficientes),  $n - 1$  adições ( $n - 1$  termos com as variáveis e o termo independente) e 1 multiplicação (dividindo, finalmente, o valor obtido pelo respectivo coeficiente da diagonal principal). Assim, temos um total de  $n - 1 + n - 1 + 1 = 2n - 1$  operações. Como o processo se repete para todas as  $n$  incógnitas, temos um total de  $n \times (2n - 1) = 2n^2 - n$  operações em uma iteração.

**Observação 3.4.** *Como o número de iterações é incerto, dependendo tanto do  $\epsilon$  do erro quanto da aproximação inicial, não é possível generalizar uma fórmula para o cálculo de operações total do método.*

Escrevendo na forma matricial  $x^{(k+1)} = C \times x^{(k)} + g$ , temos

$$C = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & -\frac{a_{23}}{a_{22}} & \dots & -\frac{a_{2n}}{a_{22}} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ -\frac{a_{n-1,1}}{a_{n-1,n-1}} & \dots & -\frac{a_{n-1,n-2}}{a_{n-1,n-1}} & 0 & -\frac{a_{n-1,n}}{a_{n-1,n-1}} \\ -\frac{a_{n1}}{a_{nn}} & \dots & -\frac{a_{n,n-2}}{a_{nn}} & -\frac{a_{n,n-1}}{a_{nn}} & 0 \end{bmatrix} \quad \text{e} \quad g = \begin{bmatrix} -\frac{b_1}{a_{11}} \\ -\frac{b_2}{a_{22}} \\ \vdots \\ -\frac{b_n}{a_{nn}} \end{bmatrix}.$$

Para garantir a convergência através do método de Gauss-Jacobi, utilizaremos um método diferente do citado anteriormente:

**Teorema 3.2. (Critério das linhas)** *Considerando a matriz  $A$  do sistema linear, seja*

$\alpha_i = \frac{\sum_{j=1; j \neq i}^n |a_{ij}|}{|a_{ii}|}$ . *Se  $\alpha = \max_{1 \leq i \leq n} \{\alpha_i\} < 1$ , então a sequência  $\{x^{(k)}\}$  gerada pelo método de Gauss-Jacobi converge para a solução do sistema, independente da aproximação inicial  $x^{(0)}$  escolhida.*

Maiores informações sobre este teorema podem ser encontradas em [11].

**Observação 3.5.** *Tal critério de convergência é suficiente para que a sequência de vetores convirja para a solução, mas não é necessário.*

**Observação 3.6.** *Novamente, a troca de linhas ou colunas pode ser utilizada a fim de adequar a matriz  $A$  ao critério.*

Agora que já conhecemos o funcionamento do método, vamos à resolução do sistema linear  $A \times x = b$  do problema proposto, lembrando que

$$A = \begin{bmatrix} 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 4 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 4 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & 4 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 4 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Pelo **critério das linhas**, vemos que a sequência de vetores convergirá para a solução do sistema. De fato:

$$\alpha_1 = \frac{|-1|+|0|+|0|+|-1|+|0|+|0|+|0|}{4} = \frac{1+1}{4} = \frac{2}{4} = 0,5.$$

Analogamente, temos

$$\alpha_4 = \alpha_5 = \alpha_8 = \frac{|-1|+|-1|}{|4|} = \frac{1+1}{4} = \frac{2}{4} = 0,5;$$

$$\alpha_2 = \alpha_3 = \alpha_6 = \alpha_7 = \frac{|-1|+|-1|+|-1|}{|4|} = \frac{1+1+1}{4} = \frac{3}{4} = 0,75.$$

$$\Rightarrow \alpha = \max\{0,5; 0,75\} = 0,75 < 1.$$

Vamos definir  $x^{(k)} = \{P_1^{(k)}; P_2^{(k)}; P_3^{(k)}; P_4^{(k)}; P_5^{(k)}; P_6^{(k)}; P_7^{(k)}; P_8^{(k)}\}$ .

Como a escolha de  $x^{(0)}$  é arbitrária, e como sabemos que a probabilidade de um evento varia de 0 a 1, tomemos

$$x^{(0)} = \{0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5\}.$$

Assim,  $P_i^{(0)} = 0,5, i \in \{1, \dots, 8\}$ .

Obtendo a fórmula de iteração:

$$\left\{ \begin{array}{l} 4P_1 - P_2 - P_5 = 0 \\ -P_1 + 4P_2 - P_3 - P_6 = 0 \\ -P_2 + 4P_3 - P_4 - P_7 = 0 \\ -P_3 + 4P_4 - P_8 = 0 \\ -P_1 + 4P_5 - P_6 = 1 \\ -P_2 - P_5 + 4P_6 - P_7 = 1 \\ -P_3 - P_6 + 4P_7 - P_8 = 1 \\ -P_4 - P_7 + 4P_8 = 1 \end{array} \right. \Rightarrow \left\{ \begin{array}{l} P_1^{(k+1)} = \frac{P_2^{(k)} + P_5^{(k)}}{4} \\ P_2^{(k+1)} = \frac{P_1^{(k)} + P_3^{(k)} + P_6^{(k)}}{4} \\ P_3^{(k+1)} = \frac{P_2^{(k)} + P_4^{(k)} + P_7^{(k)}}{4} \\ P_4^{(k+1)} = \frac{P_3^{(k)} + P_8^{(k)}}{4} \\ P_5^{(k+1)} = \frac{P_1^{(k)} + P_6^{(k)} + 1}{4} \\ P_6^{(k+1)} = \frac{P_2^{(k)} + P_5^{(k)} + P_7^{(k)} + 1}{4} \\ P_7^{(k+1)} = \frac{P_3^{(k)} + P_6^{(k)} + P_8^{(k)} + 1}{4} \\ P_8^{(k+1)} = \frac{P_4^{(k)} + P_7^{(k)} + 1}{4} \end{array} \right.$$

Assim, temos a fórmula de iteração  $x^{(k+1)} = C \times x^{(k)} + g$ , onde

$$C = \begin{bmatrix} 0 & 0,25 & 0 & 0 & 0,25 & 0 & 0 & 0 \\ 0,25 & 0 & 0,25 & 0 & 0 & 0,25 & 0 & 0 \\ 0 & 0,25 & 0 & 0,25 & 0 & 0 & 0,25 & 0 \\ 0 & 0 & 0,25 & 0 & 0 & 0 & 0 & 0,25 \\ 0,25 & 0 & 0 & 0 & 0 & 0,25 & 0 & 0 \\ 0 & 0,25 & 0 & 0 & 0,25 & 0 & 0,25 & 0 \\ 0 & 0 & 0,25 & 0 & 0 & 0,25 & 0 & 0,25 \\ 0 & 0 & 0 & 0,25 & 0 & 0 & 0,25 & 0 \end{bmatrix} \quad \text{e} \quad g = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0,25 \\ 0,25 \\ 0,25 \\ 0,25 \end{bmatrix}.$$

1ª Iteração:

$$P_1^{(1)} = \frac{P_2^{(0)} + P_5^{(0)}}{4} = \frac{0,5 + 0,5}{4} = \frac{1}{4} = 0,25;$$

$$P_2^{(1)} = \frac{P_1^{(0)} + P_3^{(0)} + P_6^{(0)}}{4} = \frac{0,5 + 0,5 + 0,5}{4} = \frac{1,5}{4} = 0,375;$$

$$P_3^{(1)} = \frac{P_2^{(0)} + P_4^{(0)} + P_7^{(0)}}{4} = \frac{0,5 + 0,5 + 0,5}{4} = \frac{1,5}{4} = 0,375;$$

$$P_4^{(1)} = \frac{P_3^{(0)} + P_8^{(0)}}{4} = \frac{0,5 + 0,5}{4} = \frac{1}{4} = 0,25;$$

$$P_5^{(1)} = \frac{P_1^{(0)} + P_6^{(0)} + 1}{4} = \frac{0,5 + 0,5 + 1}{4} = \frac{2}{4} = 0,5;$$

$$P_6^{(1)} = \frac{P_2^{(0)} + P_5^{(0)} + P_7^{(0)} + 1}{4} = \frac{0,5 + 0,5 + 0,5 + 1}{4} = \frac{2,5}{4} = 0,625;$$

$$P_7^{(1)} = \frac{P_3^{(0)} + P_6^{(0)} + P_8^{(0)} + 1}{4} = \frac{0,5 + 0,5 + 0,5 + 1}{4} = \frac{2,5}{4} = 0,625;$$

$$P_8^{(1)} = \frac{P_4^{(0)} + P_7^{(0)} + 1}{4} = \frac{0,5 + 0,5 + 1}{4} = \frac{2}{4} = 0,5.$$

$$\Rightarrow x^{(1)} = \{0,25; 0,375; 0,375; 0,25; 0,5; 0,625; 0,625; 0,5\}$$

Calculando o erro relativo na 1<sup>a</sup> Iteração:

$$d_1^{(1)} = |P_1^{(1)} - P_1^{(0)}| = |0,25 - 0,5| = |-0,25| = 0,25;$$

$$d_2^{(1)} = |P_2^{(1)} - P_2^{(0)}| = |0,375 - 0,5| = |-0,125| = 0,125;$$

$$d_3^{(1)} = |P_3^{(1)} - P_3^{(0)}| = |0,375 - 0,5| = |-0,125| = 0,125;$$

$$d_4^{(1)} = |P_4^{(1)} - P_4^{(0)}| = |0,25 - 0,5| = |-0,25| = 0,25;$$

$$d_5^{(1)} = |P_5^{(1)} - P_5^{(0)}| = |0,5 - 0,5| = |0| = 0;$$

$$d_6^{(1)} = |P_6^{(1)} - P_6^{(0)}| = |0,625 - 0,5| = |0,125| = 0,125;$$

$$d_7^{(1)} = |P_7^{(1)} - P_7^{(0)}| = |0,625 - 0,5| = |0,125| = 0,125;$$

$$d_8^{(1)} = |P_8^{(1)} - P_8^{(0)}| = |0,5 - 0,5| = |0| = 0$$

$$\Rightarrow d^{(1)} = \max_{1 \leq i \leq 8} \{d_i^{(1)}\} = 0,25$$

$$\Rightarrow d_r^{(1)} = \frac{0,25}{\max_{1 \leq i \leq 8} \{P_i^{(1)}\}} = \frac{0,25}{0,625} = 0,4.$$

Prosseguindo analogamente nas demais iterações, obtemos:

2<sup>a</sup> Iteração:

$$x^{(2)} = \{0,21875; 0,3125; 0,3125; 0,21875; 0,46875; 0,625; 0,625; 0,46875\};$$

$$d_r^{(2)} = 0,1.$$

3<sup>a</sup> Iteração:

$$x^{(3)} = \{0,1953125; 0,2890625; 0,2890625; 0,1953125; \\ 0,4609375; 0,6015625; 0,6015625; 0,4609375\};$$

$$d_r^{(3)} = 0,038961039.$$

4ª Iteração:

$$x^{(4)} = \{0,1875; 0,271484375; 0,271484375; 0,1875; 0,44921875; 0,587890625; 0,587890625; 0,44921875\};$$

$$d_r^{(4)} = 0,029900332.$$

Nas tabelas a seguir, calculadas com o auxílio de uma planilha eletrônica, vemos como se comporta a convergência, iteração após iteração:

Número da Iteração	P1	P2	P3	P4	P5	P6	P7	P8	Erro Relativo
11	0,169364095	0,243633747	0,243633747	0,169364095	0,432524323	0,559421778	0,559421778	0,432524323	0,001446054
12	0,169039518	0,243104905	0,243104905	0,169039518	0,432196468	0,558894962	0,558894962	0,432196468	0,000946228
13	0,168825343	0,242759846	0,242759846	0,168825343	0,43198362	0,558549084	0,558549084	0,43198362	0,000619244

Vemos que, a partir da 12ª iteração, o erro relativo torna-se menor que 0,001.

Número da Iteração	P1	P2	P3	P4	P5	P6	P7	P8	Erro Relativo
48	0,168421053	0,242105263	0,242105263	0,168421053	0,431578948	0,557894737	0,557894737	0,431578948	2,23301E-10
49	0,168421053	0,242105263	0,242105263	0,168421053	0,431578947	0,557894737	0,557894737	0,431578947	1,46153E-10
50	0,168421053	0,242105263	0,242105263	0,168421053	0,431578947	0,557894737	0,557894737	0,431578947	9,56584E-11

Da 49ª iteração em diante, o computador fixa a solução em 9 casas decimais, para cada incógnita, ou seja, o erro é menor do que  $10^{-9}$ .

Número da Iteração	P1	P2	P3	P4	P5	P6	P7	P8	Erro Relativo
85	0,168421053	0,242105263	0,242105263	0,168421053	0,431578947	0,557894737	0,557894737	0,431578947	4,97506E-17
86	0,168421053	0,242105263	0,242105263	0,168421053	0,431578947	0,557894737	0,557894737	0,431578947	0
87	0,168421053	0,242105263	0,242105263	0,168421053	0,431578947	0,557894737	0,557894737	0,431578947	0

A partir da 86ª, o erro já é dado como 0.

A solução, assim, pode ser dada por

$$x = \{0,168421053; 0,242105263; 0,242105263; 0,168421053; 0,431578947; 0,557894737; 0,557894737; 0,431578947\}.$$

## 3.2 Método de Gauss-Seidel

A estrutura do método de Gauss-Seidel é análoga à de Gauss-Jacobi. No entanto, quando calculamos o termo  $x_j^{(k+1)}$ , levamos em conta os termos  $x_i^{(k+1)}$  já calculados previamente, para os valores de  $i$  tais que  $i \in \{1, \dots, j-1\}$ , e  $x_i^{(k)}$ , para os valores de  $i$  tais

que  $i \in \{j + 1, \dots, n\}$ . Assim, a fórmula de recursão é dada por

$$\left\{ \begin{array}{l} x_1^{(k+1)} = \frac{b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \dots - a_{1n}x_n^{(k)}}{a_{11}} \\ x_2^{(k+1)} = \frac{b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} - \dots - a_{2n}x_n^{(k)}}{a_{22}} \\ \vdots \\ x_{n-1}^{(k+1)} = \frac{b_{n-1} - a_{n-1,1}x_1^{(k+1)} - \dots - a_{n-1,n-2}x_{n-2}^{(k+1)} - a_{n-1,n}x_n^{(k)}}{a_{n-1,n-1}} \\ x_n^{(k+1)} = \frac{b_n - a_{n1}^{(k+1)}x_1 - a_{n2}^{(k+1)}x_2 - \dots - a_{n,n-1}x_{n-1}^{(k+1)}}{a_{nn}} \end{array} \right.$$

Como a estrutura da fórmula de iteração é similar à de Gauss-Jacobi, o número de operações em uma iteração também é o mesmo. Assim, temos um total de  $2n^2 - n$  operações em uma iteração.

Para escrevermos a fórmula de iteração no formato matricial, devemos dividir a matriz  $A$  do sistema em 3 partes:

- A matriz  $L$ , triangular inferior, com diagonal nula, que corresponde aos termos que serão aplicados ao vetor  $x^{(k+1)}$ ;
- A matriz  $D$ , matriz diagonal, com os termos da diagonal principal diferentes de 0, e os demais termos iguais a 0;
- A matriz  $R$ , triangular superior, com diagonal nula, que corresponde aos termos que serão aplicados ao vetor  $x^{(k)}$ .

Assim,  $A = L + D + R$ .

**Exemplo 3.1.** Tomando a matriz

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 3 & 4 \\ 4 & 7 & -1 \end{bmatrix},$$

temos que

$$\begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 4 & 7 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 2 & -3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 3 & 4 \\ 4 & 7 & -1 \end{bmatrix}.$$

Assim,

$$L = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 4 & 7 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad e \quad R = \begin{bmatrix} 0 & 2 & -3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}.$$

Substituindo o resultado anterior, temos:

$$\begin{aligned} Ax = b &\Rightarrow (L + D + R)x = b \\ &\Rightarrow Dx = b - Lx - Rx \\ &\Rightarrow x = D^{-1}b - D^{-1}Lx - D^{-1}Rx. \end{aligned}$$

Como  $L$  é aplicado a  $x^{(k+1)}$  e  $R$  a  $x^{(k)}$ , temos

$$\begin{aligned} x^{(k+1)} = D^{-1}b - D^{-1}Lx^{(k+1)} - D^{-1}Rx^{(k)} &\Rightarrow (I + D^{-1}L)x^{(k+1)} = -D^{-1}Rx^{(k)} + D^{-1}b \\ &\Rightarrow x^{(k+1)} = -[(I + D^{-1}L)^{-1}D^{-1}R]x^{(k)} + [(I + D^{-1}L)^{-1}D^{-1}b]. \end{aligned}$$

Assim,  $C = -(I + D^{-1}L)^{-1}D^{-1}R$  e  $g = (I + D^{-1}L)^{-1}D^{-1}b$ .

Para a garantia de convergência, estudemos um critério distinto novamente:

**Teorema 3.3. (Critério de Sassenfeld)** *Considerando os termos da matriz  $A$ , defina*

$\beta_1 = \frac{\sum_{i=2}^n |a_{1i}|}{|a_{11}|}$  e  $\beta_j = \frac{\sum_{i=1}^{j-1} |a_{ji}|\beta_i + \sum_{i=j+1}^n |a_{ji}|}{|a_{jj}|}$ . Tomando  $\beta = \max_{1 \leq j \leq n} \{\beta_j\}$ , temos que, se  $\beta < 1$ , a sequência  $\{x^{(k)}\}$  gerada pelo método de Gauss-Seidel converge para a solução do sistema, independente da aproximação inicial  $x^{(0)}$  escolhida. Além disso, quanto menor o valor de  $\beta$ , mais rápida é a convergência.

**Demonstração 3.1.** *Seja o vetor  $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$  a solução do sistema linear, e  $e_i^{(k)} = x_i^{(k)} - \bar{x}_i, i \in \{1, \dots, n\}$ , ou seja, a diferença do valor de  $x_i$  na  $k$ -ésima iteração com o valor real de  $x_i$ .*

*Assim, a sequência de vetores  $x^{(k)}$  converge para a solução se, e somente se,  $e_i^{(k)} \rightarrow 0$ .*

*Sabemos que*

$$x_i^{(k+1)} = \frac{b_i - a_{i1}x_1^{(k+1)} - \dots - a_{i,i-1}x_{i-1}^{(k+1)} - a_{i,i+1}x_{i+1}^{(k)} - \dots - a_{in}x_n^{(k)}}{a_{ii}}$$

*e, como o vetor  $\bar{x}$  é solução,  $\sum_{j=1}^n a_{ij}\bar{x}_j = b_i$*

$$\Rightarrow \bar{x}_i = \frac{b_i - a_{i1}\bar{x}_1 - \dots - a_{i,i-1}\bar{x}_{i-1} - a_{i,i+1}\bar{x}_{i+1} - \dots - a_{in}\bar{x}_n}{a_{ii}}.$$

*Assim, temos que  $e_i^{(k+1)} = -\frac{\sum_{j=1}^{i-1} a_{ij}[x_j^{(k+1)} - \bar{x}_j] + \sum_{j=i+1}^n a_{ij}[x_j^{(k)} - \bar{x}_j]}{a_{ii}}$*

$$= -\frac{\sum_{j=1}^{i-1} a_{ij}e_j^{(k+1)} + \sum_{j=i+1}^n a_{ij}e_j^{(k)}}{a_{ii}}, i \in \{1, \dots, n\}.$$

Agora, definimos  $E^{(k)} = \max_{1 \leq i \leq n} \{|e_i^{(k)}|\}$ ,  $\beta_1 = \frac{\sum_{i=2}^n |a_{1i}|}{|a_{11}|}$  e  $\beta_j = \frac{\sum_{i=1}^{j-1} |a_{ji}| \beta_i + \sum_{k=j+1}^n |a_{jk}|}{|a_{jj}|}$ ,  
 com  $\beta = \max_{1 \leq i \leq n} \beta_i$ .

Dizer que a sequência converge significa dizer que  $E^{(k)} \rightarrow 0$  quando  $k \rightarrow \infty$ .

Neste sentido, queremos concluir que  $E^{(k+1)} \leq \beta E^{(k)}$ . Para isso, façamos uma indução para provar que  $|e_i^{(k+1)}| \leq \beta_i \max_{1 \leq j \leq n} \{|e_j^{(k)}|\}$ ,  $i \in \{1, \dots, n\}$ :

Para  $i = 1$ , temos que

$$\begin{aligned} e_1^{(k+1)} &= -\frac{a_{12}e_2^{(k)} + a_{13}e_3^{(k)} + \dots + a_{1n}e_n^{(k)}}{a_{11}} \\ \Rightarrow |e_1^{(k+1)}| &\leq \frac{|a_{12}||e_2^{(k)}| + |a_{13}||e_3^{(k)}| + \dots + |a_{1n}||e_n^{(k)}|}{|a_{11}|} \\ &\leq \frac{|a_{12}| + |a_{13}| + \dots + |a_{1n}|}{|a_{11}|} \max_{1 \leq j \leq n} \{|e_j^{(k)}|\} \\ &= \beta_1 \max_{1 \leq j \leq n} \{|e_j^{(k)}|\}. \end{aligned}$$

Assim,

$$|e_1^{(k+1)}| \leq \beta_1 \max_{1 \leq j \leq n} \{|e_j^{(k)}|\} \leq \beta \max_{1 \leq j \leq n} \{|e_j^{(k)}|\}.$$

Vamos supor, então, para algum  $i \leq n$ , que  $|e_t^{(k+1)}| \leq \beta_t \max_{1 \leq j \leq n} \{|e_j^{(k)}|\}$ ,  $t \in \{2, \dots, i-1\}$ .

Queremos então provar que  $|e_i^{(k+1)}| \leq \beta_i \max_{1 \leq j \leq n} \{|e_j^{(k)}|\}$ .

Temos que

$$\begin{aligned} e_i^{(k+1)} &= -\frac{\sum_{j=1}^{i-1} a_{ij}e_j^{(k+1)} + \sum_{j=i+1}^n a_{ij}e_j^{(k)}}{a_{ii}} \\ \Rightarrow |e_i^{(k+1)}| &\leq \frac{|a_{i1}||e_1^{(k+1)}| + \dots + |a_{i,i-1}||e_{i-1}^{(k+1)}| + |a_{i,i+1}||e_{i+1}^{(k)}| + \dots + |a_{in}||e_n^{(k)}|}{|a_{ii}|} \\ &\leq \frac{|a_{i1}|\beta_1 + \dots + |a_{i,i-1}|\beta_{i-1} + |a_{i,i+1}| + \dots + |a_{in}|}{|a_{ii}|} \max_{1 \leq j \leq n} \{|e_j^{(k)}|\} \\ &= \beta_i \max_{1 \leq j \leq n} \{|e_j^{(k)}|\}. \end{aligned}$$

Logo,

$$|e_i^{(k+1)}| \leq \beta_i \max_{1 \leq j \leq n} \{|e_j^{(k)}|\} \leq \beta \max_{1 \leq j \leq n} \{|e_j^{(k)}|\}, i \in \{1, \dots, n\},$$

e como  $E^{(k)} = \max_{1 \leq i \leq n} \{|e_i^{(k)}|\}$  e  $E^{(k+1)} = \max_{1 \leq i \leq n} \{|e_i^{(k+1)}|\}$ , concluímos que

$$E^{(k+1)} \leq \beta E^{(k)}.$$

Sendo assim, basta que  $\beta < 1$  para que tenhamos a sequência  $E^{(k)}$  decrescente. Sendo ela composta de valores estritamente positivos, temos que, para valores de  $k$  suficientemente grandes,  $E^{(k)} \rightarrow 0$ , e assim, a sequência converge.

Além disso, como  $\beta$  é fator multiplicador para o próximo termo, quanto menor  $\beta$ , mais rápido  $E^{(k)} \rightarrow 0$ , e assim mais rápido a sequência converge para a solução.

O **critério das linhas** também pode ser utilizado para testar a convergência no método de Gauss-Seidel. No entanto, temos que, se um sistema satisfaz o critério das linhas, ele também satisfaz o de Sassenfeld.

De fato, supondo que o critério de linhas seja satisfeito, temos que

$$\alpha = \max_{1 \leq i \leq n} \{\alpha_i\} < 1, \text{ onde } \alpha_i = \frac{\sum_{j=1; j \neq i}^n |a_{ij}|}{|a_{ii}|}.$$

$$\text{Além disso, } \beta_1 = \frac{\sum_{i=2}^n |a_{1i}|}{|a_{11}|} = \alpha_1 < 1.$$

Supondo que, para  $i \in \{1, \dots, k-1\}$ , temos que  $\beta_i \leq \alpha_i < 1$ , concluímos que

$$\begin{aligned} \beta_k &= \frac{\sum_{i=1}^{k-1} |a_{ki}| \beta_i + \sum_{i=k+1}^n |a_{ki}|}{|a_{kk}|} \\ &= \frac{|a_{k1}| \beta_1 + \dots + |a_{k,k-1}| \beta_{k-1} + |a_{k,k+1}| + \dots + |a_{kn}|}{|a_{kk}|} \\ &< \frac{|a_{k1}| + \dots + |a_{k,k-1}| + |a_{k,k+1}| + \dots + |a_{kn}|}{|a_{kk}|} \\ &= \alpha_k. \end{aligned}$$

Logo, por indução,  $\beta_i \leq \alpha_i, i \in \{1, \dots, n\}$ . Mas  $\alpha_i < 1$ , pois o critério das linhas é satisfeito. Assim,  $\beta_i < 1, i \in \{1, \dots, n\}$ , e o critério de Sassenfeld também é satisfeito.

A recíproca não é verdadeira. Se um sistema linear satisfaz Sassenfeld, não necessariamente ele satisfará o critério das linhas. Vejamos:

**Exemplo 3.2.** *Tomemos o sistema linear*

$$\begin{cases} 2x_1 - x_2 = 4 \\ x_1 + x_2 = 5 \end{cases}.$$

*Temos que o critério de linhas não é satisfeito na linha 2, pois  $\alpha_2 = 1$ , mesmo se trocarmos as colunas.*

No entanto,  $\beta_1 = \frac{|a_{12}|}{|a_{11}|} = \frac{1}{2} = 0,5 < 1$  e  $\beta_2 = \frac{|a_{21}\beta_1|}{|a_{22}|} = \frac{0,5}{1} = 0,5 < 1$ . Logo, o critério de Sassenfeld é satisfeito.

**Observação 3.7.** Novamente, os critérios de convergência são apenas condições suficientes para que a sequência convirja para a solução, mas não necessárias. Um sistema que não satisfaz ambos os critérios apresentados pode ter a sequência de vetores gerada por Gauss-Seidel convergente para a solução do problema.

Resolvamos agora o mesmo problema proposto, através do método de Gauss-Seidel.

Pelo **critério de Sassenfeld**, vemos que a sequência de vetores convergirá para a solução do sistema. De fato:

$$\beta_1 = \frac{|-1|+|0|+|0|+|-1|+|0|+|0|+|0|}{4} = \frac{1+1}{4} = \frac{2}{4} = 0,5;$$

$$\beta_2 = \frac{|-1|\times 0,5+|-1|+|0|+|0|+|-1|+|0|+|0|}{4} = \frac{0,5+1+1}{4} = \frac{2,5}{4} = 0,625;$$

$$\beta_3 = \frac{|0|\times 0,5+|-1|\times 0,625+|-1|+|0|+|0|+|-1|+|0|}{4} = \frac{0,625+1+1}{4} = \frac{2,625}{4} = 0,65625;$$

Analogamente, calculamos os demais:

$$\beta_4 = \frac{|-1|\times 0,65625+|-1|}{4} = \frac{0,65625+1}{4} = \frac{1,65625}{4} = 0,4140625;$$

$$\beta_5 = \frac{|-1|\times 0,5+|-1|}{4} = \frac{1,5}{4} = 0,375;$$

$$\beta_6 = \frac{|-1|\times 0,625+|-1|\times 0,375+|-1|}{4} = \frac{2}{4} = 0,5;$$

$$\beta_7 = \frac{|-1|\times 0,65625+|-1|\times 0,5+|-1|}{4} = \frac{2,15625}{4} = 0,5390625;$$

$$\beta_8 = \frac{|-1|\times 0,4140625+|-1|\times 0,5390625}{4} = \frac{0,953125}{4} = 0,23828125$$

$$\Rightarrow \beta = \max_{1 \leq j \leq 8} \{\beta_j\} = 0,65625 < 1.$$

Análogo à resolução pelo primeiro método, vamos definir

$$x^{(k)} = \{P_1^{(k)}; P_2^{(k)}; P_3^{(k)}; P_4^{(k)}; P_5^{(k)}; P_6^{(k)}; P_7^{(k)}; P_8^{(k)}\}$$

e escolher  $x^{(0)} = \{0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5\}$ .

A fórmula de iteração é dada por

$$\left\{ \begin{array}{l} P_1^{(k+1)} = \frac{P_2^{(k)} + P_5^{(k)}}{4} \\ P_2^{(k+1)} = \frac{P_1^{(k+1)} + P_3^{(k)} + P_6^{(k)}}{4} \\ P_3^{(k+1)} = \frac{P_2^{(k+1)} + P_4^{(k)} + P_7^{(k)}}{4} \\ P_4^{(k+1)} = \frac{P_3^{(k+1)} + P_8^{(k)}}{4} \\ P_5^{(k+1)} = \frac{P_1^{(k+1)} + P_6^{(k)} + 1}{4} \\ P_6^{(k+1)} = \frac{P_2^{(k+1)} + P_5^{(k+1)} + P_7^{(k)} + 1}{4} \\ P_7^{(k+1)} = \frac{P_3^{(k+1)} + P_6^{(k+1)} + P_8^{(k)} + 1}{4} \\ P_8^{(k+1)} = \frac{P_4^{(k+1)} + P_7^{(k+1)} + 1}{4} \end{array} \right. .$$

1ª Iteração:

$$P_1^{(1)} = \frac{P_2^{(0)} + P_5^{(0)}}{4} = \frac{0,5 + 0,5}{4} = \frac{1}{4} = 0,25;$$

$$P_2^{(1)} = \frac{P_1^{(1)} + P_3^{(0)} + P_6^{(0)}}{4} = \frac{0,25 + 0,5 + 0,5}{4} = \frac{1,25}{4} = 0,3125;$$

$$P_3^{(1)} = \frac{P_2^{(1)} + P_4^{(0)} + P_7^{(0)}}{4} = \frac{0,3125 + 0,5 + 0,5}{4} = \frac{1,3125}{4} = 0,328125;$$

$$P_4^{(1)} = \frac{P_3^{(1)} + P_8^{(0)}}{4} = \frac{0,328125 + 0,5}{4} = \frac{0,828125}{4} = 0,20703125;$$

$$P_5^{(1)} = \frac{P_1^{(1)} + P_6^{(0)} + 1}{4} = \frac{0,25 + 0,5 + 1}{4} = \frac{1,75}{4} = 0,4375;$$

$$P_6^{(1)} = \frac{P_2^{(1)} + P_5^{(1)} + P_7^{(0)} + 1}{4} = \frac{0,3125 + 0,4375 + 0,5 + 1}{4} = \frac{2,25}{4} = 0,5625;$$

$$P_7^{(1)} = \frac{P_3^{(1)} + P_6^{(1)} + P_8^{(0)} + 1}{4} = \frac{0,328125 + 0,5625 + 0,5 + 1}{4} = \frac{2,390625}{4} = 0,59765625;$$

$$P_8^{(1)} = \frac{P_4^{(1)} + P_7^{(1)} + 1}{4} = \frac{0,20703125 + 0,59765625 + 1}{4} = \frac{1,8046875}{4} = 0,451171875$$

$$\Rightarrow x^{(1)} = \{0,25; 0,3125; 0,328125; 0,20703125; 0,4375; 0,5625; 0,59765625; 0,451171875\}.$$

O cálculo do erro relativo é análogo ao calculado durante o método de Gauss-Jacobi. Assim,

$$d_r^{(1)} = 0,490196078.$$

Prosseguindo da mesma forma nas demais iterações, encontramos:

2ª Iteração:

$$x^{(2)} = \{0,1875; 0,26953125; 0,268554688; 0,179931641; 0,4375; 0,576171875; 0,573974609; 0,438476563\};$$

$$d_r^{(2)} = 0,108474576.$$

3ª Iteração:

$$x^{(3)} = \{0,176757813; 0,255371094; 0,252319336; 0,172698975; 0,438232422; 0,566894531; 0,564422607; 0,434280396\};$$

$$d_r^{(3)} = 0,028639104.$$

4ª Iteração:

$$x^{(4)} = \{0,173400879; 0,248153687; 0,246318817; 0,170149803; 0,435073853; 0,561912537; 0,560627937; 0,432694435\};$$

$$d_r^{(4)} = 0,012844361.$$

Novamente, analisaremos a tabela de valores das iterações e erros:

Número da Iteração	P1	P2	P3	P4	P5	P6	P7	P8	Erro Relativo
7	0,168883607	0,242598537	0,242429235	0,168552209	0,43188275	0,558217853	0,558106672	0,43166472	0,001171974
8	0,168620322	0,242316853	0,242243933	0,168477163	0,431709544	0,558033267	0,55798548	0,431615661	0,000504781
9	0,168506599	0,24219595	0,242164648	0,168445077	0,431634967	0,557954099	0,557933602	0,43159467	0,000216689

Vemos que, a partir da 8ª iteração, o erro relativo torna-se menor que 0,001.

Número da Iteração	P1	P2	P3	P4	P5	P6	P7	P8	Erro Relativo
24	0,168421053	0,242105263	0,242105263	0,168421053	0,431578948	0,557894737	0,557894737	0,431578947	6,51421E-10
25	0,168421053	0,242105263	0,242105263	0,168421053	0,431578947	0,557894737	0,557894737	0,431578947	2,79057E-10
26	0,168421053	0,242105263	0,242105263	0,168421053	0,431578947	0,557894737	0,557894737	0,431578947	1,19543E-10

Da 25ª iteração em diante, o computador fixa a solução em 9 casas decimais, para cada incógnita. Logo, o erro é menor do que  $10^{-9}$ .

Número da Iteração	P1	P2	P3	P4	P5	P6	P7	P8	Erro Relativo
43	0,168421053	0,242105263	0,242105263	0,168421053	0,431578947	0,557894737	0,557894737	0,431578947	4,97506E-17
44	0,168421053	0,242105263	0,242105263	0,168421053	0,431578947	0,557894737	0,557894737	0,431578947	0
45	0,168421053	0,242105263	0,242105263	0,168421053	0,431578947	0,557894737	0,557894737	0,431578947	0

A partir da 44ª, o erro já é dado como 0.

A solução, assim, é dada por

$$x = \{0,168421053; 0,242105263; 0,242105263; 0,168421053; 0,431578947; 0,557894737; 0,557894737; 0,431578947\}.$$

### 3.3 Cálculo de Iterações através de uma Planilha Eletrônica

Veremos agora como foram feitos os cálculos das tabelas de iterações e erros do sistema linear anterior, em ambos os métodos iterativos. Os cálculos foram feitos através de uma planilha eletrônica.

#	A	B	C	D	E	F	G	H	I
1		P1	P2	P3	P4	P5	P6	P7	P8
2	0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5
3	1	0,25	0,375	0,375	0,25	0,5	0,625	0,625	0,5
4	2	0,21875	0,3125	0,3125	0,21875	0,46875	0,625	0,625	0,46875
5	3	0,1953125	0,2890625	0,2890625	0,1953125	0,4609375	0,6015625	0,6015625	0,4609375
6	4	0,1875	0,271484375	0,271484375	0,1875	0,44921875	0,587890625	0,587890625	0,44921875
7	5	0,180175781	0,26171875	0,26171875	0,180175781	0,443847656	0,577148438	0,577148438	0,443847656
8	6	0,176391602	0,254760742	0,254760742	0,176391602	0,439331055	0,570678711	0,570678711	0,439331055
9	7	0,173522949	0,250457764	0,250457764	0,173522949	0,436767578	0,566192627	0,566192627	0,436767578
10	8	0,171806335	0,247543335	0,247543335	0,171806335	0,434928894	0,563354492	0,563354492	0,434928894
11	9	0,170618057	0,245676041	0,245676041	0,170618057	0,433790207	0,56145668	0,56145668	0,433790207
12	10	0,169866562	0,244437695	0,244437695	0,169866562	0,433018684	0,560230732	0,560230732	0,433018684
13	11	0,169364095	0,243633747	0,243633747	0,169364095	0,432524323	0,559421778	0,559421778	0,432524323
14	12	0,169039518	0,243104905	0,243104905	0,169039518	0,432196468	0,558894962	0,558894962	0,432196468
15	13	0,168825343	0,242759846	0,242759846	0,168825343	0,43198362	0,558549084	0,558549084	0,43198362
16	14	0,168685867	0,242533568	0,242533568	0,168685867	0,431843607	0,558323137	0,558323137	0,431843607
17	15	0,168594294	0,242385643	0,242385643	0,168594294	0,431752251	0,558175078	0,558175078	0,431752251
18	16	0,168534474	0,242288754	0,242288754	0,168534474	0,431692343	0,558078243	0,558078243	0,431692343
19	17	0,168495274	0,242225368	0,242225368	0,168495274	0,431653179	0,558014835	0,558014835	0,431653179
20	18	0,168469637	0,242183869	0,242183869	0,168469637	0,431627527	0,557973345	0,557973345	0,431627527

A figura acima mostra os cálculos das 18 primeiras iterações durante o método de Gauss-Jacobi. A linha 1 indica as incógnitas a serem calculadas, e a coluna A indica a iteração. A aproximação inicial é dada na iteração 0.

Para calcular uma das incógnitas em uma determinada iteração, digitamos a fórmula na sua respectiva célula. Por exemplo, para calcular a variável  $P_6$  na 1ª iteração, devemos digitar sua fórmula na célula G3.

Pelo método de Gauss-Jacobi, a fórmula de  $P_6$  é dada por

$$P_6^{(k+1)} = \frac{P_2^{(k)} + P_5^{(k)} + P_7^{(k)} + 1}{4},$$

Logo, para calcularmos  $P_6^{(1)}$ , temos que

$$P_6^{(1)} = \frac{P_2^{(0)} + P_5^{(0)} + P_7^{(0)} + 1}{4}.$$

Utilizaremos na fórmula de  $P_6^{(1)}$ , portanto, as células de  $P_2^{(0)}$ ,  $P_5^{(0)}$  e  $P_7^{(0)}$  respectivamente, ou seja, C2, F2 e H2. Desta forma, teremos

$$G3 = \frac{1 + C2 + F2 + H2}{4}.$$

#	A	B	C	D	E	F	G	H	I
1		P1	P2	P3	P4	P5	P6	P7	P8
2	0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5
3	1	0,25	0,375	0,375	0,25		=[(1+C2+F2+H2)/4]	0,625	0,5
4	2	0,21875	0,3125	0,3125	0,21875	0,46875	0,625	0,625	0,46875

As fórmulas podem ser digitadas apenas na linha 3 (referente à 1ª iteração), para cada uma das incógnitas, e apenas copiadas para as linhas abaixo (demais iterações).

O processo para o cálculo através de Gauss-Seidel é o mesmo, no entanto, enquanto em Gauss-Jacobi a fórmula de uma determinada célula tem apenas valores da linha anterior, ou seja, da iteração anterior, em Gauss-Seidel tomamos alguns valores da mesma linha (mesma iteração) e outros da anterior.

**Exemplo 3.3. (Cálculo de  $P_6^{(1)}$  em Gauss-Seidel)**

$$P_6^{(1)} = \frac{P_2^{(1)} + P_5^{(1)} + P_7^{(0)} + 1}{4} \Rightarrow G3 = \frac{1 + C3 + F3 + H2}{4}$$

	A	B	C	D	E	F	G	H	I
1		P1	P2	P3	P4	P5	P6	P7	P8
2	0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5
3	1	0,25	0,3125	0,328125	0,20703125		= (1+C3+F3+H2)/4		0,451171875
4	2	0,1875	0,26953125	0,268554688	0,179931641	0,4375	0,576171875	0,573974609	0,438476563

Para o cálculo do erro relativo, o processo é idêntico para ambos os métodos. Novamente, tomaremos como exemplo os cálculos dos erros relativos das 18 primeiras iterações, obtidas através do método de Gauss-Jacobi:

	K	L	M	N	O	P	Q	R	S	T
1		d1	d2	d3	d4	d5	d6	d7	d8	
2										
3	1	0,25	0,125	0,125	0,25	0	0,125	0,125	0	0,4
4	2	0,03125	0,0625	0,0625	0,03125	0,03125	0	0	0,03125	0,1
5	3	0,02344	0,02344	0,02344	0,02344	0,00781	0,02344	0,02344	0,00781	0,038961039
6	4	0,00781	0,01758	0,01758	0,00781	0,01172	0,01367	0,01367	0,01172	0,029900332
7	5	0,00732	0,00977	0,00977	0,00732	0,00537	0,01074	0,01074	0,00537	0,018612521
8	6	0,00378	0,00696	0,00696	0,00378	0,00452	0,00647	0,00647	0,00452	0,012192513
9	7	0,00287	0,0043	0,0043	0,00287	0,00256	0,00449	0,00449	0,00256	0,007923247
10	8	0,00172	0,00291	0,00291	0,00172	0,00184	0,00284	0,00284	0,00184	0,005173348
11	9	0,00119	0,00187	0,00187	0,00119	0,00114	0,0019	0,0019	0,00114	0,003380157
12	10	0,00075	0,00124	0,00124	0,00075	0,00077	0,00123	0,00123	0,00077	0,002210422
13	11	0,0005	0,0008	0,0008	0,0005	0,00049	0,00081	0,00081	0,00049	0,001446054
14	12	0,00032	0,00053	0,00053	0,00032	0,00033	0,00053	0,00053	0,00033	0,000946228
15	13	0,00021	0,00035	0,00035	0,00021	0,00021	0,00035	0,00035	0,00021	0,000619244
16	14	0,00014	0,00023	0,00023	0,00014	0,00014	0,00023	0,00023	0,00014	0,000405281
17	15	9,2E-05	0,00015	0,00015	9,2E-05	9,1E-05	0,00015	0,00015	9,1E-05	0,000265256
18	16	6E-05	9,7E-05	9,7E-05	6E-05	6E-05	9,7E-05	9,7E-05	6E-05	0,000173612
19	17	3,9E-05	6,3E-05	6,3E-05	3,9E-05	3,9E-05	6,3E-05	6,3E-05	3,9E-05	0,000113632
20	18	2,6E-05	4,1E-05	4,1E-05	2,6E-05	2,6E-05	4,1E-05	4,1E-05	2,6E-05	7,43734E-05

Na linha 1, indicamos a incógnita a qual estamos analisando, sendo  $d_i$  relativa à incógnita  $P_i, i \in \{1, \dots, 8\}$ , enquanto a coluna K indica a iteração. Na coluna T, é dado o erro relativo de cada uma das iterações.

Para calcular  $d_i^{(k)}$ , ou seja a distância entre  $P_i^{(k)}$  e  $P_i^{(k-1)}$ , escrevemos a fórmula na sua respectiva célula. Por exemplo, para calcular  $d_3^{(3)}$ , utilizamos a célula N5:

$$N5=ABS(D5-D4),$$

onde D5 e D4 representam  $P_3^{(3)}$  e  $P_3^{(2)}$ , respectivamente. A fórmula =ABS(·) toma o valor absoluto da diferença entre os valores de D5 e D4.

O erro relativo é dado pela divisão do maior valor entre os  $d_i$ 's de uma iteração pelo maior valor absoluto entre as incógnitas, na mesma iteração. Desta forma, para calcular o erro relativo na 3ª iteração, utilizando a célula T5, fazemos

$$T5 = \frac{\text{MÁXIMO}(L5:S5)}{\text{MÁXIMO}(B5:I5)}.$$

A notação L5:S5 indica que estão selecionadas todas as células, de L5 a S5, e a fórmula =MÁXIMO(·) fornece o maior valor entre as células selecionadas.

Novamente, pode-se escrever as fórmulas somente na linha 3 (1ª iteração) e copiá-la para as linhas abaixo (demais iterações).

Desta forma, podemos facilmente calcular as iterações e os respectivos erros até onde desejado. Convidamos o leitor a repetir o processo para as demais resoluções de sistemas propostas neste trabalho.

### 3.4 Comparando os Métodos Iterativos

Na análise dos métodos iterativos, não consideramos a contagem das operações, uma vez que esse cálculo depende da escolha da aproximação inicial  $x^{(0)}$  e do erro definido. Tomando os casos em que a sequência de aproximações converge, a escolha não influi nessa convergência, mas pode influir na rapidez em que ela ocorre. Assim, a título de comparação, utilizamos o mesmo vetor inicial em ambos os casos.

Podemos perceber que o método de Gauss-Seidel converge para a solução de forma mais rápida que o método de Gauss-Jacobi. Isso acontece pelo fato de que, ainda dentro de uma certa iteração, as incógnitas já calculadas são utilizadas para determinarmos as demais, enquanto que em Gauss-Jacobi, o cálculo depende apenas dos valores encontrados na iteração anterior. Dessa forma, sem considerar o custo computacional, seria mais vantajoso utilizarmos Gauss-Seidel.

No entanto, de [3], temos que, "usando a ideia de paralelismo, esta vantagem do método de Gauss-Seidel sobre o de Jacobi se inverte, pois a existência de dependência de dados no algoritmo de Gauss-Seidel faz com que o método de Jacobi seja mais rápido uma vez que nele não há dependência de dados. Em outras palavras, no método de Gauss-Seidel o cálculo da variável  $x_n^{(k)}$ , na  $k$ -ésima iteração, depende do cálculo de até todas  $n - 1$  variáveis nesta mesma iteração, pois neste método, as variáveis são utilizadas na mesma iteração logo que elas estejam disponíveis. No caso do método de Jacobi, o valor de uma variável  $x_n^{(k)}$ ,  $k$ -ésima iteração, só é alterada quando todas as variáveis daquela iteração tenham sido determinadas, isto é, cada variável é calculada usando o valor da iteração anterior das demais variáveis".

Em outras palavras, a recuperação de dados previamente obtidos armazenados na memória do computador, em Gauss-Seidel, é feita durante a iteração, enquanto que em Gauss-Jacobi essa recuperação é feita no intervalo entre duas iterações. Isso faz com que o método de Gauss-Jacobi possa ter uma vantagem no custo computacional em relação ao método de Gauss-Seidel. O armazenamento e a recuperação de valores obtidos durante as operações também tem um custo, não sendo consideradas, entretanto, neste trabalho.

Mais sobre a ideia de paralelismo de algoritmos pode ser encontrado em [3].

Considerando o custo computacional dependente apenas das operações de adição e subtração realizadas durante o processo, definiremos o método iterativo de Gauss-Seidel como sendo o mais vantajoso.

### 3.5 Limitação dos Métodos Iterativos

Vejam os exemplos de aplicação dos sistemas lineares no campo da Mecânica Estática e Engenharia:

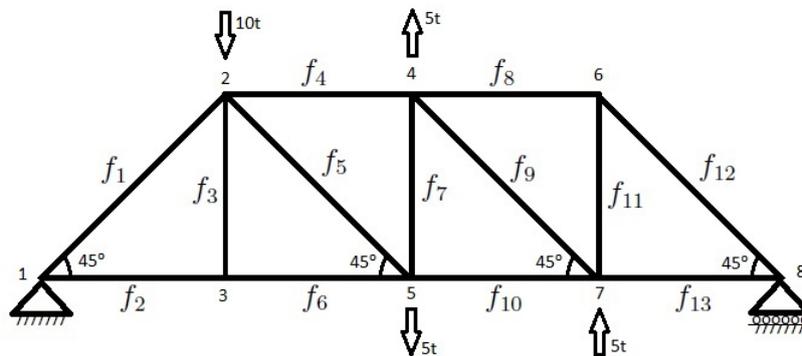
A **treliça** é uma estrutura de elementos ligados entre si pelas extremidades. Os elementos geralmente utilizados nas construções são de madeira ou barras de metal e em geral são unidos uns aos outros por meio de uma placa de reforço no qual eles são aparafusados ou soldados. No caso da treliça plana, objeto do nosso estudo, ela se distribui em um único plano e geralmente é utilizada na sustentação de telhados e pontes.

As forças aplicadas em uma treliça, internas ou externas, incidem nos nós (junção de dois elementos ou mais) e, para que ela se mantenha em equilíbrio estático, essas forças se anulam.

Definindo  $j$  como sendo o número de nós e  $m$  o número de elementos de uma treliça, quando obtemos a relação  $m = 2j - 3$ , a treliça diz-se *isoestática*. O estado isoestático de uma treliça é o "estado limite" de equilíbrio. De fato, caso um elemento da treliça seja retirado ou falhe, toda a estrutura falha. É possível que uma treliça tenha mais elementos do que o necessário.

Mais informações sobre treliças podem ser encontradas em [8].

Analisemos então a seguinte treliça:



O desenho nos mostra que, na junção 1, a treliça está fixada, e na junção 8, a treliça tem livre movimentação horizontal, porém não vertical. Além disso, existem forças externas nas junções 2, 4, 5 e 7, cujas respectivas setas indicam a direção para a qual a força atua.

Queremos determinar as forças  $f_i, i \in \{1, \dots, 13\}$  que atuam nos elementos dessa treliça. A cada junção, são feitas as análises das forças horizontais ( $F_x$ ) e verticais ( $F_y$ ), separadamente.

Considerando  $\cos 45^\circ = \sin 45^\circ = 0,70710678 = \alpha$ , temos:

**a. Na junção 1:** Como a treliça não tem movimentação nem vertical como horizontal,  $F_x$  e  $F_y$  se anulam com as forças normais. Como as normais não são de nosso interesse, não há o que determinar aqui.

**b. Na junção 2:**

$$\begin{cases} F_y : \sin 45^\circ f_1 + f_3 + \sin 45^\circ f_5 + 10t = 0 \Rightarrow 0,70710678f_1 + f_3 + 0,70710678f_5 = -10 \\ F_x : \cos 45^\circ f_1 = f_4 + \cos 45^\circ f_5 \Rightarrow 0,70710678f_1 - f_4 - 0,70710678f_5 = 0 \end{cases}$$

**c. Na junção 3:**

$$\begin{cases} F_x : f_2 = f_6 \Rightarrow f_2 - f_6 = 0 \\ F_y : f_3 = 0 \end{cases}$$

**d. Na junção 4:**

$$\begin{cases} F_x : f_4 = f_8 + \cos 45^\circ f_9 \Rightarrow f_4 - f_8 - 0,70710678f_9 = 0 \\ F_y : f_7 + \sin 45^\circ f_9 = 5t \Rightarrow f_7 + 0,70710678f_9 = 5 \end{cases}$$

**e. Na junção 5:**

$$\begin{cases} F_x : \cos 45^\circ f_5 + f_6 + \cos 45^\circ f_{10} = 0 \Rightarrow 0,70710678f_5 + f_6 + 0,70710678f_{10} = 0 \\ F_y : \sin 45^\circ f_5 + f_7 = 5t \Rightarrow 0,70710678f_5 + f_7 = 5 \end{cases}$$

**f. Na junção 6:**

$$\begin{cases} F_x : f_8 = \cos 45^\circ f_{12} \Rightarrow f_8 - 0,70710678f_{12} = 0 \\ F_y : f_{11} + \sin 45^\circ f_{12} = 0 \Rightarrow f_{11} + 0,70710678f_{12} = 0 \end{cases}$$

**g. Na junção 7:**

$$\begin{cases} F_x : \cos 45^\circ f_9 + f_{10} = f_{13} \Rightarrow 0,70710678f_9 + f_{10} - f_{13} = 0 \\ F_y : \sin 45^\circ f_9 + f_{11} + 5t = 0 \Rightarrow 0,70710678f_9 + f_{11} = -5 \end{cases}$$

**h. Na junção 8:** Igualando as forças horizontais contrárias, temos

$$\cos 45^\circ f_{12} + f_{13} = 0 \Rightarrow 0,70710678f_{12} + f_{13} = 0.$$

As forças verticais se anulam com a normal.

Assim, considerando o sistema na forma  $A \times x = b$ , temos:

$$A = \begin{bmatrix} \alpha & 0 & 1 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha & 0 & 0 & -1 & -\alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & -\alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -\alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 1 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} -10 \\ 0 \\ 0 \\ 0 \\ 0 \\ 5 \\ 0 \\ 5 \\ 0 \\ 0 \\ 0 \\ 0 \\ -5 \\ 0 \end{bmatrix}.$$

Reordenemos as linhas da matriz  $A$  para que não existam termos nulos na diagonal principal. Assim, temos que

$$A = \begin{bmatrix} \alpha & 0 & 1 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha & 0 & 0 & -1 & -\alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & -\alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -\alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 1 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} -10 \\ 0 \\ 0 \\ 0 \\ 5 \\ 0 \\ 5 \\ 0 \\ -5 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

O critério de Sassenfeld, neste caso, não é satisfeito, mesmo modificando-se algumas linhas ou colunas. No entanto, isso não garante que a sequência diverja.

Assim, escrevemos as fórmulas recursivas para aplicação do método de Gauss-Seidel, da forma

$$\left\{ \begin{array}{l} f_1^{(k+1)} = \frac{-10 - f_3^{(k)} - \alpha f_5^{(k)}}{\alpha} \\ f_2^{(k+1)} = f_6^{(k)} \\ f_3^{(k+1)} = 0 \\ f_4^{(k+1)} = \alpha f_1^{(k+1)} - \alpha f_5^{(k)} \\ f_5^{(k+1)} = \frac{-f_7^{(k)} + 5}{\alpha} \\ f_6^{(k+1)} = -\alpha f_5^{(k+1)} + f_{10}^{(k)} \\ f_7^{(k+1)} = -\alpha f_9^{(k+1)} + 5 \\ f_8^{(k+1)} = f_4^{(k+1)} - \alpha f_9^{(k)} \\ f_9^{(k+1)} = \frac{-f_{11}^{(k)} - 5}{\alpha} \\ f_{10}^{(k+1)} = -\alpha f_9^{(k+1)} + f_{13}^{(k)} \\ f_{11}^{(k+1)} = -\alpha f_{12}^{(k)} \\ f_{12}^{(k+1)} = \frac{f_8^{(k+1)}}{\alpha} \\ f_{13}^{(k+1)} = -\alpha f_{12}^{(k+1)} \end{array} \right. .$$

Utilizando-se da aproximação inicial  $f^{(0)} = \{-9; 6; 0; -2; -5; 6; 9; 1; -5; 2; -1; 2; -1\}$ , obtemos a seguinte sequência de vetores:

1ª Iteração:

$$f^{(1)} = \{-9,142135647; 6; 0; -2,9289322; -5,656854259; 6; 8,5355339; 0,6066017; -5,656854259; 3; -1,41421356; 0,857864353; -9,6066017\};$$

$$d_r^{(1)} = 0,133809515.$$

⋮

5<sup>a</sup> Iteração:

$$f^{(5)} = \{-13,31370852; 5,97918474; 4; -5,8284271; -6,213203471; \\ 7,73654406; 8; -2,82842712; -7,556349212; 4,9497475; \\ -0,3933983; -4; 2,82842712\};$$

$$d_r^{(5)} = 0,569543652.$$

⋮

10<sup>a</sup> Iteração:

$$f^{(10)} = \{-14,38477634; 17,09188306; 9; -2,3431458; -6,887301576; \\ 11,7487373; 14,3933983; 7,05025254; -3,514718612; 5,4020203; \\ 2,91673896; 9,970562777; -7,05025254\};$$

$$d_r^{(10)} = 0,824687048.$$

⋮

15<sup>a</sup> Iteração:

$$f^{(15)} = \{-35,42640964; 3,02081526; 14; 27,100505; -16,65180379; \\ 28,15075762; 3,29646444; -28,80404064; -32,63961041; 17,471609; \\ -5,60808128; -40,73506499; 28,80404064\};$$

$$d_r^{(15)} = 1,689557733.$$

Vemos que, neste exemplo, a sequência não converge para a solução. Como vimos anteriormente, isso acontece quando temos algum autovalor  $\lambda_i$  da matriz  $C$  da fórmula de iteração do problema tal que  $|\lambda_i| \geq 1$ .

Este sistema pode então ser resolvido por eliminação Gaussiana, obtendo-se assim a solução

$$f = \{-8,83883478; 6,25; 0; -2,5; -5,30330087; \\ 6,25; 8,75; 1,25; -5,30330087; 2,5; \\ -1,25; 1,76776696; -1,25\},$$

após 1703 operações.

## 4 Custo Computacional Comparado

Agora, faremos a comparação do custo computacional entre os métodos diretos e iterativos. Para isso, será resolvido um sistema linear aplicado através do método da **Eliminação Gaussiana**, o qual constatamos ser o mais vantajoso dos métodos diretos, e do método de **Gauss-Seidel**, o mais vantajoso dos iterativos.

Além disso, como já dito anteriormente, o custo operacional dos métodos iterativos depende da aproximação inicial escolhida e do erro estabelecido. Com o intuito de minimizar os efeitos dessa variação, será fixado um  $\epsilon = 0,005$  para o cálculo do erro. A aproximação inicial dependerá do exercício proposto, onde buscaremos a média dos valores possíveis.

**Observação 4.1.** *Para alguém com mais experiência na área de aplicação do sistema linear, o "chute" inicial não será necessariamente a média dos valores possíveis. Fazendo uma análise prévia dos dados do problema, pode-se garantir uma aproximação inicial melhor, e assim, uma quantidade menor de iterações necessárias para a precisão desejada.*

Veremos mais uma aplicação dos sistemas lineares, desta vez na Física, no campo da Termodinâmica.

### 4.1 Temperatura em um Ponto de uma Placa

Consideremos uma placa quadrada cujos bordos  $AC$  e  $BD$  são mantidos à temperatura de  $20^\circ C$ , o bordo  $AB$  a  $40^\circ C$  e o bordo  $CD$  a  $10^\circ C$ , com o uso de isolantes térmicos em  $A$ ,  $B$ ,  $C$  e  $D$ . Queremos descobrir a temperatura em determinados pontos dessa placa.

A equação que rege o fluxo de calor é dada por

$$\frac{\partial T}{\partial t} = c \left( \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right),$$

onde  $T$  é a temperatura num ponto  $(x, y)$ ,  $t$  é o instante de tempo e  $c > 0$  é uma constante característica do material de que é feita a placa.

No equilíbrio térmico, a temperatura  $T$  não varia mais com o tempo, e portanto

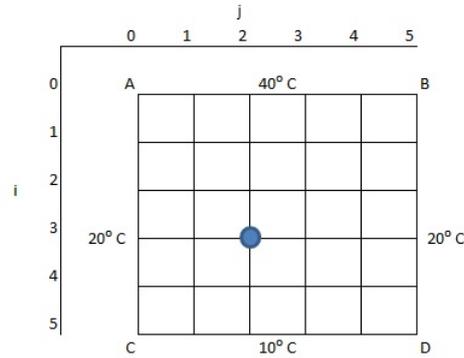
$$\frac{\partial T}{\partial t} = 0 \Rightarrow \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0.$$

Dessa forma, o problema consiste em encontrar uma função  $T(x, y)$  que satisfaça essa equação.

Para resolver esse problema de uma forma mais simples, podemos substituir a placa por uma aproximação discreta, que consiste em uma malha quadriculada. Nesta malha, a temperatura em um ponto pode ser calculada (de forma aproximada) através da média aritmética dos quatro pontos adjacentes.

Maiores informações sobre o assunto e a demonstração desse último resultado podem ser encontrados em [7].

Vemos, com a ajuda de uma escala, a placa da seguinte forma:



Calcularemos a temperatura nos pontos  $T_{ij}$  da malha quadriculada, com  $i \in \{0, \dots, 5\}$  e  $j \in \{0, \dots, 5\}$ . Como exemplo, foi destacado o ponto  $T_{32}$  na figura.

O vetor das variáveis será da forma

$$\vec{x} = \{T_{11}, T_{12}, T_{13}, T_{14}, T_{21}, T_{22}, T_{23}, T_{24}, T_{31}, T_{32}, T_{33}, T_{34}, T_{41}, T_{42}, T_{43}, T_{44}\}.$$

O vetor solução  $\bar{x}$  seguirá essa mesma ordenação.

A temperatura nos pontos  $T_{0j}$ ,  $j \in \{1, \dots, 4\}$ , é igual a  $40^\circ C$ , pois estes pontos pertencem ao bordo  $AB$ . Analogamente, nos pontos  $T_{i0}$ ,  $i \in \{1, \dots, 4\}$ , que pertencem ao bordo  $AC$ , a temperatura é igual a  $20^\circ C$ . No mesmo pensamento,  $T_{i5} = 20^\circ C$ ,  $i \in \{1, \dots, 4\}$  e  $T_{5j} = 10^\circ C$ ,  $j \in \{1, \dots, 4\}$ .

Não nos interessam os pontos  $A = (0, 0)$ ,  $B = (0, 5)$ ,  $C = (5, 0)$  e  $D = (5, 5)$ , pois estão isolados termicamente.

Tomemos então o ponto  $T_{11}$ . Seus pontos adjacentes são  $T_{01} = 40$ ,  $T_{10} = 20$ ,  $T_{12}$  e  $T_{21}$ . Assim, pela fórmula da média aritmética, temos

$$T_{11} = \frac{T_{01} + T_{10} + T_{12} + T_{21}}{4} = \frac{40 + 20 + T_{12} + T_{21}}{4} = \frac{60 + T_{12} + T_{21}}{4}$$

$$\Rightarrow 4T_{11} - T_{12} - T_{21} = 60.$$

Repetindo o processo para os demais pontos, temos

$$T_{12} = \frac{40 + T_{11} + T_{13} + T_{22}}{4} \Rightarrow -T_{11} + 4T_{12} - T_{13} - T_{22} = 40;$$

$$T_{13} = \frac{40 + T_{12} + T_{14} + T_{23}}{4} \Rightarrow -T_{12} + 4T_{13} - T_{14} - T_{23} = 40;$$

$$T_{14} = \frac{40 + T_{13} + 20 + T_{24}}{4} \Rightarrow -T_{13} + 4T_{14} - T_{24} = 60;$$

$$\begin{aligned}
T_{21} &= \frac{T_{11}+20+T_{22}+T_{31}}{4} \Rightarrow -T_{11} + 4T_{21} - T_{22} - T_{31} = 20; \\
T_{22} &= \frac{T_{12}+T_{21}+T_{23}+T_{32}}{4} \Rightarrow -T_{12} - T_{21} + 4T_{22} - T_{23} - T_{32} = 0; \\
T_{23} &= \frac{T_{13}+T_{22}+T_{24}+T_{33}}{4} \Rightarrow -T_{13} - T_{22} + 4T_{23} - T_{24} - T_{33} = 0; \\
T_{24} &= \frac{T_{14}+T_{23}+20+T_{34}}{4} \Rightarrow -T_{14} - T_{23} + 4T_{24} - T_{34} = 20; \\
T_{31} &= \frac{T_{21}+20+T_{32}+T_{41}}{4} \Rightarrow -T_{21} + 4T_{31} - T_{32} - T_{41} = 20; \\
T_{32} &= \frac{T_{22}+T_{31}+T_{33}+T_{42}}{4} \Rightarrow -T_{22} - T_{31} + 4T_{32} - T_{33} - T_{42} = 0; \\
T_{33} &= \frac{T_{23}+T_{32}+T_{34}+T_{43}}{4} \Rightarrow -T_{23} - T_{32} + 4T_{33} - T_{34} - T_{43} = 0; \\
T_{34} &= \frac{T_{24}+T_{33}+20+T_{44}}{4} \Rightarrow -T_{24} - T_{33} + 4T_{34} - T_{44} = 20; \\
T_{41} &= \frac{T_{31}+20+T_{42}+10}{4} \Rightarrow -T_{31} + 4T_{41} - T_{42} = 30; \\
T_{42} &= \frac{T_{32}+T_{41}+T_{43}+10}{4} \Rightarrow -T_{32} - T_{41} + 4T_{42} - T_{43} = 10; \\
T_{43} &= \frac{T_{33}+T_{42}+T_{44}+10}{4} \Rightarrow -T_{33} - T_{42} + 4T_{43} - T_{44} = 10; \\
T_{44} &= \frac{T_{34}+T_{43}+20+10}{4} \Rightarrow -T_{34} - T_{43} + 4T_{44} = 30.
\end{aligned}$$

#### 4.1.1 Resolvendo por Eliminação Gaussiana

Escrevendo a matriz ampliada do sistema, a partir da forma  $A \times x = b$ , temos

$$\left[ \begin{array}{cccccccccccccccc|c}
4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 60 \\
-1 & 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 40 \\
0 & -1 & 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 40 \\
0 & 0 & -1 & 4 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 60 \\
-1 & 0 & 0 & 0 & 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 20 \\
0 & -1 & 0 & 0 & -1 & 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & -1 & 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 20 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 20 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & -1 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & 0 & 0 & 0 & -1 & 20 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 4 & -1 & 0 & 0 & 30 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & -1 & 0 & 10 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & -1 & 10 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & 30
\end{array} \right].$$

Utilizando-se da fórmula obtida para o cálculo do custo computacional neste método, fazemos uma previsão de

$$\#(Elim.) = \frac{4n^3 + 9n^2 - 7n}{6} \Rightarrow \#(Elim.) = \frac{4 \times 16^3 + 9 \times 16^2 - 7 \times 16}{6}$$

$$= 3096 \text{ operações}$$

a serem realizadas para a obtenção da solução do sistema.

Este número pode diminuir pela quantidade de zeros existentes na matriz  $A$ . Vejamos como o processo ocorre.

Começando com a 1ª coluna, temos que anular apenas com os termos  $a_{21}$  e  $a_{51}$ , pois os demais já são iguais a 0. Para isso, primeiro substituímos  $L_2$  pela combinação  $L_2 + \frac{1}{4}L_1$ , e em seguida  $L_5$ , pela combinação  $L_5 + \frac{1}{4}L_1$ . Assim, a matriz se torna

$$\begin{bmatrix} 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 60 \\ 0 & 3,75 & -1 & 0 & -0,25 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 55 \\ 0 & -1 & 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 40 \\ 0 & 0 & -1 & 4 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 60 \\ 0 & -0,25 & 0 & 0 & 3,75 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 35 \\ 0 & -1 & 0 & 0 & -1 & 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 4 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 20 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 4 & -1 & 0 & 0 & -1 & 0 & 0 & 20 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 4 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 4 \end{bmatrix}$$

*Em cada linha, foram realizadas 17 multiplicações e 16 adições, totalizando 33 operações por linha. No entanto, só foi necessário repetir o procedimento para 2 linhas.*  
**TOTAL = 66** operações

A processo é análogo ao anular os termos de todas as colunas. Resumindo, teremos:

Coluna	× por linha	+ por linha	# Equações	Total Op.
2	16	15	3	<b>93</b>
3	15	14	4	<b>116</b>
4	14	13	4	<b>108</b>
5	13	12	4	<b>100</b>
6	12	11	4	<b>92</b>
7	11	10	4	<b>84</b>
8	10	9	4	<b>76</b>
9	9	8	4	<b>68</b>
10	8	7	4	<b>60</b>
11	7	6	4	<b>52</b>
12	6	5	4	<b>44</b>
13	5	4	3	<b>27</b>
14	4	3	2	<b>14</b>
15	3	2	1	<b>5</b>

Esse cálculo nos dá um total de **1005** operações, para o escalonamento da matriz apresentada.

Como já visto, transformamos a matriz ampliada escalonada em um sistema linear equivalente triangular, e encontramos os valores das incógnitas, começando por  $T_{44}$  e terminando em  $T_{11}$ . Dessa forma, obtemos, como solução, o vetor

$$\bar{x} = \{27,79768376; 27,81952297; 30,27965058; 28,41271564; 23,37121208; 24,88636358; 24,88636360; 23,37121214; 19,96212121; 20,11363632; 20,11363635; 19,96212123; 16,36363634; 15,49242425; 15,49242425; 16,36363635\}.$$

Nesta segunda parte, não há vantagem se temos muitos coeficientes iguais a 0 ou não, pois o computador fará a conta de qualquer forma. Podemos então usar a fórmula obtida para o cálculo do custo computacional na fase de obtenção das incógnitas. Assim,

$$\#(Op.) = n^2 \Rightarrow \#(Op.) = 16^2 = \mathbf{256} \text{ operações.}$$

Assim, o total de operações necessárias para a resolução do sistema apresentado, utilizando-se da eliminação Gaussiana, é igual a  $1005 + 256 = \mathbf{1261}$  operações.

Vejamos essa contagem através do método de Gauss-Seidel.

#### 4.1.2 Resolvendo por Gauss-Seidel

Primeiramente, o critério de Sassenfeld nos garante que a sequência  $\{x^{(k)}\}$  gerada pelo método converge para a solução. De fato, temos que, para a matriz  $A$  relativa ao problema,

$$\beta = \max_{1 \leq i, j \leq 4} \{\beta_{ij}\} = \beta_{33} = 0,93359375 < 1.$$

Definimos o vetor

$$\vec{x}^{(k)} = \{T_{11}^{(k)}, T_{12}^{(k)}, T_{13}^{(k)}, T_{14}^{(k)}, T_{21}^{(k)}, T_{22}^{(k)}, T_{23}^{(k)}, T_{24}^{(k)}, \\ T_{31}^{(k)}, T_{32}^{(k)}, T_{33}^{(k)}, T_{34}^{(k)}, T_{41}^{(k)}, T_{42}^{(k)}, T_{43}^{(k)}, T_{44}^{(k)}\}.$$

Como a temperatura nos bordos varia entre  $10^{\circ}C$  e  $40^{\circ}C$ , tomemos o ponto médio,  $25^{\circ}C$ . Assim, escolhemos a aproximação inicial

$$\vec{x}^{(0)} = \{25; 25; 25; 25; 25; 25; 25; 25; 25; 25; 25; 25; 25; 25; 25; 25\}.$$

A fórmula de iteração é dada por

$$\left\{ \begin{array}{l} T_{11}^{(k+1)} = \frac{T_{12}^{(k)} + T_{21}^{(k)} + 60}{4} \\ T_{12}^{(k+1)} = \frac{T_{11}^{(k+1)} + T_{13}^{(k)} + T_{22}^{(k)} + 40}{4} \\ T_{13}^{(k+1)} = \frac{T_{12}^{(k+1)} + T_{14}^{(k)} + T_{23}^{(k)} + 40}{4} \\ T_{14}^{(k+1)} = \frac{T_{13}^{(k+1)} + T_{24}^{(k)} + 60}{4} \\ T_{21}^{(k+1)} = \frac{T_{11}^{(k+1)} + T_{22}^{(k)} + T_{31}^{(k)} + 20}{4} \\ T_{22}^{(k+1)} = \frac{T_{12}^{(k+1)} + T_{21}^{(k+1)} + T_{23}^{(k)} + T_{32}^{(k)}}{4} \\ T_{23}^{(k+1)} = \frac{T_{13}^{(k+1)} + T_{22}^{(k+1)} + T_{24}^{(k)} + T_{33}^{(k)}}{4} \\ T_{24}^{(k+1)} = \frac{T_{14}^{(k+1)} + T_{23}^{(k+1)} + T_{34}^{(k)} + 20}{4} \\ T_{31}^{(k+1)} = \frac{T_{21}^{(k+1)} + T_{32}^{(k)} + T_{41}^{(k)} + 20}{4} \\ T_{32}^{(k+1)} = \frac{T_{22}^{(k+1)} + T_{31}^{(k+1)} + T_{33}^{(k)} + T_{42}^{(k)}}{4} \\ T_{33}^{(k+1)} = \frac{T_{23}^{(k+1)} + T_{32}^{(k+1)} + T_{34}^{(k)} + T_{43}^{(k)}}{4} \\ T_{34}^{(k+1)} = \frac{T_{24}^{(k+1)} + T_{33}^{(k+1)} + T_{44}^{(k)} + 20}{4} \\ T_{41}^{(k+1)} = \frac{T_{31}^{(k+1)} + T_{42}^{(k)} + 30}{4} \\ T_{42}^{(k+1)} = \frac{T_{32}^{(k+1)} + T_{41}^{(k+1)} + T_{43}^{(k)} + 10}{4} \\ T_{43}^{(k+1)} = \frac{T_{33}^{(k+1)} + T_{42}^{(k+1)} + T_{44}^{(k)} + 10}{4} \\ T_{44}^{(k+1)} = \frac{T_{34}^{(k+1)} + T_{43}^{(k+1)} + 30}{4} \end{array} \right. .$$

Os resultados obtidos a cada iteração são os seguintes:

1<sup>a</sup> Iteração:

$$x^{(1)} = \{27,5; 29,375; 29,84375; 28,7109375; 24,375; 25,9375; 26,4453125; \\ 25,0390625; 23,59375; 24,8828125; 25,33203125; 23,84277344; 19,6484375; 19,8828125; \\ 20,05371094; 18,47412109\};$$

$$d_r^{(1)} = 0,218668194.$$

2<sup>a</sup> Iteração:

$$x^{(2)} = \{28,4375; 31,0546875; 31,55273438; 29,14794922; 24,4921875; 26,71875; \\ 27,16064453; 25,0378418; 22,25585938; 23,54736328; 23,65112305; \\ 21,79077148; 18,03466797; 17,90893555; 17,50854492; 17,3248291\};$$

$$d_r^{(2)} = 0,080663881.$$

⋮

8ª Iteração:

$$x^{(8)} = \{28,81861614; 31,41742424; 31,37340094; 28,7365699; 23,61612155; \\ 25,20987861; 25,14958141; 23,50314542; 20,16343934; 20,37793396; 20,32780475; \\ 20,06921902; 16,46492245; 15,62503243; 15,59969077; 16,41722745\};$$

$$d_r^{(8)} = 0,005302193.$$

9ª Iteração:

$$x^{(9)} = \{28,75838645; 31,3354165; 31,30539195; 28,70213434; 23,5329261; \\ 25,09896449; 25,05882665; 23,457545; 20,09394563; 20,28643682; 20,25354332; \\ 20,03207894; 16,42974451; 15,57896803; 15,5624347; 16,39862841\};$$

$$d_r^{(9)} = 0,003539577.$$

Assim, para obtermos a precisão desejada, com erro menor que  $\epsilon < 0,005$ , foram necessárias 9 iterações

$$\Rightarrow \#(Op.) = 9 \times (2n^2 - n), n = 16 \Rightarrow \#(Op.) = 9 \times (2 \times 16^2 - 16) = \mathbf{4464 \text{ operações.}}$$

## 4.2 Eliminação Gaussiana x Gauss-Seidel

Após resolver o mesmo exemplo utilizando primeiramente a Eliminação Gaussiana e em seguida o método de Gauss-Seidel, concluímos que o custo computacional dos métodos iterativos é, de fato, maior que o custo dos métodos diretos.

Pudemos observar que foram necessárias 4464 operações para a conclusão do processo por Gauss-Seidel, e 1261 operações pela Eliminação Gaussiana. Mesmo se tivéssemos uma matriz  $A$  sem zeros, o que nos obrigaria a aplicar as combinações lineares em todas as linhas, a cada coluna sendo eliminada, teríamos uma contagem de 3096 operações, ainda menor que o custo de Gauss-Seidel.

Na tentativa de generalizar para o caso  $n \times n$ , façamos a seguinte análise:

São necessárias  $\frac{4n^3 + 9n^2 - 7n}{6}$  operações na eliminação Gaussiana, e  $2n^2 - n$  operações por iteração no método de Gauss-Seidel. Assim, dado um número  $k \in \mathbb{N}$  de

iterações, o custo operacional será de  $k(2n^2 - n)$  operações. Analisemos para que valores de  $k$  teremos  $\#(Elim.) < \#(Gauss - Seidel)$ . Tomando  $k > \frac{4n^2 + 9n - 7}{12n - 6}$ , teremos

$$\begin{aligned} k &> \frac{4n^2 + 9n - 7}{6(2n - 1)} \\ \Rightarrow k &> \frac{4n^3 + 9n^2 - 7n}{6n(2n - 1)} \\ \Rightarrow k(2n^2 - n) &> \frac{4n^3 + 9n^2 - 7n}{6}. \end{aligned}$$

Assim, para os valores de  $k$  mencionados, temos que o número de operações ao aplicar Gauss-Seidel é maior que o número ao aplicar a Eliminação de Gauss.

No exemplo resolvido, bastaria que tomássemos

$$k > \frac{4 \times 256 + 9 \times 16 - 7}{12 \times 16 - 6} = \frac{1161}{186} > 6.$$

Assim, para qualquer valor de  $k$  menor ou igual a 6, Gauss-Seidel seria mais vantajoso. Entretanto, após a 6ª iteração, a solução obtida ainda possuía um erro relativo de  $d^{(6)} = 0,010898508$ , maior que o dobro do estipulado.

No entanto, vejamos o que acontece quando este problema é proposto para um físico com experiência no campo da Termodinâmica. Dado que determinados pontos da malha estão mais próximos de bordos com temperaturas diferentes uns dos outros, fazendo-se uma análise prévia sua aproximação inicial pode ser dada por

$$\vec{x}^{(0)} = \{30; 30; 30; 30; 25; 25; 25; 25; 20; 20; 20; 20; 15; 15; 15; 15\}.$$

Desta forma, aplicando a fórmula de iteração de Gauss-Seidel, obtemos

		Iteração			
		1	2	3	4
<b>Iteração</b>	<b>T11</b>	28,75	28,59375	28,62304688	28,64562988
	<b>T12</b>	30,9375	31,23046875	31,24145508	31,19865417
	<b>T13</b>	31,484375	31,42089844	31,25030518	31,17780685
	<b>T14</b>	29,12109375	28,75854492	28,6516571	28,63086224
	<b>T21</b>	23,4375	23,26171875	23,34106445	23,36235046
	<b>T22</b>	24,84375	24,921875	24,89868164	24,86053467
	<b>T23</b>	25,33203125	25,0012207	24,86091614	24,85329151
	<b>T24</b>	23,61328125	23,35632324	23,34564209	23,35127115
	<b>T31</b>	19,609375	19,81933594	19,90509033	19,92338181
	<b>T32</b>	19,86328125	20,01098633	20,0202179	20,05932331
	<b>T33</b>	20,04882813	19,93835449	20,02918243	20,06852865
	<b>T34</b>	19,66552734	19,86999512	19,92093086	19,94004369
	<b>T41</b>	16,15234375	16,26831055	16,31095886	16,33689404
	<b>T42</b>	15,25390625	15,33874512	15,42419434	15,45919657
	<b>T43</b>	15,07568359	15,36560059	15,44056892	15,46702504
	<b>T44</b>	16,18530273	16,30889893	16,34037495	16,35176718
<b>Erro Relativo</b>	0,049627792	0,011538462	0,005458931	0,002323764	

Vemos que, neste caso, a solução com a precisão pretendida foi obtida com 4 iterações

$$\Rightarrow \#(Op.) = 4 \times (2n^2 - n), n = 16 \Rightarrow \#(Op.) = 4 \times (2 \times 16^2 - 16) = \mathbf{1984 \text{ operações.}}$$

Considerando o número de operações efetivas realizadas pela eliminação Gaussiana neste exemplo, Gauss-Seidel continua com um custo computacional maior. Porém, quando não levamos em conta os diversos zeros que facilitaram o escalonamento da matriz  $A$ , o método iterativo torna-se mais vantajoso (1984 contra 3096 operações).

Além disso, vejamos o que acontece na eliminação Gaussiana. Analisaremos a solução encontrada ao utilizar os métodos e a compararemos com a solução obtida se tivéssemos continuado a iterar o algoritmo de Gauss-Seidel.

Elim.	G.-Seidel - 1ª aprox.	G.-Seidel - 2ª aprox.	G.-Seidel (continuado)
27,79768376	28,75838645	28,64562988	28,63636364
27,81952297	31,3354165	31,19865417	31,17424242
30,27965058	31,30539195	31,17780685	31,17424242
28,41271564	28,70213434	28,63086224	28,63636364
23,37121208	23,5329261	23,36235046	23,37121212
24,88636358	25,09896449	24,86053467	24,88636364
24,88636360	25,05882665	24,85329151	24,88636364
23,37121214	23,457545	23,35127115	23,37121212
19,96212121	20,09394563	19,92338181	19,96212121
20,11363632	20,28643682	20,05932331	20,11363636
20,11363635	20,25354332	20,06852865	20,11363636
19,96212123	20,03207894	19,94004369	19,96212121
16,36363634	16,42974451	16,33689404	16,36363636
15,49242425	15,57896803	15,45919657	15,49242424
15,49242425	15,5624347	15,46702504	15,49242424
16,36363635	16,39862841	16,35176718	16,36363636

Vemos que, para as primeiras incógnitas calculadas (lembrando que, numa matriz triangular superior, as variáveis são resolvidas de baixo para cima), a Eliminação Gaussiana está ainda mais precisa do que o método de Gauss-Seidel (com o erro proposto). No entanto, quando chegamos nas últimas incógnitas, o erro na eliminação Gaussiana se torna maior. Isto é decorrente da propagação de erros de arredondamentos do método. A cada conta resolvida e aproximada, aumenta-se a distância para a solução exata. Este problema se estende aos demais métodos diretos.

No caso dos métodos iterativos, considerando os sistemas que cumprem os critérios de convergência, os erros de arredondamento não afetam o resultado final, pois a sequência de aproximações converge para a solução independente do vetor inicial escolhido. Apenas os arredondamentos na última iteração influiriam nas contas. Assim, apesar de ter o custo computacional maior, os métodos iterativos garantem uma maior exatidão da solução.

Além disso, com uma análise prévia dos dados do problema, melhorando assim a aproximação inicial proposta, pode-se diminuir o número de iterações necessárias para a obtenção da previsão desejada, e assim, abrandar o custo computacional do método.

Relaxar as condições iniciais do problema também podem tornar o método de Gauss-Seidel vantajoso, no quesito do custo computacional. No entanto, sua vantagem é exatamente a sua capacidade de precisão. Quando atendidos os critérios de convergência, qualquer erro de arredondamento cometido em uma determinada iteração não prejudicará o resultado final, dado que qualquer aproximação inicial alcança a convergência.

Em contrapartida, podemos destacar a versatilidade dos métodos diretos. Qualquer sistema linear pode ser resolvido através deles, sem exigências preliminares, ao contrário dos iterativos.

Assim, dada a intenção de quem se depara com um problema que envolve sistemas lineares, a opção pelo método será diferente. Pode-se levar em conta rapidez, precisão ou facilidade. Desta forma, cada método tem a sua importância no ensino da Matemática.

## Referências

- [1] ASANO, C.H. ; COLLI, E. **Cálculo Numérico - Fundamentos e Aplicações**. USP, 2009
- [2] CARVALHO, P.C.P. et al. **A Matemática do Ensino Médio, Volume 3**. 5<sup>a</sup> ed., SBM, 2005
- [3] DOS SANTOS, J.D. ; DA SILVA, Z.C. **Métodos Numéricos**. Editora Universitária - UFPE, 2006
- [4] EDWARDS JR., C.H. ; PENNEY, D.E. **Introdução à Álgebra Linear**. LTC, 1998
- [5] FRANCO, N.B. **Cálculo Numérico**. Pearson, 2006
- [6] GARCIA, A. ; LEQUAIN, Y. **Elementos de Álgebra**. 5<sup>a</sup> ed., IMPA, 2010
- [7] HAFFNER, J.F. ; PEREIRA, L.F.A. **Análise de Sistemas Lineares**. PUC-RS, disponível em <http://www.feng.pucrs.br/gacs/new/disciplinas/asl/apostilas/Aula01.pdf>, acessado em 4/4/2014
- [8] HIBBELER, R.C. **Estática - Mecânica para Engenheiros**. 5<sup>a</sup> ed., Prentice Hall, 2003
- [9] HUMES, A.F.P.C. et al. **Noções de Cálculo Numérico**. McGraw-Hill, 1984
- [10] LIMA, E.L. **Álgebra Linear**. 6<sup>a</sup> ed., IMPA, 2003
- [11] LOPES, V.L.R. ; RUGGIERO, M.A.G. **Cálculo Numérico - Aspectos Teóricos e Computacionais**. 2<sup>a</sup> ed., Pearson, 2010
- [12] MILNE, W.E. **Cálculo Numérico**. 2<sup>a</sup> ed., Polígono, 1968
- [13] SOUZA, J.M.E. **Programa de Aprendizagens Significativas - Matemática, Volume 2**. Inment, 2008