



UNIVERSIDADE FEDERAL DO AMAPÁ
CURSO DE LICENCIATURA EM MATEMÁTICA

Leomir Braga Monteiro
Rosivan Ferreira e Ferreira

**ESTUDO TEÓRICO DOS MÉTODOS
DIRETOS E ITERATIVOS NA SOLUÇÃO
DE SISTEMAS LINEARES**

Macapá-AP
2012

Leomir Braga Monteiro
Rosivan Ferreira e Ferreira

ESTUDO TEÓRICO DOS MÉTODOS DIRETOS E ITERATIVOS NA SOLUÇÃO DE SISTEMAS LINEARES

Trabalho de conclusão de curso apresentado ao colegiado de Matemática da Universidade Federal do Amapá, como parte das exigências para a obtenção do título de Licenciatura em Matemática, sob a orientação do Prof^o. Dr. GUZMÁN EULÁLIO ISLA CHAMILCO.

Macapá-AP
2012

LEOMIR BRAGA MONTEIRO
ROSIVAN FERREIRA E FERREIRA

ESTUDO TEÓRICO DOS MÉTODOS DIRETOS E ITERATIVOS NA SOLUÇÃO DE SISTEMAS LINEARES

Este Trabalho de Conclusão de Curso foi julgado e aprovado pela comissão avaliadora do Colegiado de Matemática da Universidade Federal do Amapá. Composta pelos integrantes abaixo-relacionados:

AVALIADORES:

Prof. Dr. Guzmán Eulálio Isla Chamilco

Prof. Dr. José Walter Cárdenas Sotil

Prof. Dr. Erasmo Senger

Avaliado em: ____/____/____

“O SENHOR guardará a tua entrada e a tua
saída, desde agora e para sempre.”

(Salmo 121:8)

Agradecimentos

Primeiramente agradecemos a Deus por mais esta conquista, dando-nos saúde e força para superarmos as adversidades encontradas no caminho.

Aos nossos familiares que nos deram todo apoio necessário para que pudéssemos vencer mais essa etapa de nossas vidas.

Ao professor Gúzman Eulálio Isla Chamilco pela orientação, dedicação, amizade, companheirismo e incentivo durante todo o curso.

Agradecemos a todos os professores do colegiado, que contribuíram grandemente para a nossa formação acadêmica em especial aos professores José Walter Cárdenas Sotil, Márcio Aldo Lobato Baia, Erasmo Senger enfim, todos os professores do colegiado, que não mediram esforços para termos uma formação acadêmica de qualidade.

Aos colegas da turma de matemática 2008 (Unifap), que se mostraram verdadeiros amigos e nos proporcionaram muitos momentos de alegria, pelas brincadeiras, grupos de estudo, a galera do futebol e as comemorações ao fim de cada semestre.

Somos gratos à instituição UNIFAP, por nos proporcionar a nossa formação acadêmica, enfim a todos aqueles que contribuíram direta e indiretamente a nossa formação.

Resumo

Neste trabalho de conclusão de curso foram analisados Métodos Diretos e Iterativos que são de grande importância na resolução de sistemas lineares algébricos. Os quais são de uso frequente desde os problemas comuns do cotidiano até os sistemas de maior complexidade como por exemplo os sistemas de equações diferenciais. Na definição e resolução de sistemas lineares, a forma matricial sempre está presente, seja na esquematização dos métodos diretos, seja no estudo da convergência dos Métodos Iterativos. Por isso, é definido o conceito de norma no espaço das matrizes. Apresentamos como Métodos Diretos a Eliminação de Gauss, a Decomposição LU e a Fatoração de Cholesky e como Métodos Iterativos o Método de Gauss-Jacobi, Gauss-Seidel e SOR. Para cada um desses métodos foram descritos algoritmos numéricos para gerar a solução. Não foram utilizados programas computacionais para implementar esses métodos. Focamos o trabalho na parte teórica fazendo uma comparação (teórica) entre os métodos. Os métodos Iterativos são preferidos em matrizes de pequeno porte, por ter menos restrições ao seu uso que os Métodos Iterativos. Os Métodos Iterativos mostram vantagem quando a matriz do sistema é esparsa e de grande porte, uma vez que o método não altera a estrutura de tal matriz.

Palavras-chave: Matrizes, Sistema Lineares, Método Diretos, Métodos Iterativo, Normas de matrizes.

Resumen

Em éste trabajo estudiamos y analizamos los métodos directos e iterativos que son de gran importancia en la resolución de sistemas algebraicos lineales. En la formulación y resolución de sistemas lineales, la forma de una matriz siempre esta presente. Sea en el diseño de los métodos directos como en el estudio de la convergencia de los métodos iterativos, utilizando conceptos de normas en el espacio de las matrices. Los métodos estudiados son los directos: Eliminación de Gauss, la descomposición LU y la factoración de Cholesky, encuanto los iterativos: Gauss - Jacobi, Gauss - Seidel y SOR. Para cada un de estos métodos se han descrito algoritmos numéricos para la generación de la solución. Damos mayor énfasis al trabajo teórico, haciendo una comparación entre los métodos directos e iterativos, observando que los métodos iterativos muestran mejor ventaja cuando la matriz del sistema es grande y escassa, puesto que el método no altera la estructura de esas matrices .

Palabras Claves: Matrices, Sistemas Lineales, Métodos Directos, Métodos Iterativos, Normas de Matrices.

Índice

Agradecimentos	1
Resumo	vi
Resumen	vii
1 Introdução	12
2 Teoria das Matrizes	14
2.1 Matrizes	14
2.1.1 Operações com Matrizes	15
2.2 Sistemas Lineares	19
2.2.1 Solução de um Sistema Linear	20
2.2.2 Sistemas Equivalentes	21
3 Normas	22
3.1 Definições	22
3.1.1 Normas em \mathfrak{R}^n	22
3.2 Normas Matriciais	24
3.3 Convergência em Norma	28
4 Métodos Diretos	30
4.0.1 Métodos da Eliminação de Gauss	30
4.0.2 Fatoração LU	38
4.0.3 Fatoração de Cholesky	41
4.1 Métodos Iterativos	45
4.1.1 Obtenção dos Métodos Iterativos	46
4.1.2 Critério de Convergência dos Métodos Iterativos	47
4.1.3 Estimativa de erro na Iteração	49
4.1.4 Teste de Parada	50
4.1.5 Método de Gauss-Jacobi	50
4.1.6 Método Iterativo de Gauss-Seidel	54

4.2	Método SOR	62
5	Aplicação e Comparação dos Métodos Diretos e Iterativos	66
5.1	Eliminação de Gauss	66
5.2	Fatoração LU	68
5.3	Fatoração de Cholesky	69
5.4	Método Iterativo de Gauss-Jacobi	69
5.5	Método Iterativo de Gauss-seidel	73
5.6	Método SOR ou Relaxação Sucessiva	75
	Considerações Finais	xvi
	Referências Bibliográficas	xviii

Capítulo 1

Introdução

O sistema é algo muito próximo e comum no cotidiano, serve a título de exemplo para se ter uma noção de quanto se gasta ou quanto se ganha em diversas atividades, como por exemplo, o custo de uma ligação telefônica, orçamento para grandes construções, planejamento e gerenciamento de empresas.

Inúmeros sistemas lineares surgem também de problemas práticos como discretização de equações diferenciais. Estima-se pelo menos 75% dos problemas científicos a solução de um sistema de equações lineares surgir em algum estágio da solução.[6]

Uma das mais usadas estruturas algébricas são as matrizes. Sua utilização nas mais variadas formulações matemáticas permite a simplificação não somente da parte teórica, mas também das próprias aplicações. Um sistema de equações $Ax = b$ pode-se resolver como $x = A^{-1}b$ desde que a inversa A^{-1} exista, entretanto o cálculo da inversa é trabalhoso e sujeito a erros de arredondamento, por isso são necessários métodos mais eficientes para tais resoluções.

Analisando a distinção dos métodos, é comum classificá-los em dois grupos. Os Métodos Diretos que após um número finito de operações fornecem a solução exata, pelo menos teoricamente, isto porque quando fazemos muitas divisões e multiplicações, introduzimos erros de arredondamento produzidos pela máquina. Por isso, os Métodos Diretos são indicados para o caso em temos um sistema de pequeno porte, de modo que, em sua solução tenhamos de fazer um número reduzido de operações, caso contrário a solução que vamos obter muitas vezes se afasta da solução real (exata).

Para sistemas de grande porte, esse efeito diminui com a aplicação de outros métodos que garantem em certos casos, qualquer precisão desejada. Estes métodos são chamados Métodos Iterativos, no qual usamos uma aproximação inicial da solução e a melhoramos quantas vezes sejam necessárias para chegarmos a uma precisão satisfatória. É claro que os Métodos Iterativos só se mostrará útil se a sequência das aproximações construídas pelos métodos convergi para a solução exata do sistema.

No capítulo 2 apresentamos a teoria das matrizes, conceito de matrizes, operações com matrizes, inversas de matrizes juntamente com o conceito de solução de sistemas de equações

lineares. No capítulo 3 definimos as normas de matrizes e normas de vetores, assim como o conceito de convergência em norma de uma sequência de vetores e matrizes. No capítulo 4 apresentamos os Métodos Diretos (Eliminação de Gauss, Fatoração LU e fatoração de Cholesky) e os Métodos Iterativos (Gauss-Jacobi, Gauss-Seidel e SOR) na resolução de sistemas lineares. No capítulo 5 apresentamos um exemplo de sistema linear no qual obtemos a sua solução aplicando os métodos vistos neste trabalho, fazendo uma comparação entre os mesmos.

Capítulo 2

Teoria das Matrizes

Neste capítulo apresentaremos alguns conceitos importantes da teoria de matrizes, que serão úteis no desenvolvimento de métodos numéricos para a resolução de sistemas lineares.

2.1 Matrizes

Definição 2.1.1 *Dados dois números m e n naturais e não nulos, chama-se matriz m por n (indica-se $m \times n$) toda tabela A formada por número reais distribuídos em m linhas e n colunas. Em uma matriz A qualquer, cada elemento é indicado por a_{ij} . O índice i indica a linha e o índice j indica a coluna a qual o elemento pertence. Com a convenção de que as linhas sejam numeradas de cima para baixo (de 1 até m) e as colunas, da esquerda para a direita (de 1 até n) uma matriz $m \times n$ é representada por:*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

ou representamos também a matriz A por $A = (a_{ij})_{m \times n}$. Note que $1 \leq i \leq m$ e $1 \leq j \leq n$. Indicaremos por $M_{m \times n}(\mathbb{R})$ o conjunto das matrizes reais do tipo $m \times n$. Se $m = n$ ao invés de $M_{m \times n}(\mathbb{R})$, usaremos a notação $M_n(\mathbb{R})$. Cada matriz de $M_n(\mathbb{R})$ chama-se matriz quadrada de ordem n . Em contraposição, quando $m \neq n$, uma matriz do tipo $m \times n$ se diz retangular. Uma matriz 1×1 se identifica com o número real a_{11} .

Definição 2.1.2 *Chama-se diagonal principal de uma matriz quadrada A de ordem n , os elementos cujo índice das linhas é igual ao índice das colunas, isto é:*

$$\{a_{ij} \mid i = j\} = \{a_{11}, a_{22}, a_{33}, \dots, a_{nn}\}$$

Chama-se diagonal secundária da matriz quadrada A de ordem n o conjunto dos elementos que têm soma dos índices igual a $n + 1$, isto é:

$$\{a_{ij} \mid i + j\} = \{a_{1n}, a_{2,n-1}, a_{3,n-2} \dots, a_{n1}\}$$

Exemplo 2.1 A matriz

$$M = \begin{pmatrix} 1 & 3 & -1 \\ -2 & -3 & 2 \\ 1 & 2 & 4 \end{pmatrix}$$

é quadrada de ordem 3 Sua diagonal principal é $(1, -3, 4)$ e sua diagonal secundária é $(1, -3, -1)$

Será definido a seguir matrizes que são comuns na teoria:

1. **Matriz Linha:** é toda matriz do tipo $1 \times n$, isto é, uma matriz formada por uma única linha.
2. **Matriz Coluna:** é toda matriz do tipo $m \times 1$, isto é, uma matriz formada por uma única coluna.
3. **Matriz Nula:** é toda Matriz cujos os elementos são todos iguais a zero.
4. **Matriz identidade:** denotada por I_n , é uma matriz quadrada de ordem n , em que os elementos da diagonal principal são iguais a 1, e todos os outros elementos são iguais a zero.
5. **Matriz Diagonal:** é toda matriz quadrada em que os elementos, que não pertencem a diagonal principal são iguais a zero

2.1.1 Operações com Matrizes

A seguir apresentaremos definições para operações com matrizes, as quais permitiram formular equações algébricas em termos de matrizes.

Definição 2.1.3 (Igualdade de matrizes): das matrizes $A = (a_{ij})_{m \times n}$ e $B = (a_{ij})_{m \times n}$, são iguais quando:

$$a_{ij} = b_{ij}, \forall i = 1, \dots, m \text{ e } \forall j = 1, \dots, n$$

Isto quer dizer que duas matrizes para serem iguais, devem ser do mesmo tipo e apresentar todos os elementos correspondente (elementos com índices iguais) iguais.

Definição 2.1.4 (Adição de Matrizes): dadas duas matrizes $A = (a_{ij})_{m \times n}$ e $B = (b_{ij})_{m \times n}$, chama-se soma de $A + B$ a matriz $C = (c_{ij})_{m \times n}$, tal que, $c_{ij} = a_{ij} + b_{ij}$, para todo i e todo j .

Em outras palavras, a matriz soma C é do mesmo tipo de A e B e é tal que cada um de seus elementos é a soma de elementos correspondentes de A e B .

Propriedade 2.1: Se $A, B, C \in M_{m \times n}(\mathfrak{R})$, então:

1. Associatividade: $(A + B) + C = A + (B + C)$
2. Comutativa: $A + B = B + A$
3. Elemento Neutro: Existe S tal que, $S + A = S + A = A$
4. Elemento Oposto: Para toda matriz $A \exists A'$ tal que $A + A' = A' + A = S$

Definição 2.1.5 (Multiplicação por escalar): Seja a matriz $A = (a_{ij})_{m \times n}$ e dado um número k , a multiplicação de k pela a matriz A é definida como a matriz B (do tipo $m \times n$) $B = k \cdot A$, em que $b_{ij} = k \cdot a_{ij}$

O significado disso é: B é obtida de A , multiplicando-se todos os seus elementos por k

Propriedades 2.2: Se $A, B \in M_{m \times n}(\mathfrak{R})$ e $\alpha, \beta \in \mathfrak{R}$

1. $\alpha \cdot (\beta \cdot A) = (\alpha \cdot \beta) \cdot A$
2. $\alpha \cdot (A + B) = \alpha \cdot A + \alpha \cdot B$
3. $(\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot B$
4. $1 \cdot A = A$

Definição 2.1.6 (Produto de Matrizes): Dadas duas matrizes $A = (a_{ij})_{m \times n}$ e $B = (b_{jk})_{n \times p}$, chama-se produto de $A \cdot B$, a matriz $C = (c_{ik})_{m \times p}$ tal que:

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{ij}b_{jk} = \sum_{j=1}^n a_{ij}b_{jk}, \forall i = 1, \dots, m, \forall k = 1, \dots, p.$$

Teorema 2.1.1 : Se $A = (a_{ij})_{m \times n}$, então $A \cdot I_n = A$ e $I_m \cdot A = A$.

Demonstração:

I- Sendo $I_n = (\alpha_{ij})_{n \times n}$ e $B = A \cdot I_n = (b_{ij})_{n \times n}$, temos:

$$b_{ij} = a_{i1}\alpha_{1j} + a_{i2}\alpha_{2j} + \dots + a_{ii}\alpha_{ii} + \dots + a_{in}\alpha_{nj}$$

$$b_{ij} = a_{i1} \cdot 0 + a_{i2} \cdot 0 + \dots + a_{ii} \cdot 0 + \dots + a_{in} \cdot 0$$

$$b_{ij} = a_{ij}, \forall i, j$$

Logo, $A \cdot I_n = A$

II- Análoga.

Propriedade 2.3.

Para a multiplicação de matrizes valem as seguintes propriedades:

1- Associatividade

$$(AB)C = A(BC),$$

para quaisquer matrizes do tipo $A = (a_{ij})_{m \times n}$, $B = (b_{jk})_{n \times p}$ e $C = (c_{ij})_{p \times r}$

2- Distribuidade à direita em relação à adição:

$$(A + B)C = AC + AB$$

para quaisquer matrizes do tipo $A = (a_{ij})_{m \times n}$, $B = (b_{jk})_{m \times n}$ e $C = (c_{ij})_{n \times p}$

3- Distribuidade à esquerda em relação à adição:

$$C(A + B) = CA + CB$$

para quaisquer matrizes do tipo $A = (a_{ij})_{m \times n}$, $B = (b_{jk})_{m \times n}$ e $C = (c_{ij})_{p \times m}$

$$4- (\alpha A)B = A(\alpha B) = \alpha(AB)$$

para quaisquer que sejam as matrizes $A = (a_{ij})_{m \times n}$, $B = (b_{jk})_{n \times p}$ e $\alpha \in \mathfrak{R}$

Observações:

1- Nota-se que a multiplicação de matrizes não é comutativa, ou seja, para duas matrizes quaisquer A e B não é válido afirmar que $AB = BA$.

2- Quando A e B são tais que $AB = BA$, falamos que A e B comutam. Nota-se que uma condição necessária para A e B comutarem é que elas sejam quadradas e de mesma ordem.

3- É importante observar também que :

$$AB = 0 \Rightarrow A = 0 \text{ ou } B = 0$$

não é válida para as matrizes, ou seja, é possível encontrar duas matrizes não nulas onde o produto é igual a matriz nula.

Definição 2.1.7 (Matriz Transposta): Dada uma matriz $A = (a_{ij})_{m \times n}$, chama-se transposta de A matriz $A^T = (a'_{ij})_{n \times m}$ tal que $a'_{ji} = a_{ij}$, para todo i e todo j. Isto significa que, por exemplo, $a'_{11}, a'_{21}, \dots, a'_{n1}$. São respectivamente iguais a $a_{11}, a_{21}, \dots, a_{n1}$; vale dizer que a 1ª coluna de A^T é igual à 1ª de A. Repetindo o raciocínio, chegaríamos à conclusão de que as colunas de A^T são ordenadamente iguais às linhas de A.

Exemplo 2.7.

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \\ -1 & 3 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 3 \end{pmatrix}$$

Propriedade 2.4. A matriz transposta tem as seguintes propriedades:

$$1- (A^T)^T = A \quad \forall \text{ matriz } A = (a_{ij})_{m \times n};$$

$$2- \text{ Se } A = (a_{ij})_{m \times n} \text{ e } B = (b_{ij})_{m \times n}, \text{ então } (A + B)^T = A^T + B^T$$

$$3- \text{ Se } A = (a_{ij})_{m \times n} \text{ e } \alpha \in \mathfrak{R}, \text{ então } (\alpha \cdot A)^T = \alpha \cdot A^T$$

$$4- \text{ Se } A = (a_{ij})_{m \times n} \text{ e } B = (b_{jk})_{n \times p}, \text{ então } (AB)^T = B^T A^T$$

Definição 2.1.8 (Matriz simétrica) Chama-se matriz simétrica toda matriz quadrada A, de ordem n, tal que $A^T = A$.

Resulta da definição que, se $A = (a_{ij})$ é uma matriz simétrica, temos:

$$a_{ij} = a_{ji}; \forall i, j \in 1, 2, 3, \dots, n$$

ou seja, os elementos simetricamente dispostos em relação a diagonal principal são iguais.

Exemplo 2.8.: As matrizes abaixo são simétricas.

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$$

Definição 2.1.9 (Matriz Anti-Simétrica): Denomina-se matriz anti-simétrica a toda matriz quadrada A de ordem n , tal que $A^T = -A$.

Resulta da definição que, se $A = (a_{ij})$ é uma matriz simétrica, temos:

$$a_{ij} = -a_{ji}; \forall i, j \in 1, 2, 3, \dots, n,$$

ou seja, os elementos simetricamente dispostos em relação a diagonal principal são opostos.

Exemplo 2.8.: As matrizes abaixo são anti-simétricas.

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix}$$

Definição 2.1.10 (Matriz Inversível): Uma matriz quadrada A é chamada de inversível ou não singular, se existe uma matriz B tal que $AB = BA = I$, onde I é a matriz identidade. Se A não for inversível, dizemos que A é uma matriz singular. Chamaremos a tal matriz B de inversa de A e a denotaremos por A^{-1} .

Teorema 2.1.2 Se A é inversível, então a matriz A^{-1} tal que $AA^{-1} = A^{-1}A = I$ é única.

Demonstração: Suponhamos que existam as matrizes $A_1^{-1}A_2^{-1}$ e tais que:

$$AA_1^{-1} = A_1^{-1}A = I$$

$$AA_2^{-1} = A_2^{-1}A = I$$

então,

$$A_1^{-1} = A_1^{-1}I = A_1^{-1}(AA_2^{-1}) = (A_1^{-1}A)A_2^{-1} = IA_2^{-1} = A_2^{-1}$$

Portanto, a inversa de uma matriz é única

Matrizes diagonalmente dominante: é uma classe muito importante de matrizes em sistemas de equações, pois permitem realizar, por exemplo, a eliminação gaussiana sem a

necessidade do pivoteamento. Definiremos estas matrizes a seguir:

Definição 2.1.11 Uma matriz A é dita **diagonalmente dominante** se seus elementos verificam a condição abaixo:

$$|a_{ii}| \geq \sum_{j=1}^n |a_{ij}|, i = 1, 2, \dots, n$$

e estritamente diagonalmente dominante se

$$|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}|, i = 1, 2, \dots, n$$

isto é, em cada linha, o elemento da diagonal é maior em módulo que a sua soma dos módulos dos outros elementos.

Definição 2.1.12 (Matriz simétrica e positiva definida): uma matriz A é dita simétrica e positiva definida se atende a condição abaixo:

$$A = A^T \text{ e } X^T A X > 0, \forall \text{ vetor } x \neq 0$$

Exemplo 2.11. Seja a matriz $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ e o vetor $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq 0$. temos que:

$$X^T A X = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 & x_1 + 2x_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= 2x_1^2 + x_1x_2 + x_2x_1 + 2x_2^2 > 0$$

$$= 2x_1^2 + 2x_2x_1 + 2x_2^2 = 2x_1^2 + x_2^2 + x_1^2 + x_2^2 > 0$$

$$= (x_1 + x_2)^2 + x_1^2 + x_2^2 > 0$$

Portanto, a matriz A é positiva definida.

2.2 Sistemas Lineares

Em vários ramos da matemática nos deparamos com problemas cuja a solução em algum estágio envolve a resolução de um sistema linear, por isso, será o nosso objeto de estudo seguinte.

No caso geral em que o sistema linear envolve m equações com n incógnitas, o sistema pode apresentar uma única solução, infinita soluções ou não admitir nenhuma solução. Este tipo de problema é tratado na Álgebra Linear usando o processo de escalonamento. Iremos analisar esquemas numéricos para soluções de sistemas lineares de n equações com n incógnitas, ou seja.

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

Onde a_{ij} são os coeficientes, x_j são as incógnitas e os b_j são as constantes. Este sistema pode ser escrito na forma matricial $Ax = b$ com $A \in M_n(\mathfrak{R})$ e $x, b \in m_{n \times 1}(\mathfrak{R})$.

Definição 2.2.1 (Equação Linear) Dados os números reais $a_{11}, a_{12}, \dots, a_{1n}$ ($n \geq 1$), chamamos de equação linear nas incógnitas x_1, x_2, \dots, x_n , toda equação do tipo:

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b$$

Os números $a_{11}, a_{12}, \dots, a_{1n}$, são chamados de coeficientes e b , é o termo independente da equação.

2.2.1 Solução de um Sistema Linear

Dizemos que (x_1, x_2, \dots, x_n) , em que $x_i \in \mathfrak{R}$, é uma solução do sistema S , se satisfaz todas as equações de S . O conjunto de todas as soluções de um sistema linear é denominado conjunto solução.

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \text{ (sentença verdadeira)}$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \text{ (sentença verdadeira)}$$

\vdots

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_n \text{ (sentença verdadeira)}$$

Exemplo 2.15

$$\text{Dado o sistema } S_1 : \begin{cases} 2x + 3y = 4 \\ x - y = 2 \end{cases}$$

o par ordenado $(2, 0)$ é a solução desse sistema, pois,

$$\begin{cases} 2 \cdot 2 + 3 \cdot 0 = 4 \\ 2 - 0 = 2 \end{cases}, \text{ são sentenças verdadeiras}$$

Definição 2.2.2 (Sistema Possível) Se um sistema linear S tiver pelo menos uma solução, chamaremos de sistema possível.

Definição 2.2.3 (Sistema Impossível) Se um sistema linear S não tiver nenhuma solução, chamaremos de impossível.

Definição 2.2.4 (Sistema Possível e Indeterminado): é um sistema linear S que possui infinitas soluções.

Definição 2.2.5 (Sistema Linear Homogêneo) Denominamos de sistema linear homogêneo em que o termo independente de todas as equações é igual a zero, isto é, quando todas as equações são homogêneas. Um sistema homogêneo é sempre possível, pois possui, ao menos a solução nula. Se o sistema só possui a solução nula, ele é possível e determinado. Havendo outras soluções, além da solução nula, ele é possível e indeterminado. Essas soluções recebem o nome de soluções próprias ou não triviais.

Definição 2.2.6 (Matrizes de um Sistema) Dado um sistema linear S de m equações e n incógnitas, consideremos as matrizes:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

e

$$B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

Chamamos A de matriz incompleta do sistema e B de matriz completa.

Observemos que B foi obtida a partir de A , acrescentando-se a esta a coluna formada pelos termos independentes das equações do sistema.

2.2.2 Sistemas Equivalentes

Definição 2.2.7 Dizemos que dois sistemas lineares S_1 e S_2 , são equivalentes quando toda solução de S_1 é também solução de S_2 e vice-versa.

Exemplo 2.18.

$$S_1 \begin{cases} 2x - y = 3 \\ x + 3y = 5 \end{cases}$$

Cuja solução é dada pelo par ordenado $(2, 1)$.

$$S_2 \begin{cases} 2x - y = 3 \\ 3x + 2y = 8 \end{cases}$$

é fácil notar que $(2, 1)$ é solução de S_2 , pois a segunda equação também é verificada

$$3 \cdot 2 + 2 \cdot 1 = 6 + 2 = 8$$

Capítulo 3

Normas

3.1 Definições

Definição 3.1.1 Se V é um espaço vetorial, uma norma é uma função $\| \cdot \| : V \rightarrow \mathfrak{R}$ que verifica as propriedades a baixo:

i) $\| x \| > 0$, se $x \neq 0$, $x \in V$; $\| x \| = 0 \Leftrightarrow x = 0$

ii) $\| \alpha x \| = |\alpha| \cdot \| x \|$, $\forall \alpha \in \mathfrak{R}, \forall x \in V$

iii) $\| x + y \| \leq \| x \| + \| y \| \forall x, y \in V$

3.1.1 Normas em \mathfrak{R}^n

De especial interesse, quando $V = \mathfrak{R}^n$, são as normas l_p definidas como:

$$\| x \|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad p \geq 1$$

$$\| x \|_\infty = \max_{1 \leq i \leq n} \{|x_i|\}$$

Propriedade 3.1 A função $\| \cdot \|_1 : \mathfrak{R}^n \rightarrow \mathfrak{R}$ definida por:

$$\| x \|_1 = |x_1| + |x_2| + \dots + |x_n| = \sum_{i=1}^n |x_i|,$$

é uma norma em \mathfrak{R}^n , chamada de norma l_1 ou norma da soma.

demonstração: considere $x, y \in \mathfrak{R}^n$ e $\alpha \in \mathfrak{R}$.

i) Se $x = (x_1, x_2, \dots, x_n) \neq 0$, então $\exists x_k \neq 0 \implies |x_k| > 0$. Logo,

$$\| x \|_1 = \sum_{i=1}^n |x_i| \geq |x_k| > 0$$

ii) $\| \alpha x \|_1 = \| \alpha(x_1, x_2, \dots, x_n) \|_1 = \| (\alpha x_1, \alpha x_2, \dots, \alpha x_n) \|_1$

$$= \sum_{i=1}^n |\alpha x_i| = \sum_{i=1}^n |\alpha| |x_i| = |\alpha| \sum_{i=1}^n |x_i| = |\alpha| \|x\|_1$$

$$\text{iii) } \|x + y\|_1 = \|(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n)\|_1$$

$$= \|(x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)\|_1 = \sum_{i=1}^n |x_i + y_i| \leq \sum_{i=1}^n (|x_i| + |y_i|) \leq \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i|$$

$$\text{Daí, } \|x + y\|_1 \leq \|x\|_1 + \|y\|_1$$

De i), ii) e iii) e da Definição 3.1 temos que $\|\cdot\|_1$ é uma norma em \mathfrak{R}^n

Propriedade 3.2. A função $\|\cdot\|_2: \mathfrak{R}^n \rightarrow \mathfrak{R}$ definida por:

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}$$

é uma norma em \mathfrak{R}^n , denominada Norma l_2 ou Norma Euclidiana.

Demonstração: Considere $x, y \in \mathfrak{R}^n$ e $\alpha \in \mathfrak{R}$.

i) Se $x = (x_1, x_2, \dots, x_n) \neq 0$, então, $\exists i \in 1, 2, \dots, n$ tal que $x_i \neq 0$. logo.

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + \underbrace{x_i^2}_{\neq 0} + \dots + x_n^2} \geq \sqrt{x_1^2} = |x_1| > 0$$

$$\text{ii) } \|\alpha\|_2 = \|\alpha(x_1, x_2, \dots, x_n)\|_2 = \|(\alpha x_1, \alpha x_2, \dots, \alpha x_n)\|_2$$

$$= \sqrt{(\alpha x_1)^2 + (\alpha x_2)^2 + \dots + (\alpha x_n)^2}$$

$$= \sqrt{\alpha^2(x_1^2 + x_2^2 + \dots + x_n^2)}$$

$$= \sqrt{\alpha^2} \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = |\alpha| \|x\|_2$$

$$\text{Portanto, } \|\alpha x\|_2 = |\alpha| \|x\|_2$$

$$\text{iii) } \|x + y\|_2^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle$$

$$= \|x\|_2^2 + 2\langle x, y \rangle + \|y\|_2^2 \leq \|x\|_2^2 + 2\|x\|_2 \|y\|_2 + \|y\|_2^2$$

(Desigualdade de Cauchy-Schwarz)

$$\|x + y\|_2^2 \leq (\|x\|_2 + \|y\|_2)^2$$

$$\therefore \|x + y\|_2 \leq \|x\|_2 + \|y\|_2$$

De i), ii) e iii) e da definição 3.1 temos que $\| \cdot \|_2$ é uma norma em \mathfrak{R}^n .

Propriedade 3.3. A função $\| \cdot \|_\alpha : \mathfrak{R}^n \rightarrow \mathfrak{R}$ definida por:

$$\| x \|_\alpha = \max_{1 \leq i \leq n} |x_i|,$$

é uma norma em \mathfrak{R}^n , denominada Norma l_∞ ou Norma do Máximo.

Demonstração: considere $x, y \in \mathfrak{R}^n$ e $\alpha \in \mathfrak{R}$

i) Se $x = (x_1, x_2, \dots, x_n) \neq 0$, então $\exists x_k \neq 0 \Rightarrow |x_k| > 0$ logo,

$$\| x \|_\infty = \max_{1 \leq i \leq n} |x_i| \{ |x_1| + |x_2| + \dots + \underbrace{|x_k|}_{\neq 0} + \dots + |x_n| \} \geq |x_k| > 0.$$

$$\text{ii) } \| \alpha x \|_\infty = \| \alpha(x_1, x_2, \dots, x_n) \|_\infty = \| (\alpha x_1, \alpha x_2, \dots, \alpha x_n) \|_\infty$$

$$= \max_{1 \leq i \leq n} \{ |\alpha x_i| \} = \max_{1 \leq i \leq n} \{ |\alpha| |x_i| \}$$

$$= |\alpha| \max_{1 \leq i \leq n} \{ |x_i| \} = |\alpha| \| x \|_\infty$$

$$\text{iii) } \| x + y \|_\infty = \| (x_1 + x_2 + \dots + x_n) + (y_1 + y_2 + \dots + y_n) \|_\infty$$

$$= \| x_1 + y_1, x_2 + y_2, + \dots + x_n + y_n \|_\infty$$

$$\leq \max_{1 \leq i \leq n} \{ |x_i| \} + \max_{1 \leq i \leq n} \{ |y_i| \}$$

$$\leq \| x \|_\infty + \| y \|_\infty$$

logo, $\| x + y \|_\infty \leq \| x \|_\infty + \| y \|_\infty$

De i), ii) e iii) e da definição 3.1 temos que $\| \cdot \|_\infty$ é uma norma em \mathfrak{R}^n .

Exemplo 3.1. Para o vetor $x = (3, 3, 3, -3)$ temos:

$$\text{i) } \| x \|_2 = \sqrt{(3)^2 + (3)^2 + (3)^2 + (-3)^2} = \sqrt{36} = 6$$

$$\text{ii) } \| x \|_\infty = \max_{1 \leq i \leq 3} \{ |x_i| \} = \max_{1 \leq i \leq 3} \{ |3|, |3|, |3|, |-3| \} = 3$$

$$\text{iii) } \| x \|_1 = \sum_{i=1}^3 \| x \|_1 = |3| + |3| + |3| + |-3| = 12$$

3.2 Normas Matriciais

Similar às normas em \mathfrak{R}^n , poderia-se procurar funções que verifiquem as condições i), ii) e iii) da definição 3.1.1 de modo a gerar normas para matrizes. Entretanto, se prefere definir a norma de uma matriz quadrada de ordem n associada à norma de um vetor em \mathfrak{R}^n , como veremos na seguinte definição

Definição 3.2.1 Seja $\| \cdot \|$ uma norma em \mathfrak{R}^n , e A uma matriz de ordem n . Definimos a norma de uma matriz A , denotada por $\| A \|$ como:

$$\| A \| = \sup \| Au \| : u \in \mathfrak{R}^n, \| u \| = 1$$

A seguinte propriedade verifica que $\| A \|$ definida por 3.1. é uma norma no espaço das matrizes.

Propriedade 3.4. Seja $\| \cdot \|$ uma norma em \mathfrak{R}^n , então:

$$\| A \| = \sup \| Au \| : u \in \mathfrak{R}^n, \| u \| = 1$$

define uma norma no espaço vetorial das matrizes quadradas de ordem n , isto é:

- i) $\| A \| > 0, \forall A \neq 0, A \in M_n(\mathfrak{R})$
- ii) $\| \alpha A \| = |\alpha| \cdot \| A \|, \forall \alpha \in \mathfrak{R}, A \in M_n(\mathfrak{R})$
- iii) $\| A + B \| \leq \| A \| + \| B \|, \forall A, B \in M_n(\mathfrak{R})$

Demonstração: Sejam $A, B \in M_n$ e $\alpha \in \mathfrak{R}$

i) Se $A \neq 0$, então A tem pelo menos uma coluna distinta de zero. Digamos que $A^j \neq 0$.

Considere um vetor que 1 esteja na j -ésima componente, isto é, $x = (0, \dots, 0, 1, 0, \dots, 0)^T$.

Note que $x \neq 0$ e que o vetor $v = \frac{x}{\| x \|}$ tem norma 1. Além disso,

$$Ax = A^j \text{ e } \| Ax \| = \| A^j \| \neq 0. \text{ Segue, pela definição de } \| A \|, \\ \| A \| = \sup \| Au \| : u \in \mathfrak{R}^n, \| u \| = 1 \geq \| Av \| = \frac{\| Ax \|}{\| x \|} = \frac{\| A^j \|}{\| x \|} > 0$$

$$\text{ii) } \| \alpha A \| = \sup \| (\alpha A)u \| : \| u \| = 1 = \sup \| \alpha(Au) \| : \| u \| = 1 \\ = \sup \| |\alpha|(Au) \| : \| u \| = 1 \\ = |\alpha| \sup \| Au \| : \| u \| = 1 = |\alpha| \cdot \| A \|$$

$$\text{iii) } \| A + B \| = \sup \| (A + B)u \| : \| u \| = 1 \\ = \sup \| Au + Bu \| : \| u \| = 1 \leq \sup \| Au \| + \| Bu \| : \| u \| = 1 \\ \leq \sup \| Au \| : \| u \| = 1 + \sup \| Bu \| : \| u \| = 1 \leq \| A \| + \| B \|$$

De i), ii) e iii) e da definição 3.1 temos que $\| A \|$ é uma norma em $M_n(\mathfrak{R})$

Como consequência da definição de norma matricial 3.1, se verificam as seguintes propriedades:

Propriedade 3.5.

- i) $\| Ax \| \leq \| A \| \cdot \| X \|, \forall A \in M_n(\mathfrak{R}), \forall x \in \mathfrak{R}^n$
- ii) $\| I \| = 1$, onde I é a matriz identidade em $M_n(\mathfrak{R})$
- iii) $\| AB \| \leq \| A \| \cdot \| B \|, \forall A, B \in M_n(\mathfrak{R})$

Demonstração:

i) Se $x = 0$ então $\| x \| = 0$ e $\| Ax \| = 0$, e portanto $\| Ax \| = \| A \| \cdot \| x \|$

Se $x \neq 0$, então $\| x \| \neq 0$ e $\| Ax \| \neq 0$. Fazendo $v = \frac{\| Ax \|}{\| x \|}$, temos que $\| v \| = 1$. Logo,

$$\| A \| = \sup \{ \| Au \| : \| u \| = 1 \} \geq \| Av \| \geq \| A \| \cdot \frac{\| x \|}{\| x \|} = \| Ax \|$$

Multiplicando por $\|x\|$ resulta que $\|AX\| \leq \|A\| \cdot \|X\|$

ii) $\|I\| = \sup\{\|Iu\| : \|u\| = 1\} = \sup\{\|A(Bu)\| : \|u\| = 1\} = \sup\{1\} = 1$

iii) $\|AB\| = \sup\{\|(AB)u\| : \|u\| = 1\} = \sup\{\|A(Bu)\| : \|u\| = 1\}$

$\leq \sup\{\|A\| \cdot \|Bu\| : \|u\| = 1\} \leq \|A\| \cdot \sup\{\|Bu\| : \|u\| = 1\} \leq \|A\| \cdot \|A\|$, o que completa a prova.

Os seguintes teoremas determinam uma forma prática de calcular as normas matriciais associadas às normas vetoriais l_∞ e l_1 .

Teorema 3.2.1 *Se a norma vetorial $\|\cdot\|_\infty$ é definida por $\|x\|_\infty = \max_{1 \leq i \leq 3} \{|x_i|\}$, então a norma matricial associada é calculada por:*

$$\|A\|_\infty = \max_{1 \leq i \leq 3} \sum_{j=1}^n |a_{ij}|$$

Demonstração: Sabemos que a norma de A é definida como:

$$\|A\| = \sup\{\|Au\| : u \in \mathfrak{R}^n, \|u\| = 1\}$$

vamos supor que N representa o valor máximo. Então temos que $\|Au\| \leq N$, para todo vetor unitário u.

Vamos mostrar que $\|Au\|_\infty = \max_{1 \leq i} \{|u_1|, |u_2|, \dots, |u_n|\} = 1$

Seja a matriz $A_{n \times n}$ e o vetor u:

$$Au = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix}^T = \begin{pmatrix} a_{11}u_1 & a_{12}u_2 & \cdots & a_{1n}u_n \\ a_{21}u_1 & a_{22}u_2 & \cdots & a_{2n}u_n \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1}u_1 & a_{n2}u_2 & \cdots & a_{nn}u_n \end{pmatrix}$$

Tomando a norma do máximo, temos:

$$\|A\|_\infty = \max_{1 \leq i} \{|a_{11}u_1 + a_{12}u_2 + \dots + a_{1n}u_n|, |a_{21}u_1 + a_{22}u_2 + \dots + a_{2n}u_n|, \dots, |a_{n1}u_1 + a_{n2}u_2 + \dots + a_{nn}u_n|\}$$

Usando a Desigualdade Triangular:

$$\begin{aligned} &\leq \max_{1 \leq i} \{|a_{11}u_1| + |a_{12}u_2| + \dots + |a_{1n}u_n|, |a_{21}u_1| + |a_{22}u_2| + \dots + |a_{2n}u_n|, \dots, |a_{n1}u_1| + |a_{n2}u_2| + \dots + |a_{nn}u_n|\} \\ &\leq \max_{1 \leq i \leq n} \{|a_{11}||u_1| + |a_{12}||u_2| + \dots + |a_{1n}||u_n|, |a_{21}||u_1| + |a_{22}||u_2| + \dots + |a_{2n}||u_n|, \dots, |a_{n1}||u_1| + |a_{n2}||u_2| + \dots + |a_{nn}||u_n|\} \\ &= \max_{1 \leq i \leq n} \{|A_1||u| + |A_2||u| + \dots + |A_n||u|\} = \max_{1 \leq i \leq n} \{|A_1| + |A_2| + \dots + |A_n|\} = N \end{aligned}$$

Se o valor máximo ocorre na linha k, então $u = e_k$, onde e_k é um vetor unitário do \mathfrak{R}^n .

Logo,

$$\|Ae_k\|_\infty = \|A_k\|_\infty = N$$

Portanto,

$$\| A \| = \sup\{\| Au \| : u \in \mathfrak{R}^n, \| u \| = 1\} = \| A \|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$$

o que completa a prova.

Teorema 3.2.2 . Se a norma vetorial $\| \cdot \|_1$ se $\| x \|_1 = \sum_{i=1}^n |x_i|$, então a norma matricial associada é definida por:

$$\| A \|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$$

Demonstração: Sabendo que a norma de A é definida como:

$$\| A \| = \sup\{\| Au \| : u \in \mathfrak{R}^n, \| u \| = 1\}$$

Vamos supor que M representa o valor máximo. Então temos $\| Ax \|_1 \leq M$, para qualquer vetor unitário $\| x \|$. Achar \times , usando a norma da soma para validar a igualdade. $M = \max\{\| A \|_1\}$ o valor máximo da soma das colunas é $\| x \| = 1$

$$Ax = \begin{pmatrix} a_{11}x_1 & a_{12}x_2 & \cdots & a_{1n}x_n \\ a_{21}x_1 & a_{22}x_2 & \cdots & a_{2n}x_n \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1}x_1 & a_{n2}x_2 & \cdots & a_{nn}x_n \end{pmatrix}$$

$$\| A \|_1 = |a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n|, |a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n|, \dots, |a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n|$$

Utilizando a propriedade de módulos, temos:

$$\| Ax \|_1 \leq (|a_{11}| + |a_{12}| + \dots + |a_{1n}|)|x_1| + \dots + (|a_{1n}| + |a_{2n}| + \dots + |a_{nn}|)|x_n|$$

$$\| Ax \|_1 \leq |A_1||x_1| + |A_2||x_2| + \dots + |A_n||x_n|$$

$$\| Ax \|_1 \leq M|x_1| + M|x_2| + \dots + M|x_n|$$

$$\| Ax \|_1 \leq M(|x_1| + |x_2| + \dots + |x_n|) = M \| x \|_1 = M$$

Se na coluna, ocorrer que seja o valor máximo dentre a soma dos valores absolutos das colunas, conclui-se que, para $x = e_k$, temos:

$$\| Ax \|_1 = \| Ae_k \|_1 = \| A_k \|_1 = M, \text{ portanto,}$$

$$\| Ax \|_1 = M = \max_{1 \leq j \leq n} \{\| A_j \|_1\} = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$$

o que completa a prova.

Exemplo 3.2. Para a matriz $A = \begin{pmatrix} 2 & -3 & 1 \\ 5 & 2 & 4 \\ -1 & 0 & 7 \end{pmatrix}$ temos que,

i) Soma dos módulos dos elementos das matrizes em cada coluna:

$$\text{Linha 1: } |2| + |-3| + |1| = 6$$

$$\text{Linha 2: } |5| + |2| + |4| = 11$$

$$\text{linha 3: } |-1| + |0| + |7| = 8$$

$$\text{Logo, } \|A\|_{\infty} = \max\{6, 11, 8\} = 11$$

ii) Soma dos módulos dos elementos da matriz em cada coluna:

$$\text{coluna 1: } |2| + |5| + |-1| = 8$$

$$\text{coluna 2: } |-3| + |2| + |0| = 5$$

$$\text{coluna 3: } |1| + |4| + |7| = 12$$

$$\text{Logo, } \|A\|_1 = \max\{8, 5, 12\} = 12$$

3.3 Convergência em Norma

Um espaço vetorial normado $(V, \|\cdot\|)$ é um espaço vetorial V com norma $\|\cdot\|$. O conceito de norma nos permite definir o conceito de convergência sobre espaços vetoriais.

Definição 3.3.1 A sequência $x^{(1)}, x^{(2)}, \dots$ em $(V, \|\cdot\|)$ converge ao vetor $x \in (V, \|\cdot\|)$, escreveremos $x^{(k)} \rightarrow x$, se

$$\lim_{k \rightarrow \infty} \|x^{(k)} - x\| = 0$$

A convergência acima coincide com a ideia intuitiva que a distância entre os vetores $x^{(k)}$ e o vetor limite x se aproxima de zero quando k aumenta.

Em espaços de dimensão finita, como é o caso de \mathfrak{R}^n e $M_n(\mathfrak{R})$, se uma sequência converge em uma norma convergente em qualquer outra norma. Entretanto, esta afirmação não se verifica em espaços de dimensão infinita.

Exemplo 3.3. Se $x^{(k)} = \begin{pmatrix} 2 - k^{-1} \\ -5 + k^{-2} \\ 3 + e^{-k} \end{pmatrix}$ e $x = \begin{pmatrix} 2 \\ -5 \\ 3 \end{pmatrix}$, então $x^{(k)} - x = \begin{pmatrix} -k^{-1} \\ k^{-2} \\ e^{-k} \end{pmatrix}$

Em \mathfrak{R}^3 ,

$$\lim_{k \rightarrow \infty} \|x^{(k)} - x\|_{\infty} = \lim_{k \rightarrow \infty} k^{-1} = 0$$

Como \mathfrak{R}^4 é um espaço de dimensão finita, verifica-se imediatamente que

$$\lim_{k \rightarrow \infty} \|x^{(k)} - x\|_1 = 0 \quad \lim_{k \rightarrow \infty} \|x^{(k)} - x\|_2 = 0$$

Em espaços vetoriais normados de dimensão finita toda sequência $\{x^{(k)}, k = 1, 2, \dots\}$ que verifica o critério de Cauchy:

$$\lim_{k \rightarrow \infty} \sup_{i, j \geq 1} \|x^{(i)} - x^{(j)}\|_1 = 0,$$

é convergente, isto é, existe $x \in V$ tal que $x^{(k)} \rightarrow x$ em norma.

Se $x^{(k)}$ é a k -ésima aproximação num método iterativo que resolve a equação $Ax = b$, então pode-se definir um critério de parada de acordo com uma tolerância $\epsilon > 0$ predefinida. Os critérios de parada mais usuais são:

i) Erro Absoluto: $\|x^{k+1} - x^k\| \leq \epsilon$

ii) Erro Relativo: $\frac{\|x^{k+1} - x^k\|}{\|x^k\|} \leq \epsilon$

iii) Residual: $\|b - Ax^k\| \leq \epsilon$

Capítulo 4

Métodos Diretos

Pertencem a essa classe todos os métodos estudados nos cursos de 1° e 2° graus, destacando-se a regra de Cramer. Este método, aplicado à resolução de um sistema $n \times n$ envolve o cálculo de $(n + 1)$ determinantes de ordem n . Se n for igual a 20 podemos mostrar que o número total de operações efetuadas, será $21 \times 20! \times 19$ multiplicações mais um número semelhante de adições. Assim, um computador que efetue cerca de cem milhões de multiplicações por segundo levaria 3×10^5 anos para efetuar as operações necessárias.[7] Desta forma, o estudo de métodos mais eficientes é necessário, pois, em geral, os problemas práticos exigem a resolução de sistemas lineares de grande porte, isto é, sistemas que envolvem um grande número de equações e variáveis.

Devemos observar que no caso de sistemas lineares $n \times n$, da forma $Ax = b$ com solução única, o vetor x^* é dada por: $x^* = A^{-1}b$. Onde x^* é o vetor solução do sistema acima. No entanto, calcular explicitamente a matriz A^{-1} e em seguida efetuar o produto $A^{-1}b$ é desaconselhável, uma vez que o número de operações envolvidas é grande, o que torna este processo não competitivo com os métodos que estudaremos a seguir.

4.0.1 Métodos da Eliminação de Gauss

Entre os métodos diretos, destacam-se os métodos de eliminação que evitam o cálculo direto da matriz inversa de A e além disto não apresentam problemas com tempo de execução como a regra de Cramer.

O método da Eliminação de Gauss consiste em transformar o sistema linear original num sistema linear equivalente com a matriz dos coeficientes triangular superior, pois, estes são de resolução imediata. Dizemos que dois sistemas lineares são equivalentes quando possuem a mesma solução.

Veremos a seguir um algoritmo para a resolução de sistemas triangulares e estudaremos como o método da Eliminação de Gauss efetua a transformação do sistema linear original no sistema triangular equivalente.

Resolução de Sistemas Triangulares

Seja o sistema linear $Ax = b$, onde A : matriz $n \times n$, Triangular superior, com elementos da diagonal diferentes de zero. Escrevendo as equações deste sistema, temos:

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \quad + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad \quad \quad \ddots \quad \quad \quad \vdots \\ \quad \quad \quad \quad \quad \quad a_{nn}x_n = b_n \end{array} \right.$$

Da última equação, temos:

$$x_n = \frac{b_n}{a_{nn}}$$

x_{n-1} pode então ser obtido da penúltima equação:

$$x_{n-1} = \frac{b_{n-1} - a_{n-1,n}x_n}{a_{n-1,n-1}}$$

e assim sucessivamente obtém-se x_{n-2}, \dots, x_2 e finalmente x_1 :

$$x_1 = \frac{b_1 - a_{12}x_2 - a_{13}x_3 - \dots - a_{1n}x_n}{a_{11}}$$

Descrição do método da Eliminação de Gauss

Como dissemos anteriormente, o método consiste em transformar convenientemente o sistema linear original para obter um sistema linear equivalente com matriz dos coeficientes triangular superior.

Para modificar convenientemente o sistema linear dado de forma a obter um sistema linear equivalente, faremos uso do teorema a seguir:

Teorema 4.0.1 *Seja $Ax = b$ um sistema linear. Aplicado sobre as equações deste sistema uma sequência de operações elementares escolhidas entre:*

- i) Trocar duas equações;
- ii) Multiplicar uma equação por uma constante não nula;
- iii) Adicionar um múltiplo de uma equação a uma outra equação;

obtemos um novo sistema $\tilde{A}x = \tilde{b}$ e os sistemas $Ax = b$ são equivalentes.

Descreveremos a seguir como o método da Eliminação de Gauss usa este teorema para triangularizar a matriz A . Vamos supor que $\det(A) \neq 0$.

A eliminação é efetuada por colunas e chamaremos de etapa k do processo a fase em que se elimina a variável x^k das equações $k + 1, k + 2, \dots, n$.

Usaremos a notação a_{ij}^k para denotar o coeficiente da linha i e coluna j no final da k -ésima etapa, bem como b_i^k será o i -ésimo elemento do vetor constante no final da etapa k .

Considerando que $\det(A) \neq 0$, é sempre possível reescrever o sistema linear de forma que o elemento da posição a_{11} seja diferente de zero, usando apenas a operação elementar:

$$\text{Seja } A^{(0)}|b^{(0)} = A|b = \left(\begin{array}{cccc|c} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} & b_1^{(0)} \\ a_{21}^{(0)} & a_{22}^{(0)} & \cdots & a_{2n}^{(0)} & b_2^{(0)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1}^{(0)} & a_{n2} & \cdots & a_{nm}^{(0)} & b_n^{(0)} \end{array} \right)$$

onde, $a_{ij}^{(0)} = a_{ij}$, $b_i^{(0)} = b_i$ e $a_{11} \neq 0$.

Etapa 1: A eliminação da variável x_1 das equações $i = 2, \dots, n$ é feita da seguinte forma: da equação i subtraímos a 1ª equação multiplicada por m_{i1} . Observamos que para que esta eliminação seja efetuada, a única escolha é $m_{i1} = \frac{a_{i1}^{(0)}}{a_{11}^{(0)}}$, $i = 2, \dots, n$.

Os elementos $m_{i1} = m_{i1} = \frac{a_{i1}^{(0)}}{a_{11}^{(0)}}$, $i = 2, \dots, n$ são os multiplicadores e o elemento $a_{11}^{(0)}$ é denominado pivô da 1ª etapa.

Ao final desta etapa teremos a matriz:

$$A^{(1)}|b^{(1)} = \left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{array} \right)$$

Onde,

$$a_{1j}^{(1)} = a_{1j}^{(0)} \text{ para } j = 1, \dots, n$$

$$b_1^{(1)} = b_1^{(0)} \text{ e}$$

$$a_{ij}^{(1)} = a_{ij}^{(0)} - m_{i1}a_{1j}^{(0)} \text{ } i = 2, \dots, n \text{ e } j = 1, \dots, n$$

$$b_i^{(1)} = b_i^{(0)} - m_{i1}b_1^{(0)} \text{ } i = 2, \dots, n$$

Etapa 2:

Deve-se ter pelo menos um elemento $a_{i2}^{(1)} \neq 0$, para $i = 2, \dots, n$, caso contrário $\det(A^{(1)}) = 0$, o que implica que $\det(A) = 0$; mas $\det(A) \neq 0$, por hipótese.

Então, é sempre possível reescrever a matriz $A^{(0)}$, sem alterar a posição da linha 1, de forma que o pivô, $a_{22}^{(1)}$, seja não nulo.

Os multiplicadores desta etapa serão os elementos $m_{i2} = \frac{a_{i2}^{(1)}}{a_{22}^{(1)}}$ para $i = 3, \dots, n$.

A variável x_2 é eliminada das equações $i = 3, \dots, n$ da seguinte forma: da equação i subtraímos a segunda equação multiplicada por m_{i2} .

Ao final, teremos a matriz $A^{(2)}|b^{(2)}$;

$$A^{(2)}|b^{(2)} = \left(\begin{array}{cccc|c} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} & \cdots & a_{1n}^{(2)} & | & b_1^{(2)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & | & b_2^{(2)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} & | & b_3^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} & | & b_n^{(2)} \end{array} \right)$$

Onde, $a_{ij}^{(2)} = a_{ij}^{(1)}$ para $i = 1, 2$ e $j = i, i+1, \dots, n$

$b_i(2) = b_i^{(1)}$ para $i = 1, 2$ e

$a_{ij}^{(2)} = a_{ij}^{(1)} - m_{i2}a_{2j}^{(1)}$ $i = 3, \dots, n$ e $j = 2, \dots, n$

$b_i(2) = b_i^{(1)} - m_{i2}b_2(1)$ $i = 3, \dots, n$

Seguindo o raciocínio análogo, procede-se até a etapa $(n - 1)$ e a matriz, ao final desta etapa, será:

$$A^{(2)}|b^{(2)} = \left(\begin{array}{cccc|c} a_{11}^{(n-1)} & a_{12}^{(n-1)} & a_{13}^{(n-1)} & \cdots & a_{1n}^{(n-1)} & | & b_1^{(n-1)} \\ 0 & a_{22}^{(n-1)} & a_{23}^{(n-1)} & \cdots & a_{2n}^{(n-1)} & | & b_2^{(n-1)} \\ 0 & 0 & a_{33}^{(n-1)} & \cdots & a_{3n}^{(n-1)} & | & b_3^{(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n-1)} & | & b_n^{(n-1)} \end{array} \right)$$

e o sistema linear $A^{(n-1)}x = b^{(n-1)}$ é triangular superior e equivalente ao sistema linear original.

Exemplo: Seja o sistema linear:

$$\begin{cases} 3x_1 + 2x_2 + 4x_3 = 1 \\ x_1 + x_2 + 2x_3 = 2 \\ 4x_1 + 3x_2 - 2x_3 = 3 \end{cases}$$

Etapa 1: Eliminar x_1 das equações 2 e 3:

Para facilitar o entendimento do processo, de agora em diante usaremos a notação L_i para indicar o vetor linha formado pelos elementos da linha i da matriz $A^{(k)}|b^{(k)}$. Assim, nessa

etapa, $L_1 = (3 \ 2 \ 4 \ 1)$.

$$A^{(0)}|b^{(0)} = \left(\begin{array}{ccc|c} a_{11}^{(0)} & a_{12}^{(0)} & a_{13}^{(0)} & | & b_1^{(0)} \\ a_{21}^{(0)} & a_{22}^{(0)} & a_{23}^{(0)} & | & b_2^{(0)} \\ a_{31}^{(0)} & a_{32}^{(0)} & a_{33}^{(0)} & | & b_3^{(0)} \end{array} \right) = \left(\begin{array}{ccc|c} 3 & 2 & 4 & | & 1 \\ 1 & 1 & 2 & | & 2 \\ 4 & 3 & -2 & | & 3 \end{array} \right)$$

Pivô: $a_{11}^{(0)} = 3$

$m_{21} = 1/3$

$m_{31} = 4/3$

$L_2 \leftarrow L_2 - m_{21}L_1$

$L_3 \leftarrow L_3 - m_{31}L_1$

$$\Rightarrow A^{(1)}|b^{(1)} = \left(\begin{array}{ccc|c} 3 & 2 & 4 & 1 \\ 0 & 1/3 & 2/3 & 5/3 \\ 0 & 1/3 & -22/3 & 5/3 \end{array} \right) = \left(\begin{array}{ccc|c} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & b_2^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & b_3^{(1)} \end{array} \right)$$

Etapa 2: Eliminar x_2 da equação 3:

$$\text{Pivô: } a_{22}^{(1)} = 1/3$$

$$m_{32} = \frac{1/3}{1/3} = 1$$

$$L_3 \leftarrow L_3 - m_{32}L_2$$

$$\Rightarrow A^{(2)}|b^{(2)} = \left(\begin{array}{ccc|c} 3 & 2 & 4 & 1 \\ 0 & 1/3 & 2/3 & 5/3 \\ 0 & 0 & -8 & 0 \end{array} \right)$$

Assim, resolver $Ax = b$ é equivalente a resolver $A^{(2)}x = b^{(2)}$

$$\left\{ \begin{array}{l} 3x_1 + 2x_2 + 4x_3 = 1 \\ + 1/3x_2 + 2/3x_3 = 5/3 \\ - 8x_3 = 0 \end{array} \right.$$

A solução desse sistema é o vetor $x^* = \begin{pmatrix} -3 \\ 5 \\ 0 \end{pmatrix}$

Estratégias de Pivoteamento Vimos que o algoritmo para o método de Eliminação de Gauss requer o cálculo dos multiplicadores:

$$m_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} \quad i = k + 1, \dots, n$$

em cada etapa k do processo.

O que acontece se o pivô for nulo? E se o pivô tiver próximo de zero?

Este dois casos merecem atenção especial, pois, é impossível trabalhar com um pivô nulo. E trabalhar com um pivô próximo de zero pode conduzir a resultados totalmente imprecisos. Isto porque em qualquer calculadora ou computador os cálculos são efetuados com aritmética de precisão finita, e pivôs próximos de zero dão origem a multiplicadores bem maiores que a unidade que, por sua vez, origina uma aplicação dos erros de arredondamento.

Para se contornar esses problemas deve-se adotar uma estratégia de pivoteamento, ou seja, adotar um processo de escolha da linha e/ou coluna pivotal.

Estratégias de Pivoteamento Parcial

Essa estratégia consiste em:

i) no início da etapa da fase de eliminação, escolher para pivô o elemento de maior módulo

entre os coeficientes $a_{ik}^{(k-1)}$, $i = k, k + 1, \dots, n$.

ii) trocar as linhas k e i se necessário.

Exemplo: $n = 4, k = 2$

$$A^{(1)}|b^{(1)} = \left(\begin{array}{cccc|c} 3 & 2 & 1 & -1 & 5 \\ 0 & 1 & 0 & 3 & 6 \\ 0 & -3 & -5 & 7 & 7 \\ 0 & 2 & 4 & 0 & 15 \end{array} \right)$$

Início da etapa 2:

Escolher pivô $\max_{j=2,3,4} |a_{j2}^{(1)}| = |a_{32}^{(1)}| = 3 \implies \text{pivô} = -3$

ii) trocar linhas 2 e 3.

$$A^{(1)}|b^{(1)} = \left(\begin{array}{cccc|c} 3 & 2 & 1 & -1 & 5 \\ 0 & -3 & -5 & 7 & 7 \\ 0 & 1 & 0 & 3 & 6 \\ 0 & 2 & 4 & 0 & 15 \end{array} \right)$$

e os multiplicadores dessa etapa serão:

$$m_{32} = \frac{1}{-3} = -1/3$$

$$m_{42} = \frac{2}{-3} = -2/3$$

Observamos que a escolha do maior elemento em módulo entre os candidatos a pivô faz com que os multiplicadores, em módulo, estejam entre zero e um, o que evita a aplicação dos erros de arredondamento.

Estratégias de Pivoteamento Completo Nesta estratégia, no início da etapa k é escolhido pra pivô o elemento de maior módulo. entre todos os elementos que ainda atuam no processo de eliminação:

Escolher pivô $\max_{\forall i,j \geq k} |a_{ij}^{(k-1)}| = |a_{rs}^{(k-1)}| \implies \text{pivô} = a_{rs}^{(k-1)}$

Observamos que, no exemplo anterior, se fosse adotada essa estratégia, o pivô da etapa 2 seria $a_{34}^{(1)} = 7$, o que acarretaria a troca das colunas 2 e 4 e, e em seguida, das linhas 2 e 3, donde:

$$A^{(1)}|b^{(1)} = \left(\begin{array}{cccc|c} 3 & -1 & 1 & 2 & 5 \\ 0 & 7 & -5 & -3 & 7 \\ 0 & 3 & 0 & 1 & 6 \\ 0 & 0 & 4 & 2 & 15 \end{array} \right)$$

Esta estratégia não é muito empregada, pois envolve uma comparação extensa entre os

elementos $a_{ij}(k-1)$, $i, j \geq k$ e troca de linhas e colunas, como vimos no exemplo anterior; é evidente que todo esse processo acarreta um esforço computacional maior que a estratégia de pivoteamento parcial.

Exemplo: consideremos o sistema linear

$$\begin{cases} 0,0002x_1 + 2x_2 = 5 \\ 2x_1 + 2x_2 = 6 \end{cases}$$

Inicialmente vamos resolvê-lo sem a estratégia de pivoteamento parcial e vamos supor que temos que trabalhar com aritmética de três dígitos. Nosso sistema é:

$$\begin{cases} 0,2 \times 10^{-3}x_1 + 0,2 \times 10^1x_2 = 0,5 \times 10^1 \\ 0,2 \times 10^1x_1 + 0,2 \times 10^1x_2 = 0,6 \times 10^1 \end{cases}$$

então,

$$A^{(0)}|b^{(0)} = \left(\begin{array}{cc|c} 0,2 \times 10^{-3}x_1 & 0,2 \times 10^1x_2 & 0,5 \times 10^1 \\ 0,2 \times 10^1x_1 & 0,2 \times 10^1x_2 & 0,6 \times 10^1 \end{array} \right)$$

Etapa 1: Pivô $0,2 \times 10^{-3}$

$$m_{21} = (0,2 \times 10^1)/(0,2 \times 10^{-3}) = 1 \times 10^4 = 0,1 \times 10^5 \text{ e } a_{21}^{(1)} = 0$$

$$a_{22}^{(1)} = a_{22}^{(0)} - a_{12}^{(0)} \times m_{21} = 0,2 \times 10^1 - (0,2 \times 10^1) \times (0,1 \times 10^5) = 0,2 \times 10^1 - 0,2 \times 10^6 = -0,2 \times 10^5$$

$$b_2^{(1)} = b_2^{(0)} - b_1^{(0)} \times m_{21} = 0,6 \times 10^1 - (0,5 \times 10^1) \times (0,1 \times 10^5) = 0,6 \times 10^1 - 0,5 \times 10^5 = -0,5 \times 10^5$$

$$A^{(1)}|b^{(1)} = \left(\begin{array}{cc|c} 0,2 \times 10^{-3} & 0,2 \times 10^1 & 0,5 \times 10^1 \\ 0 & -0,2 \times 10^5 & -0,5 \times 10^5 \end{array} \right)$$

E a solução do sistema $A^{(1)}x = b^{(1)}$ resultante é:

$$-0,2 \times 10^5x_2 = -0,5 \times 10^5 \Rightarrow x_2 = (0,5)/(0,2) = 2,5 = -0,25 \times 10$$

$$\Rightarrow 0,2 \times 10^{-3}x_1 + 0,2 \times 10^1 \times 0,25 \times 10^1 = 0,5 \times 10^1$$

$$\Rightarrow 0,2 \times 10^{-3}x_1 = 0,5 \times 10^1 - 0,05 \times 10^2 = 0,5 \times 10^1 - 0,5 \times 10^1 = 0$$

$$\text{e, portanto, } \bar{x} = \begin{pmatrix} 0 & 2,5 \end{pmatrix}^T$$

É fácil verificar que \bar{x} não satisfaz a segunda equação, pois $2 \times 0 + 2 \times 2,5 \neq 6$

Usando agora a estratégia de pivoteamento parcial (e ainda aritmética de três dígitos), temos:

$$A^{(0)}|b^{(0)} = \left(\begin{array}{cc|c} 0,2 \times 10^1 & 0,2 \times 10^1 & 0,6 \times 10^1 \\ 0,2 \times 10^{-3} & 0,2 \times 10^1 & 0,5 \times 10^1 \end{array} \right)$$

Assim o pivô é $0,2 \times 10^1$ e $m_{21} = (0,2 \times 10^{-3})/(0,2 \times 10^1) = 0,1 \times 10^{-3}$. De forma

análoga ao que fizemos acima, obtemos o novo sistema

$$A^{(0)}|b^{(0)} = \left(\begin{array}{cc|cc} 0,2 \times 10^1 & 0,2 \times 10^1 & 0,6 \times 10^1 & \\ 0 & 0,2 \times 10^1 & 0,5 \times 10^1 & \end{array} \right) \text{ Cujas solução é } \bar{x} = \begin{pmatrix} 0,5 \\ 0,25 \times 10^1 \end{pmatrix}$$

E o vetor \bar{x} é realmente a solução do nosso sistema, pois

$$\begin{aligned} 0,2 \times 10^{-3} \times 0,5 + 0,2 \times 10^1 \times 0,25 \times 10^1 &= 0,1 \times 10^{-3} + 0,05 \times 10^2 = 0,5 \times 10^1 = 5 \text{ e} \\ 0,2 \times 10^1 \times 0,5 + 0,2 \times 10^1 \times 0,25 \times 10^1 &= 0,1 \times 10^1 + 0,05 \times 10^2 = 0,01 \times 10^2 + 0,05 \times 10^2 = \\ 0,06 \times 10^2 &= 0,6 \times 10^1 = 6 \end{aligned}$$

Estratégia de Refinamento de Soluções: Quando se opera com números exatos, não se cometem erros de arredondamento no decorrer dos cálculos e as transformações elementares juntamente com as substituições (retroativas ou progressivas), produzem resultados exatos. Entretanto, na maioria das vezes, tem-se erros de arredondamento que podem se propagar, chegando mesmo a comprometer os resultados. Daí o uso de técnicas especiais para minimizar a propagação de tais erros de arredondamento. Uma das técnicas é a seguinte:

Seja $\bar{x}^{(0)}$ a solução aproximada para o sistema $Ax = b$.

Então, a solução melhorada $\bar{x}^{(1)}$ é obtida como se segue:

$\bar{x}^{(1)} = \bar{x}^{(0)} + \delta^{(0)}$, onde δ é uma parcela da correção. Logo:

$A\bar{x}^{(1)} = b$. Então, tem-se:

$$A(\bar{x}^{(0)} + \delta^{(0)}) = b$$

$$A\bar{x}^{(0)} + A\delta^{(0)} = b$$

$$A\delta^{(0)} = b - A\bar{x}^{(0)}$$

$$A\delta^{(0)} = r^{(0)}$$

Assim, para se obter a parcela de correção $\delta^{(0)}$ basta que resolva o sistema linear acima, onde A é a matriz de coeficiente das incógnitas do sistema $Ax = b$ e $r^{(0)}$ é o resíduo produzido pela solução aproximada $\bar{x}^{(0)}$.

A nova aproximação será, então,

$$\bar{x}^{(1)} = \bar{x}^{(0)} + \delta^{(0)}$$

Caso se queira uma melhor aproximação, resolve-se, agora, o sistema $A\delta^{(1)} = r^{(1)}$. Onde $\delta^{(1)}$ é a parcela de correção para $\bar{x}^{(1)}$ e $r^{(1)}$ é o resíduo produzido por $\bar{x}^{(1)}$.

O processo é repetido até que se obtenha a precisão desejada.

Exemplo: Resolver o sistema abaixo pelo método de Gauss, com duas casas decimais

$$\begin{cases} 8,7x_1 + 3,0x_2 + 9,3x_3 + 11,0x_4 = 16,4 \\ 24,5x_1 - 8,8x_2 + 11,5x_3 - 45,1x_4 = -49,7 \\ 52,3x_1 - 84,0x_2 - 23,5x_3 + 11,4x_4 = -80,8 \\ 21,0x_1 - 81,0x_2 - 13,2x_3 + 21,5x_4 = -106,3 \end{cases}$$

Resolvendo esse sistema obteremos $\bar{x} = [0,97 \ 1,98 \ -0,97 \ 1,00]^T$

com resíduo; $r = [0,042 \ 0,214 \ 0,594 \ -0,594]^T$. Fazendo:

$$\bar{x}^{(1)} = \bar{x}^{(0)} + \delta^{(0)} \text{ e } r = r^{(0)}$$

O cálculo da parcela é feito pelo sistema $A\delta^{(0)} = r^{(0)}$, que fornece como resultado:

$$\delta^{(0)} = [0,0295 \ 0,0195 \ -0,0294 \ 0,0000]^T, \bar{x} \text{ será, então:}$$

$$\bar{x}^{(1)} = \bar{x}^{(0)} + \delta^{(0)}$$

$\bar{x}^{(1)} = [1,000 \ 2,000 \ -0,999 \ 1,000]^T$, cujo resíduo é $r^{(1)} = [-0,009 \ -0,011 \ 0,024 \ 0,013]^T$, fazendo

$\bar{x}^{(2)} = \bar{x}^{(1)} + \delta^{(1)}$ e $r = r^{(1)}$, tem-se outra parcela de correção fornecida pelo sistema

$$A\delta^{(1)} = r^{(1)}$$

$\delta^{(1)} = [-0,0002 \ -0,0002 \ -0,0007 \ 0,0000]^T$. O valor melhorado de \bar{x} será:

$$\bar{x}^{(2)} = \bar{x}^{(1)} + \delta^{(1)}$$

$\bar{x}^{(2)} = [1,000 \ 2,000 \ -1,000 \ 1,000]^T$ com resíduo $r^{(2)} = [0 \ 0 \ 0 \ 0]^T$

4.0.2 Fatoração LU

Seja o sistema linear $Ax = b$.

O processo de fatoração para resolução deste sistema consiste em decompor a matriz A dos coeficientes em um produto de dois ou mais fatores e em seguida, resolver uma sequência de sistemas lineares que nos conduzirá à solução do sistema linear original.

Por exemplo, se pudermos realizar a fatoração $A = CD$, o sistema linear $Ax = b$, pode ser escrito:

$$(CD)x = b$$

se $y = Dx$, então resolver o sistema linear $Ax = b$ é equivalente a resolver o sistema linear $Cy = b$ e, em seguida o sistema linear $Dx = y$.

A vantagem dos processos de fatoração é que podemos resolver qualquer sistema linear que tenha A como matriz dos coeficientes. Se o vetor b for alterado, a resolução do sistema linear será quase imediata.

A fatoração LU é um dos processos mais empregados. Nesta fatoração a matriz L é triangular inferior com diagonal unitária e a matriz U é triangular superior.

Os fatores L e U podem ser obtidos através de fórmulas para os elementos l_{ij} e u_{ij} , ou então, podem ser construídos usando a idéia básica do método da Eliminação de Gauss.

A obtenção dos fatores L e U pelas fórmulas dificulta o uso da estratégia de pivoteamento e, por essa razão, veremos como obter L e U através do processo de Gauss.

Usaremos um exemplo teórico de dimensão 3:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

Trabalharemos somente com a matriz dos coeficientes. Seja então:

$$A^{(0)} = \begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & a_{13}^{(0)} \\ a_{21}^{(0)} & a_{22}^{(0)} & a_{23}^{(0)} \\ a_{31}^{(0)} & a_{32}^{(0)} & a_{33}^{(0)} \end{pmatrix} = A$$

Os multiplicadores da etapa 1 do processo de Gauss são:

$$m_{21} = \frac{a_{21}^{(0)}}{a_{11}^{(0)}} \text{ e } m_{31} = \frac{a_{31}^{(0)}}{a_{11}^{(0)}} \text{ (supondo que } a_{11}^{(0)} \neq 0 \text{)}$$

Para eliminar x_1 da linha i , $i = 2, 3$, multiplicaremos a linha 1 por m_{i1} e subtraímos o resultado na linha i .

os coeficientes $a_{ij}^{(0)}$ serão alterados para $a_{ij}^{(1)}$, onde:

$$a_{1j}^{(1)} = a_{1j}^{(0)} \text{ para } j = 1, 2, 3$$

$$a_{ij}^{(1)} = a_{ij}^{(0)} - m_{i1}a_{1j}^{(0)} \text{ para } i = 2, 3 \text{ e } j = 1, 2, 3$$

Estas operações correspondem a se pré-multiplicar a matriz $A^{(0)}$ pela matriz $M^{(0)}$, onde

$$M^{(0)} = \begin{pmatrix} 1 & 0 & 0 \\ -m_{21} & 1 & 0 \\ -m_{31} & 0 & 1 \end{pmatrix}, \text{ pois:}$$

$$M^{(0)}A^{(0)} = \begin{pmatrix} 1 & 0 & 0 \\ -m_{21} & 1 & 0 \\ -m_{31} & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & a_{13}^{(0)} \\ a_{21}^{(0)} & a_{22}^{(0)} & a_{23}^{(0)} \\ a_{31}^{(0)} & a_{32}^{(0)} & a_{33}^{(0)} \end{pmatrix} =$$

$$\begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & a_{13}^{(0)} \\ a_{21}^{(0)} - m_{21}a_{11}^{(0)} & a_{22}^{(0)} - m_{21}a_{12}^{(0)} & a_{23}^{(0)} - m_{21}a_{13}^{(0)} \\ a_{31}^{(0)} - m_{31}a_{11}^{(0)} & a_{32}^{(0)} - m_{31}a_{12}^{(0)} & a_{33}^{(0)} - m_{31}a_{13}^{(0)} \end{pmatrix} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} \end{pmatrix} = A^{(1)}$$

Portanto $M^{(0)}A^{(0)} = A^{(1)}$ onde $A^{(1)}$ é a mesma matriz obtida no final da etapa 1 do processo de Gauss.

Supondo agora que $a_{22}^{(1)} \neq 0$, o multiplicador da etapa 2 será: $m_{32} = \frac{a_{32}^{(1)}}{a_{22}^{(1)}}$

Para eliminar x_2 da linha 3, multiplicaremos a linha 2 por m_{32} e subtraímos o resultado da linha 3.

Os coeficientes $a_{ij}^{(1)}$ serão alterados para:

$$a_{1j}^{(2)} = a_{1j}^{(1)}, \text{ para } j = 1, 2, 3$$

$$a_{2j}^{(2)} = a_{2j}^{(1)}, \text{ para } j = 2, 3$$

$$a_{3j}^{(2)} = a_{3j}^{(1)} - m_{32}a_{2j}^{(1)}, \text{ para } j = 2, 3$$

As operações efetuadas em $A^{(1)}$ são equivalentes a pré-multiplicar $A^{(1)}$ por $M^{(1)}$, onde.

$$M^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -m_{32} & 1 \end{pmatrix}, \text{ pois:}$$

$$M^{(1)}A^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -m_{32} & 1 \end{pmatrix} \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} \end{pmatrix} =$$

$$\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{32}^{(1)} - m_{32}a_{22}^{(1)} & a_{33}^{(1)} - m_{32}a_{23}^{(1)} \end{pmatrix} = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} \\ 0 & 0 & a_{33}^{(2)} \end{pmatrix}$$

Portanto, $M^{(1)}A^{(1)} = A^{(2)}$ onde $A^{(2)}$ é a matriz obtida no final da etapa 2 do método da Eliminação de Gauss.

Temos então que:

$$A = A^{(0)}$$

$$A^{(1)} = M^{(0)}A^{(0)} = M^{(0)}A$$

$$A^{(2)} = M^{(1)}A^{(1)} = M^{(1)}M^{(0)}A^{(0)} = M^{(1)}M^{(0)}A$$

Onde $A^{(0)}$ é triangular superior, é fácil verificar que:

$$(M^{(0)})^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & 0 & 1 \end{pmatrix} \text{ e } (M^{(1)})^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & m_{32} & 1 \end{pmatrix}$$

$$\text{Assim, } (M^{(0)})^{-1}(M^{(1)})^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{pmatrix}$$

$$\text{Então, } A = (M^{(1)}M^{(0)})^{-1}A^2 = (M^{(0)})^{-1}(M^{(1)})^{-1}A^2$$

$$A = (M^{(0)})^{-1}(M^{(1)})^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{pmatrix} \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} \\ 0 & 0 & a_{33}^{(2)} \end{pmatrix} = LU$$

Ou seja: $L = (M^{(0)})^{-1}(M^{(1)})^{-1}$ e $U = A^2$

Isto é, fatoramos a matriz A em duas matrizes triangulares L e U, sendo que o fator L é triangular inferior com diagonal unitária e seus elementos l_{ij} para $i > j$ são os multiplicadores m_{ij} obtidos no processo de Eliminação de Gauss; o fator U é triangular superior e é a matriz triangular superior obtida no final da fase de triangularização do método da

Eliminação de Gauss.

Teorema 4.0.2 (Fatoração LU) Dada uma matriz quadrada A de ordem n , seja A_k a matriz constituída das primeiras k linhas e colunas de A . Suponha que $\det(A_k) \neq 0$ para $k = 1, 2, \dots, (n - 1)$. Então, existe uma única matriz triangular inferior $L = (m_{ij})$, com $m_{ij} = 1, 1 \leq i \leq n$ e uma única matriz triangular superior $U = (u_{ij})$ tais que $LU = A$. Ainda mais, $\det(A) = u_{11} u_{22} \dots u_{nn}$

Resolução do sistema linear $Ax = b$ usando a fatoração LU de A

Dados o sistema linear $Ax = b$ e a fatoração LU da matriz A , temos:

$$Ax = b \Leftrightarrow (LU)x = b$$

Seja $y = Ux$. A solução do sistema linear pode ser obtida da resolução dos sistemas lineares triangulares:

i) $Ly = b$

ii) $Ux = y$

Verifiquemos teoricamente que o vetor y é o vetor constante do lado direito obtido ao final do processo da Eliminação de Gauss.

Considerando o sistema linear $Ly = b$, temos que $y = L^{-1}b$.

Mas, $L = (M^{(0)})^{-1}(M^{(1)})^{-1} \Rightarrow L^{-1} = L = M^{(1)}M^{(0)}$

Então, $y = M^{(1)}M^{(0)}b^{(0)}$, onde $b^{(0)} = b$

temos que

$$M^{(0)}b^{(0)} = \begin{pmatrix} 1 & 0 & 0 \\ -m_{21} & 1 & 0 \\ -m_{31} & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1^{(0)} \\ b_2^{(0)} \\ b_3^{(0)} \end{pmatrix} = \begin{pmatrix} b_1^{(0)} \\ b_2^{(0)} - m_{21}b_1^{(0)} \\ b_3^{(0)} - m_{31}b_1^{(0)} \end{pmatrix} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(1)} \\ b_3^{(1)} \end{pmatrix} = b^{(1)}.$$

Isto é, o vetor obtido após o produto de $M^{(0)}$ por $b^{(0)}$ é o mesmo vetor do lado direito obtido após a etapa 1 do processo da Eliminação de Gauss.

$$\text{Obtido } b^{(1)}, \text{ temos que } y = M^{(1)}b^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -m_{32} & 1 \end{pmatrix} \begin{pmatrix} b_1^{(1)} \\ b_2^{(1)} \\ b_3^{(1)} \end{pmatrix} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(1)} \\ b_3^{(1)} - m_{32}b_2^{(1)} \end{pmatrix} = \begin{pmatrix} b_1^{(2)} \\ b_2^{(2)} \\ b_3^{(2)} \end{pmatrix} = b^{(2)}$$

4.0.3 Fatoração de Cholesky

Uma matriz $A : n \times n$ é definida positiva se $x^T Ax > 0$ para todo $x \in \mathfrak{R}^n, x \neq 0$.

A resolução de sistemas lineares em que a matriz A é simétrica, definida positiva, é fre-

quente em problemas práticos e tais matrizes podem ser fatoradas na forma:

$$AGG^T$$

Onde G: $n \times n$ é uma matriz triangular inferior com elementos da diagonal estritamente positivos. Esta fatoração é conhecida como Fatoração de Cholesky.

Seja A: $n \times n$ e vamos supor que A satisfaça as hipóteses do teorema 2. Então, A pode ser fatorada, de forma única, como $LD\bar{U}$ com:

L: $n \times n$, triangular inferior com diagonal unitária;

D: $n \times n$, diagonal e

\bar{U} : $n \times n$, triangular superior com diagonal unitária

Se, além das hipóteses do teorema 2, a matriz for simétrica $\bar{U} = L^T$, e, então, a fatoração fica: LDL^T .

Exemplo: Considere a matriz

$$A = \begin{pmatrix} 16 & -4 & 12 & -4 \\ -4 & 2 & -1 & 1 \\ 12 & -1 & 14 & -2 \\ -4 & 1 & -2 & 83 \end{pmatrix}$$

Calculando os fatores L e U de A e, seguida, os fatores L, D e \bar{U} , teremos:

$$\begin{pmatrix} 16 & -4 & 12 & -4 \\ -4 & 2 & -1 & 1 \\ 12 & -1 & 14 & -2 \\ -4 & 1 & -2 & 83 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1/4 & 1 & 0 & 0 \\ 3/4 & 2 & 1 & 0 \\ -1/4 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 16 & -4 & 12 & -4 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 81 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1/4 & 1 & 0 & 0 \\ 3/4 & 2 & 1 & 0 \\ -1/4 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 16 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 81 \end{pmatrix} \begin{pmatrix} 1 & -1/4 & 3/4 & -1/4 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 81 \end{pmatrix}$$

Observamos que:

i) $\bar{u}_{ij} = u_{ij}/u_{ii}$;

ii) Como a matriz A é simétrica, $\bar{U} = L^T$.

Se A for definida positiva, os elementos da matriz D são estritamente positivos, conforme demonstramos a seguir: Como A é definida positiva, temos que pra qualquer $x \in \mathfrak{R}^n, x \neq 0, x^T A x > 0$. Usando a fatoração LDL^T de A, temos:

$$0 < x^T (LDL^T)x = y^T D y.$$

Agora $y = L^T x$ e L tem posto completo. Então, $y \neq 0$ pois x é não nulo e, para cada $y \in \mathfrak{R}^n, \exists x \in \mathfrak{R}^n$, tal que $y = L^T x$.

Fazendo $y = e_i, i = 1, \dots, n$, teremos: $e_i^T D e_i = d_{ii}$, e, como $y^T D y > 0$, qualquer $y \neq 0$, obtemos: $d_{ii} > 0, i = 1, \dots, n$.

Concluindo, se A for simétrica definida positiva, então A pode ser fatorada na forma LDL^T com L triangular inferior com diagonal unitária e D matriz diagonal com elementos na diagonal estritamente positivos.

Podemos escrever então: $A = LDL^T = L\overline{D}\overline{D}L^T$, onde $\overline{d}_{ii} = \sqrt{d_{ii}}$ e, se $G = L\overline{D}$, obtemos $A = GG^T$ com G triangular inferior com diagonal estritamente positiva.

Teorema 4.0.3 (Fatoração de Cholesky) *Se A: $n \times n$ é simétrica e definida positiva, então existe uma única matriz triangular inferior G: $n \times n$ com diagonal positiva, tal que $A = GG^T$*

Exemplo:

Retomando a matriz A do exemplo anterior e sua fatoração LDL^T , observamos que o valor D é tal que $d_{ii} > 0, i = 1, \dots, 4$

Fazendo $\overline{D} = D^{1/2}$, teremos:

$A = LDL^T = L\overline{D}\overline{D}L^T = (L\overline{D})(\overline{D}L^T) = GG^T$, onde:

$$\overline{D} = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 9 \end{pmatrix} \text{ e } G = \begin{pmatrix} 4 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ -1 & 0 & 1 & 9 \end{pmatrix}$$

A matriz G triangular inferior com diagonal positiva, é o fator de Cholesky da matriz A.

Neste exemplo o fator de Cholesky foi obtido através da fatoração LDL^T , que por sua vez foi obtida a partir da fatoração LU. No entanto, o fator de Cholesky deve ser calculado através da equação matricial $A = GG^T$, uma vez que, assim, os cálculos envolvidos serão reduzidos pela metade.

Cálculo do fator de Cholesky: É dada A: $n \times n$, matriz simétrica e definida positiva:

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

O fator G: $n \times n$ triangular inferior com diagonal positiva será obtido a partir da equação matricial: $A = GG^T$

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

O cálculo será realizado por colunas:

$$\text{coluna 1: } \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} = G \begin{pmatrix} g_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} g_{11}^2 \\ g_{21}g_{11} \\ \vdots \\ g_{n1}g_{11} \end{pmatrix};$$

então: $g_{11} = \sqrt{a_{11}}$ e $g_{j1} = a_{j1}/g_{11}$, $j = 2, \dots, n$;

$$\text{Coluna 2: } \begin{pmatrix} a_{11} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{n2} \end{pmatrix} = G \begin{pmatrix} g_{11} \\ g_{22} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} g_{11}g_{21} \\ g_{21}^2 + g_{22}^2 \\ g_{31}g_{21} + g_{32}g_{22} \\ \vdots \\ g_{n1}g_{21} + g_{n2}g_{22} \end{pmatrix}$$

então $g_{21}^2 + g_{22}^2 \Rightarrow g_{22} = \sqrt{a_{22} - g_{21}^2}$ e $g_{j1}g_{21} + g_{j2}g_{22}$, $j = 3, \dots, n$.

Os elementos g_{j1} já estão calculados; assim, $g_{j2} = (a_{j1} - g_{j1}g_{21})/g_{22}$, $j = 3, \dots, n$.

Coluna k: Para obter os elementos da coluna k de G: $(0 \dots g_{kk}g_{k+1k} \dots g_{nk})^T$, $k = 3, \dots, n$, usamos a equação matricial:

$$\begin{pmatrix} a_{k1} \\ a_{k2} \\ \vdots \\ a_{kk} \\ a_{k+1k} \\ \vdots \\ a_{nk} \end{pmatrix} = G \begin{pmatrix} g_{k1} \\ g_{k2} \\ \vdots \\ g_{kk} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

e teremos:

$$a_{kk} = g_{k1}^2 + g_{k2}^2 + \dots + g_{kk}^2 \text{ e daí}$$

$$g_{kk} = \left(a_{kk} - \sum_{i=1}^{k-1} g_{ki}^2 \right)^{1/2}$$

e $a_{jk} = g_{j1}g_{k1} + g_{j2}g_{k2} + \dots + g_{jk}g_{kk}$, $j = (k + 1), \dots, n$

Como todos os elementos g_{ik} , $i = 1, \dots, (k - 1)$ já estão calculados, teremos:

$$g_{jk} = \left(a_{jk} - \sum_{i=1}^{k-1} g_{ji}g_{ki} \right) / g_{kk} \quad j = (k + 1), \dots, n.$$

Obtido o valor G, a resolução do sistema linear $Ax = b$ prossegue com a resolução dos

sistemas triangulares:

$$Ax = b \Leftrightarrow (GG^T)x = b \Rightarrow \begin{cases} i) Gy = b \\ ii) G^T x = y \end{cases}$$

4.1 Métodos Iterativos

São métodos que geram uma sequência de valores $x^{(k)}$ a partir de uma aproximação inicial $x^{(0)}$, nos Métodos Iterativos dado um vetor de de aproximação inicial $x^{(0)}$ calculamos uma nova aproximação que em seguida será utilizada novamente até chegarmos a uma boa aproximação da solução exata do sistema, ou seja, uma vez que $x^{(0)}$ é dado, o Método Iterativo gera a partir de $x^{(0)}$ uma nova aproximação $x^{(1)}$ em seguida, $x^{(1)}$ é usado para gerar uma nova e melhor aproximação e assim por diante até uma melhor aproximação $x^{(k)}$ que espera-se convergir para x . Como o número de iterações pode ser infinita, na prática não podemos iterar para sempre, logo algum critério de para deve ser pré-estabelecido. Um exemplo para critério de parada pode ser definido por uma cota de tolerância ϵ , ou seja, encerra-se o processo de iteração e $x^{(k)}$ é escolhido como a solução aproximada adequada para o problema uma vez que ele esteja suficientemente próximo da solução exata o quanto se precise.

Observe que: quando $\|b - Ax^{(k)}\| < \epsilon$ ou quando $\|x^{(k+1)} - x^{(k)}\| < \epsilon$

São obedecidas, então as iterações podem ser interrompidas. Logo o último valor aproximado obtido é aceito como a melhor aproximação da solução exata dentro das circunstâncias pré-estabelecidas.

Os Métodos Iterativos não depende do vetor de aproximação inicial $x^{(0)}$, na prática isso significa que quanto mais aproximado da solução do sistema estiver $x^{(0)}$ menor será o número de iterações a serem calculadas para chegar a solução aproximada com a precisão desejada.

Em termos mais específicos temos:

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

A: matriz dos coeficientes, $n \times n$:

x: vetor das variáveis, $n \times 1$:

b: vetor dos termos constantes, $n \times 1$.

Esse sistema é convertido, de alguma forma, num sistema do tipo $x = Cx + g$, onde C é matriz $n \times n$ e g é o vetor $n \times 1$. Observamos que $\varphi(x) = Cx + g$ é uma função de interação dada na forma matricial.

É então proposto o esquema iterativo:

Partimos de $x^{(0)}$ (vetor aproximação inicial) e então construímos consecutivamente os vetores:

$$x_{(1)} = Cx^{(0)} + g = \varphi(x^{(0)}), \text{ (primeira aproximação),}$$

$x_{(2)} = Cx^{(1)} + g = \varphi(x^{(1)})$, (segunda aproximação) etc.

De um modo geral, a aproximação $X^{(k+1)}$ é calculada pela fórmula $x^{(k+1)} = Cx^{(k)} + g$, ou seja, $x^{(k+1)} = \varphi(x^{(k)})$, $k = 0, 1, \dots$

É importante observar que se a sequência de aproximações $x^{(0)}, x^{(1)}, \dots, x^{(k)}, \dots$ é tal que, $\lim_{k \rightarrow \infty} x^{(k)} = \alpha$, então $\alpha = C\alpha + g$, ou seja, α é solução do sistema linear $Ax = b$.

4.1.1 Obtenção dos Métodos Iterativos

Dado um sistema linear $Ax = b$, decompõem-se a matriz A na forma $A = M + N$, onde M é uma matriz não singular. Logo teremos:

$$Ax = b$$

$$(M + N)x = b$$

$$Mx + Nx = b$$

$$Mx = b - Nx$$

$$M^{-1}Mx = M^{-1}b - M^{-1}Nx$$

$$Ix = M^{-1}b - M^{-1}Nx$$

$$x = M^{-1}b - M^{-1}Nx$$

O método iterativo é então definido por:

$$x^{(k+1)} = M^{-1}b - M^{-1}Nx^{(k)},$$

com $k = 0, 1, 2, \dots$

onde, $x^{(0)}$ é um vetor dado.

O sistema linear pode ser escrito na forma:

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

Onde, $g \in M_{n \times 1}(\mathfrak{R})$ e $C \in M_n(\mathfrak{R})$. é importante que a matriz M seja muito mais simples do que A , caso contrário não estaríamos simplificando o problema. Observe que esta iteração pode ser feita de outra maneira, não havendo necessidade de se proceder verdadeiramente ao cálculo da inversa de M .

Iremos abordar os seguintes métodos iterativos:

◇ Método de Jacobi

◇ Método de Gauss-Seidel

◇ Método SOR.

Os métodos iterativos são definidos, por uma escolha particular da matriz M . A matriz M é geralmente diagonal, triangular ou tridiagonal.

Seja o sistema $Ax = b$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$A = \underbrace{\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ a_{21} & 0 & 0 & \cdots & 0 \\ a_{31} & a_{32} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & 0 \end{pmatrix}}_L + \underbrace{\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}}_D + \underbrace{\begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ 0 & 0 & 0 & \cdots & a_{32} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}}_U$$

Onde L é uma matriz triangular estritamente inferior, U uma matriz triangular estritamente e D a matriz diagonal.

Notamos que a matriz D não deverá ter zeros na diagonal principal, isto é, seus elementos $a_{ii} \neq 0$. Caso aconteça de $a_{ii} = 0$, deve-se efetuar uma troca de linhas ou colunas na matriz A, para obtermos uma matriz D com os elementos $a_{ii} \neq 0$.

4.1.2 Critério de Convergência dos Métodos Iterativos

Vamos estabelecer critérios que nos permita determinar quando existe convergência para estes métodos iterativos. Pretendem-se métodos que satisfaçam:

$$\lim_{k \rightarrow \infty} x^{(k)} = x$$

ou

$$\lim_{k \rightarrow \infty} \|x^{(k)} - x\| = 0.$$

De

$$x = M^{-1}b - M^{-1}Nx$$

$$x^{(k+1)} = M^{-1}b - M^{-1}Nx^{(k)},$$

temos

$$x - x^{(k+1)} = -M^{-1}Nx + M^{-1}Nx^{(k)}$$

$$x - x^{(k+1)} = -M^{-1}N(x - x^{(k)})$$

Fazendo $C = -M^{-1}N$, onde a matriz C é denominada *matriz de iteração*, e $x - x^{(k+1)}$ é o erro cometido na k -ésima iteração, tem-se:

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

Atribuindo valores para $k \geq 0$, temos:

$$\begin{aligned} x - x^{(1)} &= C(x - x^{(0)}) \\ x - x^{(2)} &= C(x - x^{(1)}) = C.C(x - x^{(0)}) = C^2(x - x^{(0)}) \\ x - x^{(3)} &= C(x - x^{(2)}) = C.C^2(x - x^{(0)}) = C^3(x - x^{(0)}) \\ x - x^{(4)} &= C(x - x^{(3)}) = C.C^3(x - x^{(0)}) = C^4(x - x^{(0)}) \\ &\vdots \end{aligned}$$

Em geral tem-se

$$x - x^{(k+1)} = C^{(k+1)}(x - x^{(0)})$$

Aplicando $\| \cdot \|$ a norma da matriz associada temos

$$0 \leq \| x - x^{(k+1)} \| = \| C^{(k+1)}(x - x^{(0)}) \| \leq \| C^{(k+1)} \| \| x - x^{(0)} \|$$

Por outro lado temos que

$$\| C^{(k+1)} \| \leq \| C \|^{(k+1)}$$

Dai segue que

$$0 \leq \| x - x^{(k+1)} \| \leq \| C \|^{(k+1)} \| x - x^{(0)} \|$$

Vemos que se $\| C \| < 1$ então

$$\lim_{k \rightarrow \infty} \| C \|^{(k+1)} = 0 \Rightarrow \lim_{k \rightarrow \infty} \| x^{(k)} - x \| = 0$$

Portanto o método iterativo convergente se $\| C \| < 1$, para qualquer que seja a iteração inicial $x^{(0)}$. Logo temos provado o seguinte teorema:

Teorema 4.1.1 *Se $\| C \| < 1$ para qualquer norma matricial, então o método iterativo definido por $x^{(k+1)} = Cx^{(k)} + g$ converge qualquer que seja o vetor $x^{(0)}$ dado.*

4.1.3 Estimativa de erro na Iteração

$$x^{(k+1)}$$

De

$$\begin{aligned}x &= M^{-1}b - M^{-1}Nx \\x^{(k+1)} &= M^{-1}b - M^{-1}Nx^{(k)},\end{aligned}$$

temos

$$\begin{aligned}x - x^{(k+1)} &= -M^{-1}Nx + M^{-1}Nx^{(k)} \\x - x^{(k+1)} &= -M^{-1}N(x - x^{(k)})\end{aligned}$$

Fazendo $C = -M^{-1}$, tem-se

$$\begin{aligned}x - x^{(k+1)} &= C(x - x^{(k)}) \\x - x^{(k+1)} &= C(x - x^{(k+1)} + x^{(k+1)} - x^{(k)}).\end{aligned}$$

Aplicando norma $\| \cdot \|$ resulta,

$$\begin{aligned}\|x - x^{(k+1)}\| &= \|C(x - x^{(k+1)} + x^{(k+1)} - x^{(k)})\| \\&\leq \|C\| \|x - x^{(k+1)} + x^{(k+1)} - x^{(k)}\| \\&\|x - x^{(k+1)}\| \leq \|C\| (\|x - x^{(k+1)}\| + \|x^{(k+1)} - x^{(k)}\|) \\&= \|C\| \|x - x^{(k+1)}\| + \|C\| \|x^{(k+1)} - x^{(k)}\| \\&\|x - x^{(k+1)}\| \leq \|C\| \|x - x^{(k+1)}\| + \|C\| \|x^{(k+1)} - x^{(k)}\| \\&\|x - x^{(k+1)}\| - \|C\| \|x - x^{(k+1)}\| \leq \|C\| \|x^{(k+1)} - x^{(k)}\| \\&\|x - x^{(k+1)}\| (1 - \|C\|) \leq \|C\| \|x^{(k+1)} - x^{(k)}\|.\end{aligned}$$

Logo, a estimativa de erro na iteração $k + 1$ dos métodos iterativos é dada pela expressão:

$$\|x - x^{(k+1)}\| \leq \frac{\|C\|}{1 - \|C\|} \|x^{(k+1)} - x^{(k)}\|$$

4.1.4 Teste de Parada

O processo iterativo é repetido até que o vetor $x^{(k)}$ esteja suficientemente próximo do vetor $x^{(k-1)}$.

Medimos a distância entre $x^{(k)}$ e $x^{(k-1)}$ por $d^{(k)} = \max_{1 \leq i \leq n} |x_i^{(k)} - x_i^{(k-1)}|$

Assim, dada uma precisão ϵ , o vetor $x^{(k)}$ será escolhido como \bar{x} , solução aproximada da solução exata, se $d^{(k)} < \epsilon$.

Da mesma maneira que no teste de parada dos Métodos Iterativos para zeros de funções, podemos efetuar aqui o teste do erro relativo:

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

Computacionalmente usamos também como teste de parada um número máximo de iterações.

4.1.5 Método de Gauss-Jacobi

Seja o sistema linear $Ax = b$ e considerando $A = L + D + U$, podemos então escrever o sistema na forma

$$(L + D + U)x = b$$

$$(L + U)x + Dx = b$$

$$Dx = b - (L + U)x$$

$$D^{-1}Dx = D^{-1}b - D^{-1}(L + U)x$$

$$Ix = D^{-1}b - D^{-1}(L + U)x$$

$$x = D^{-1}b - D^{-1}(L + U)x$$

O método iterativo definido por:

$$x^{(k+1)} = D^{-1}b - D^{-1}(L + U)x^{(k)} \text{ com } k = 0, 1, 2, \dots$$

Onde $x^{(0)}$ é um vetor dado, é chamado de Método de Gauss-Jacobi.

Este método é do tipo iterativo definido por $x^{(k+1)} = M^{-1}b - M^{-1}Nx^{(k)}$ com:

$M = D$ e $N = L + U$, da equação $x^{(k+1)} = D^{-1}b - D^{-1}(L + U)x^{(k)}$:

$$x^{(k+1)} = D^{-1}(b - (L + U)x^{(k)})$$

A forma como o Método de Gauss-Jacobi transforma o sistema linear $Ax = b$ em $x = Cx + g$ é a seguinte:

Tomamos o sistema original:

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

e supondo $a_{ii} \neq 0, i = 1, \dots, n$, isolamos o vetor x mediante a separação pela diagonal, assim:

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

Desta forma, temos $x = Cx + g$, onde

$$C = \begin{pmatrix} 0 & -a_{12}/a_{11} & -a_{13}/a_{11} & \cdots & -a_{1n}/a_{11} \\ -a_{21}/a_{22} & 0 & -a_{23}/a_{22} & \cdots & -a_{2n}/a_{22} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -a_{n1}/a_{nn} & -a_{n2}/a_{nn} & -a_{n3}/a_{nn} & \cdots & 0 \end{pmatrix} \text{ e } g = \begin{pmatrix} b_1/a_{11} \\ b_2/a_{22} \\ \vdots \\ b_n/a_{nn} \end{pmatrix}$$

O método de Gauss-Jacobi consiste em, dado $x^{(0)}$, aproximação inicial, obter $x^{(1)}, \dots, x^{(k)}$, ... através da relação recursiva $x^{(k+1)} = Cx^{(k)} + g$:

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

Em geral $x_i^{(k+1)}$ pode ser obtido pela fórmula:

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

$i = 1, 2, 3, \dots, n$

Exemplo Resolva o sistema linear:

$$\begin{cases} 10x_1 + 2x_2 + x_3 = 7 \\ x_1 + 5x_2 + x_3 = -8 \\ 2x_1 + 3x_2 + 10x_3 = 6 \end{cases}$$

pelo método de Gauss-jacobi com $x^{(0)} = \begin{pmatrix} 0.7 \\ -1.6 \\ 0.6 \end{pmatrix}$ e $\epsilon = 0.05$.

O processo iterativo é

$$\begin{cases} x_1^{(k+1)} = \frac{1}{10}(7 - 2x_2^{(k)} - x_3^{(k)}) = 0x_1^{(k)} - \frac{2}{10}2x_2^{(k)} - \frac{1}{10}x_3^{(k)} + \frac{7}{10} \\ x_2^{(k+1)} = \frac{1}{5}(-8 - x_2^{(k)} - x_3^{(k)}) = -\frac{1}{5}x_1^{(k)} + 0x_2^{(k)} - \frac{1}{5}x_3^{(k)} - \frac{8}{5} \\ x_3^{(k+1)} = \frac{1}{10}(6 - 2x_2^{(k)} - 3x_3^{(k)}) = -\frac{2}{10}x_1^{(k)} - \frac{3}{10}x_2^{(k)} + 0x_3^{(k)} + \frac{6}{10} \end{cases}$$

Na forma matricial $x^{(k+1)} = Cx^{(k)} + g$ temos:

$$C = \begin{pmatrix} 0 & -2/10 & -1/10 \\ -1/5 & 0 & -1/5 \\ -1/5 & -3/10 & 0 \end{pmatrix} \text{ e } g = \begin{pmatrix} 7/10 \\ -8/5 \\ 6/10 \end{pmatrix}$$

Assim ($k = 0$)

$$\begin{cases} x_1^{(1)} = -0.2x_2^{(0)} - 0.1x_3^{(0)} + 0.7 = 0.2(-1.6) - 0.1 \times 0.6 + 0.7 = 0.96 \\ x_2^{(1)} = -0.2x_1^{(0)} - 0.2x_3^{(0)} - 1.6 = -0.2 \times 0.7 - 0.2 \times 0.6 - 1.6 = -1.86 \\ x_3^{(1)} = -0.2x_1^{(0)} - 0.3x_2^{(0)} + 0.6 = -0.2 \times 0.7 - 0.3 \times (-0.6) + 0.6 = 0.94 \end{cases}$$

$$\text{ou } x^{(1)} = Cx^{(0)} + g = \begin{pmatrix} 0.96 \\ -1.86 \\ 0.94 \end{pmatrix}.$$

Calculando $d_r^{(1)}$, temos:

$$\begin{aligned} |x_1^{(1)} - x_1^{(0)}| &= 0.26 \\ |x_2^{(1)} - x_2^{(0)}| &= 0.26 \Rightarrow d_r^{(1)} = \frac{0.34}{\max_{1 \leq i \leq n} |x_i^{(1)}|} = \frac{0.34}{1.86} = 1.1828 > \epsilon \end{aligned}$$

$$|x_3^{(1)} - x_3^{(0)}| = 0.34$$

Prosseguindo as iterações, temos:

para $k = 1$:

$$x^{(2)} = \begin{pmatrix} 0.978 \\ -1.98 \\ 0.9984 \end{pmatrix} \Rightarrow d_r^{(2)} = \frac{0.0324}{1.9888} = 0.0163 > \epsilon.$$

e para $k = 2$:

$$x^{(3)} = \begin{pmatrix} 0.9994 \\ -1.9888 \\ 0.966 \end{pmatrix} \Rightarrow d_r^{(2)} = \frac{0.12}{1.98} = 0.0606 > \epsilon.$$

Então, a solução \bar{x} do sistema linear acima, com o erro menor que 0.05, obtida pelo método de Gauss-Jacobi, é

$$\bar{x} = x^{(3)} = \begin{pmatrix} 0.9994 \\ -1.9888 \\ 0.966 \end{pmatrix}.$$

Neste exemplo tomamos $x^{(0)} = \begin{pmatrix} 0.7 \\ -1.6 \\ 0.6 \end{pmatrix} = \begin{pmatrix} b_1/a_{11} \\ b_2/a_{22} \\ b_3/a_{33} \end{pmatrix}$. No entanto, o valor de $x^{(0)}$ é

arbitrário, pois veremos mais adiante que a convergência ou não de um método iterativo para a solução de um sistema linear de equações é independente da aproximação inicial escolhida.

CRITÉRIO DE CONVERGÊNCIA: Daremos aqui um teorema que estabelece uma condição suficiente para a convergência do método iterativo de Gauss-Jacobi.

Teorema 4.1.2 : (Critério das linhas) *Seja o sistema linear $Ax = b$ e seja $\alpha_k = (\sum_{j=1, j \neq k}^n |a_{kj}|) / |a_{kk}|$. Se $\alpha = \max_{1 \leq i \leq n} \alpha_k < 1$, então o método de Gauss-Jacobi gera uma sequência $x^{(k)}$ converge para a solução do sistema dado, independente da escolha da aproximação inicial, $x^{(0)}$.*

Exemplo:

Seja a matriz A do sistema linear do exemplo anterior:

$$A = \begin{pmatrix} 10 & 2 & 1 \\ 1 & 5 & 1 \\ 2 & 3 & 10 \end{pmatrix}, \text{ temos}$$

$\alpha_1 = \frac{2+1}{10} = 0.3 < 1$; $\alpha_2 = \frac{1+1}{5} = 0.4 < 1$; $\alpha_3 = \frac{2+3}{10} = 0.5 < 1$ e então $\max_{1 \leq k \leq 3} \alpha_k = 0.5 < 1$ donde, pelo critério das linhas, temos garantia de convergência para o método de Gauss-Jacobi.

Exemplo: Para o sistema linear $\begin{cases} x_1 + x_2 = 3 \\ x_1 - 3x_2 = -3 \end{cases}$

o método de Gauss-Jacobi gera uma sequência convergente para a solução exata $x^* = \begin{pmatrix} 3/2 \\ 3/2 \end{pmatrix}$.

No entanto, o critério das linhas não é satisfeito, visto que $\alpha_1 = 1/1 = 1$. Isto mostra que a condição do teorema é apenas suficiente.

Exemplo:

$$\text{A matriz A do sistema } \begin{cases} x_1 + 3x_2 + x_3 = -2 \\ 5x_1 + 2x_2 + 2x_3 = 3 \\ 0.x_1 + 6x_2 + 8x_3 = -6 \end{cases}$$

Não satisfaz o critério das linhas pois $\alpha_1 = \frac{3+1}{1} = 4 > 1$. contudo, se permutarmos a primeira equação com a segunda, temos o sistema linear.

$$\begin{cases} 5x_1 + 2x_2 + 2x_3 = 3 \\ x_1 + 3x_2 + x_3 = -2 \\ 0.x_1 + 6x_2 + 8x_3 = -6 \end{cases}$$

que é equivalente ao sistema original e a matriz $\begin{pmatrix} 5 & 2 & 2 \\ 1 & 3 & 1 \\ 0 & 6 & 8 \end{pmatrix}$

deste novo sistema satisfaz o critério das linhas.

Assim, é conveniente aplicar o método de Gauss-Jacobi a esta nova disposição do sistema, pois dessa forma a convergência está assegurada.

Concluindo, sempre que o critério das linhas não for satisfeito, devemos tentar uma permutação de linhas e/ou colunas de forma a obtemos uma disposição para a qual a matriz dos coeficientes satisfaça o critério das linhas. No entanto, nem sempre é possível obter tal disposição.

4.1.6 Método Iterativo de Gauss-Seidel

Este método iterativo converge duas vezes mais rápido que o método de Gauss-Jacobi (pelo menos na grande maioria das aplicações), neste método os valores de x são atualizados dentro de cada iteração, sem esperar pela próxima. Em outras palavras obtido o valor $x_j^{(k+1)}$ este é usado no lugar de $x_j^{(k)}$ no cálculo seguinte.

Seja o sistema de equações lineares $Ax = b$, com $A = L + D + U$, daí temos que:

$$Ax = b$$

$$(L + D + U)x = b$$

$$(L + D)x + Ux = b$$

$$(L + D)x = b - Ux$$

$$(L + D)^{-1}(L + D)x = (L + D)^{-1}b - (L + D)^{-1}Ux$$

$$Ix = (L + D)^{-1}b - (L + D)^{-1}Ux$$

$$x = (L + D)^{-1}b - (L + D)^{-1}Ux$$

O método iterativo definido por:

$$x^{(k+1)} = (L + D)^{-1}b - (L + D)^{-1}Ux^{(k)} \text{ com } k = 0, 1, 2, \dots$$

Onde $x^{(0)}$ é um vetor inicial dado, é chamado de Método de Gauss-seidel.

Este método é do tipo iterativo definido por $x^{(k+1)} = M^{-1}b - M^{-1}Nx^{(k)}$ com:

$$M = L + D \text{ e } N = U$$

Temos que:

$$x^{(k+1)} = (L + D)^{-1}b - (L + D)^{-1}Ux^{(k)}$$

$$(L + D)x^{(k+1)} = (L + D)(L + D)^{-1}b - (L + D)(L + D)^{-1}Ux^{(k)}$$

$$Lx^{(k+1)} + Dx^{(k+1)} = Ib - IUx^{(k)}$$

$$Dx^{(k+1)} = b - Lx^{(k+1)} - Ux^{(k)} \Rightarrow Dx^{(k+1)} = D^{-1}(b - Lx^{(k+1)} - Ux^{(k)})$$

$$x^{(k+1)} = D^{-1}(b - Lx^{(k+1)} - Ux^{(k)}).$$

Da mesma forma no método de Gauss-Jacobi, ao método de Gauss-Seidel o sistema linear

$Ax = b$ é escrito na forma equivalente $x = Cx + g$ por separação da diagonal.

O processo iterativo consiste em, sendo $x^{(0)}$ uma aproximação inicial, calcular $x^{(1)}, x^{(2)}, \dots, x^{(k)}, \dots$

por:

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

De forma geral $x_i^{(k+1)}$ pode ser obtido pela fórmula abaixo:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right)$$

Portanto, no processo iterativo de Gauss-Seidel, no momento de se calcular $x_j^{(k+1)}$ usamos todos os valores $x_1^{(k+1)}, \dots, x_{j-1}^{(k+1)}$ que já foram calculados e os valores $x_{j+1}^{(k)}, \dots, x_n^{(k)}$ restantes.

Exemplo: Resolva o sistema linear:

$$\begin{cases} 5x_1 + x_2 + x_3 = 5 \\ 3x_1 + 4x_2 + x_3 = 6 \\ 3x_1 + 3x_2 + 6x_3 = 0 \end{cases}$$

pelo método de Gauss-Seidel com $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ e $\varepsilon = 5 \times 10 \times 10^{-2}$. O processo iterativo

é:

$$\begin{cases} x_1^{(k+1)} = 1 - 0.2x_2^{(k)} - 0.2x_3^{(k)} \\ x_2^{(k+1)} = 1.5 - 0.75x_1^{(k+1)} - 0.25x_3^{(k)} \\ x_3^{(k+1)} = 0 - 0.5x_1^{(k+1)} - 0.5x_2^{(k+1)} \end{cases} \text{ com } x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

(k = 0)

$$\begin{cases} x_1 = 1 - 0 - 0 = 1 \\ x_2 = 1.5 - 0.75 \times 1 - 0 = 0.75 \\ x_3 = -0.5 \times 1 - 1 - 0.5 \times 0.75 = -0.875 \end{cases} \Rightarrow x^{(1)} = \begin{pmatrix} 1 \\ 0.75 \\ -0.875 \end{pmatrix}, \text{ donde}$$

$$|x_1^{(1)} - x_1^{(0)}| = 1$$

$$|x_2^{(1)} - x_2^{(0)}| = 0.75$$

$$|x_3^{(1)} - x_3^{(0)}| = 0.875$$

$$\frac{0.2}{\max_{1 \leq i \leq 3} |x_i^{(2)}|} \Rightarrow \frac{1}{\max_{1 \leq i \leq 3} |x_i^{(1)}|} = 1 > \varepsilon$$

Assim (k=1) é

$$\begin{cases} x_1^{(2)} = 1 - 0.2 \times 0.75 + 0.2 \times 0.875 = 1.025 \\ x_2^{(2)} = 1.5 - 0.75 \times 1.025 - 0.25 \times (-0.875) = 0.95 \\ x_3^{(2)} = -0.5 \times 1.025 - 0.5 \times 0.95 = -0.9875 \end{cases}$$

$$\Rightarrow x^{(2)} = \begin{pmatrix} 1.025 \\ 0.95 \\ -0.9875 \end{pmatrix}, \text{ donde}$$

$$|x_1^{(2)} - x_1^{(1)}| = 0.025$$

$$|x_2^{(2)} - x_2^{(1)}| = 0.20$$

$$|x_3^{(2)} - x_3^{(1)}| = 0.1125$$

$$\Rightarrow d_r^{(2)} = \frac{0.2}{\max_{1 \leq i \leq 3} |x_i^{(2)}|} = \frac{0.2}{1.025} = 0.1951 > \varepsilon$$

Continuando as iterações obtemos:

$$x^{(3)} = \begin{pmatrix} 1.0075 \\ 0.9912 \\ -0.9993 \end{pmatrix} \Rightarrow d_r^{(3)} = 0.0409 < \varepsilon$$

Assim, a solução \bar{x} do sistema dado com erro menor que ε , pelo método de Gauss-Seidel, é

$$\bar{x} = x^{(3)} = \begin{pmatrix} 1.0075 \\ 0.9912 \\ -0.9993 \end{pmatrix}.$$

O esquema iterativo do método de Gauss-Seidel pode ser escrito na forma matricial da seguinte maneira:

Inicialmente escrevemos a matriz A, dos coeficientes, como $A = L + D + R$, onde:

L: matriz triangular inferior com diagonal nula;

D: matriz diagonal com $d_{ii} \neq 0, i = 1, \dots, n$;

R: matriz triangular superior com diagonal nula.

O modo mais simples de se escrever A nesta forma é:

$$L = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ a_{21} & 0 & \cdots & 0 \\ a_{31} & a_{32} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, D = \begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & a_{33} & \\ & & & \ddots \\ & & & & a_{nn} \end{pmatrix} \text{ e } R = \begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Portanto, $Ax = b \Leftrightarrow (L+D+R)x = b \Leftrightarrow Dx = b - Lx - Rx \Leftrightarrow x = D^{-1}b - D^{-1}Lx - D^{-1}Rx$

No método de Gauss-Seidel o vetor $x^{(k+1)}$ é calculado por:

$$x^{(k+1)} = D^{-1}b - D^{-1}Lx^{(k+1)} - D^{-1}Rx^{(k)}.$$

Agora podemos ainda escrever $x^{(k+1)} = Cx^{(k)} + g$, Considerando que $A = D(L_1 + I + R_1)$

onde:

$$L_1 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \frac{a_{21}}{a_{22}} & 0 & \cdots & 0 \\ \frac{a_{31}}{a_{33}} & \frac{a_{32}}{a_{33}} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \frac{a_{n1}}{a_{nn}} & \frac{a_{n2}}{a_{nn}} & \cdots & 0 \end{pmatrix} R_1 = \begin{pmatrix} 0 & \frac{a_{12}}{a_{11}} & \frac{a_{13}}{a_{11}} & \cdots & \frac{a_{1n}}{a_{11}} \\ 0 & 0 & \frac{a_{23}}{a_{22}} & \cdots & \frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

então $Ax = b \Leftrightarrow D(L_1 + I + R_1)x = b \Leftrightarrow (L_1 + I + R_1)x = D^{-1}b \Leftrightarrow x = -L_1x + R_1x + D^{-1}b$

e o método de Gauss-Seidel é

$$x^{(k+1)} = -L_1 x^{(k+1)} - R_1 x^{(k)} + D^{-1}b, \text{ donde } (I + L_1)x^{(k+1)} = -R_1 x^{(k)} + D^{-1}b$$

$$\text{ou } x^{(k+1)} = \underbrace{-(I + L_1)^{-1}R_1}_{C} x^{(k)} + \underbrace{(I + L_1)^{-1}D^{-1}b}_g = Cx^{(k)} + g$$

Método da Convergência do Método de Gauss-Seidel: Como em todo processo iterativo, precisamos de critérios que nos forneçam garantia de convergência.

Para o método de Gauss-Seidel analisaremos os seguintes critérios, que estabelecem condições suficientes de convergência: o critério de Sassenfeld e o critério das linhas.

Critério de Sassenfeld

$$\text{Seja } x^* = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ a solução exata do sistema } Ax = b \text{ e seja:}$$

$$x^{(k)} = \begin{pmatrix} x_1^k \\ x_2^k \\ \vdots \\ x_n^k \end{pmatrix} \text{ a } k\text{-ésima aproximação de } x^*$$

Queremos uma condição que nos garanta que $x^{(k)} \rightarrow x^*$ quando $k \rightarrow \infty$, ou seja, que $\lim_{k \rightarrow \infty} e_i^{(k)} = 0$ para $i = 1, \dots, n$ onde $e_i^{(k)} = x_i^{(k)} - x_i^*$ agora,

$$\begin{cases} e_1^{(k+1)} = & -\frac{1}{a_{11}}(a_{12}e_2^{(k)} + a_{13}e_3^{(k)} + \dots + a_{1n}e_n^{(k)}) \\ e_2^{(k+1)} = & -\frac{1}{a_{22}}(a_{21}e_1^{(k)} + a_{23}e_3^{(k)} + \dots + a_{2n}e_n^{(k)}) \\ \vdots & \vdots \\ e_n^{(k+1)} = & -\frac{1}{a_{nn}}(a_{n1}e_1^{(k+1)} + a_{n2}e_2^{(k+1)} + \dots + a_{n,n-1}e_{n-1}^{(k+1)}) \end{cases}$$

Chamaremos de $E^{(k)} = \max_{1 \leq i \leq n} |e_i^{(k)}|$ e sejam

$$\beta_1 = \sum_{j=2}^n |a_{1j}|/|a_{11}| \text{ e para } i = 2, 3, \dots, n$$

$$\beta_i = \left[\sum_{j=1}^{i-1} \beta_j |a_{ij}| + \sum_{j=i+1}^n |a_{ij}| \right] / |a_{ii}|$$

Note que a condição $x^{(k)} \rightarrow x^*$ equivale a $E^k \rightarrow 0$ quando $k \rightarrow \infty$.

Mostraremos por indução que $E^{(k+1)} \leq \beta E^{(k)}$ onde $\beta = \max_{1 \leq i \leq n} \beta_i$.

Para $i = 1$, temos

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

$$\text{Ent\~{a}o, } |e_1^{(k+1)}| \leq \beta_1 \max_{1 \leq j \leq n} |e_j^{(k)}| \leq \beta \max_{1 \leq j \leq n} |e_j^{(k)}|$$

Suponhamos por indu\~{c}o\~{a}o que:

$$|e_2^{(k+1)}| \leq \beta_2 \max_{1 \leq j \leq n} |e_j^{(k)}|$$

$$|e_3^{(k+1)}| \leq \beta_3 \max_{1 \leq j \leq n} |e_j^{(k)}|$$

⋮

$$|e_{i-1}^{(k+1)}| \leq \beta_{i-1} \max_{1 \leq j \leq n} |e_j^{(k)}| \quad |i \leq n$$

e mostraremos que $|e_i^{(k+1)}| \leq \beta_i \max_{1 \leq j \leq n} |e_j^{(k)}|$. mas,

$$\begin{aligned} |e_i^{(k+1)}| &\leq \frac{1}{|a_{ii}|} (|a_{i1}| |e_1^{(k+1)}| + |a_{i2}| |e_2^{(k+1)}| + \dots + |a_{i,i-1}| |e_{i-1}^{(k+1)}|) + \frac{1}{|a_{ii}|} (|a_{i,i+1}| |e_{i+1}^{(k)}| + \dots + \\ &|a_{in}| |e_{in}^{(k)}|) \end{aligned}$$

e usando a hip\~{o}tese de indu\~{c}o\~{a}o:

$$|e_i^{(k+1)}| \leq \frac{1}{|a_{ii}|} \underbrace{(|a_{i1}| \beta_1 + |a_{i2}| \beta_2 + \dots + |a_{i,i-1}| \beta_{i-1} + |a_{i,i+1}| + \dots + |a_{in}|)}_{\beta_i} \max_{1 \leq j \leq n} |e_j^{(k)}|, \text{ ou seja,}$$

$$|e_i^{(k+1)}| \leq \beta_i \max_{1 \leq j \leq n} |e_j^{(k)}| = E^{(k+1)} \leq \beta \max_{1 \leq j \leq n} |e_j^{(k)}| \quad \forall i, 1 \leq i \leq n$$

$$\text{Portanto, } \max_{1 \leq j \leq n} |e_j^{(k+1)}| = E^{(k+1)} \leq \beta \max_{1 \leq j \leq n} |e_j^{(k)}| = \beta E^{(k)}$$

Assim basta que $\beta < 1$ para que tenhamos $E^{(k+1)} < E^{(k)}$. Al\~{e}m disso, temos que $E^{(k)} \leq \beta E^{(k-1)} \leq \beta(\beta E^{(k-2)}) \leq \dots \leq \beta^k E^{(0)}$ e desde que β menor que 1, ent\~{a}o, $E^{(k)} \rightarrow 0$ quando $k \rightarrow \infty$ e, o que \~{e} importante, independentemente da aproxima\~{c}o\~{a}o inicial escolhida.

Com isso estabelecemos o crit\~{e}rio de Sassenfeld:

$$\text{Sejam } \beta_1 = \frac{|a_{12}| + |a_{13}| + \dots + |a_{1n}|}{|a_{11}|} \text{ e}$$

$$\beta_j = \frac{|a_{j1}| \beta_1 + |a_{j2}| \beta_2 + \dots + |a_{j,j-1}| \beta_{j-1} + |a_{j,j+1}| + \dots + |a_{jn}|}{|a_{jj}|}$$

$$\text{Sejam } \beta = \max_{1 \leq j \leq n} \beta_j.$$

Se $\beta < 1$, ent\~{a}o o m\~{e}todo de Gauss-Seidel gera uma sequ\~{e}ncia convergente qualquer que seja $x^{(0)}$. Al\~{e}m disso, quando menor for o β , mais r\~{a}pida ser\~{a} a converg\~{e}ncia.

Exemplo:

a) Seja o sistema linear

$$\begin{cases} x_1 + 0.5x_2 - 0.1x_3 + 0.1x_4 = 0.2 \\ 0.2x_1 + x_2 - 0.2x_3 - 0.1x_4 = -2.6 \\ -0.1x_1 - 0.2x_2 + x_3 + 0.2x_4 = 1.0 \\ 0.1x_1 + 0.3x_2 + 0.2x_3 + x_4 = -2.5 \end{cases}$$

Para este sistema linear com esta disposição de linhas e colunas, temos:

$$\beta_1 = [0.5 + 0.1 + 0.1]/1 = 0.7$$

$$\beta_2 = [(0.2)(0.7) + 0.2 + 0.1]/1 = 0.44$$

$$\beta_3 = [(0.1)(0.7) + (0.2)(0.44) + 0.2]/1 = 0.358$$

$$\beta_4 = [(0.1)(0.7) + (0.3)(0.44) + (0.2)(0.358)]/1 = 0.2736$$

Portanto, $\beta = \max_{1 \leq i \leq n} \beta_i = 0.7 < 1$ e então temos a garantia de que o método de Gauss-Seidel vai gerar uma sequência convergente.

b) Seja agora o sistema linear

$$\begin{cases} 2x_1 + x_2 + 3x_3 = 9 \\ -x_2 + x_3 = 1 \\ x_1 + 3x_3 = 3 \end{cases}$$

com essa disposição de linhas e colunas, temos

$$\beta_1 = (1 + 3)/2 = 2 > 1$$

trocando a 1ª coluna pela 3ª, temos

$$\begin{cases} x_1 + 3x_3 = 3 \\ -x_2 + x_3 = 1 \\ 2x_1 + x_2 + 3x_3 = 9 \end{cases}$$

$$\text{Donde } \beta_1 = (0 + 3)/1 = 3$$

A partir dessa disposição, trocando a 1ª coluna pela 3ª, temos

$$\begin{cases} 3x_3 + x_1 = 3 \\ x_3 - x_2 = 1 \\ 3x_3 + x_2 + 2x_1 = 9 \end{cases}$$

Dessa forma,

$$\beta_1 = 1/3$$

$$\beta_2 = [(1)(1/3) + 0]/1 = 1/3$$

$$\beta_3 = [(3)(1/3) + (1)(1/3)]/2 = 2/3$$

Portanto, $\beta = \max_{1 \leq i \leq 3} \beta_i = 2/3 < 1$; então vale o critério de Sassenfeld e temos garantia de convergência.

c) Considerando agora o exemplo usando o sistema abaixo, verificamos que o critério de Sassenfeld é apenas suficiente, pois para

$$\begin{cases} x_1 + x_1 = 3 \\ x_1 - 3x_2 = -3 \end{cases}$$

vimos que o método de Gauss-Seidel gera uma sequência convergente e, no entanto,

$$\beta_1 = 1/1 = 1$$

$$\beta_2 = [1 \times 1] = 1/3$$

e, portanto, o critério de Sassenfeld não é satisffeito.

Critério das Linhas: O critério das linhas estudado no método de Gauss-Jacobi pode ser aplicado no estudo da convergência de Gauss-Seidel.

O critério das linhas diz que se $\alpha = \max_{1 \leq k \leq n} \{\alpha_k\} < 1$, onde

$$\alpha_k = \left(\sum_{j=0, j \neq k}^n |a_{kj}| \right) / |a_{kk}|$$

então o método de Gauss-Seidel gera uma sequência convergente.

A prova da convergência consiste em verificar que se o critério das linhas for satisffeito, automaticamente o critério de Sassenfeld é satisffeito:

$$\beta_1 = (|a_{12}| + |a_{13}| + \dots + |a_{1n}|) / |a_{11}| = \alpha_1 < 1$$

e para $i = 2, \dots, k-1$, supor por indução que $\beta_i \leq \alpha < 1$. Então

$$\beta_k = (\beta_1 |a_{k1}| + \dots + \beta_{k-1} |a_{k,k-1}| + |a_{k,k+1}| + |a_{kn}|) / |a_{kk}| < (|a_{k1}| + \dots + |a_{k,k-1}| + |a_{k,k+1}| + \dots + |a_{kn}|) / |a_{kk}| = \alpha_k$$

Assim, $\beta_i \leq \alpha_i, i = 1, \dots, n$.

Então, $\alpha_i < 1$ implica que $\beta_i < 1, i = 1, \dots, n$, ou seja, o critério de Sassenfeld é satisffeito.

Observamos, no entanto, que o critério de Sassenfeld pode ser satisffeito mesmo que o critério das linhas não o seja.

$$\text{Exemplo } \begin{cases} 3x_3 + x_1 = 3 \\ x_3 - x_2 = 1 \\ 3x_3 + x_2 + 2x_1 = 9 \end{cases}$$

Temos, $\alpha_1 = \beta_1 = 1/3 < 1$ e

$\alpha_2 = 1/1 = 1$; então o critério das linhas não é satisffeito.

No entanto,

$$\beta_2 = \frac{1 \times 1/3}{1} = \frac{1}{3} < 1$$

$$\beta_3 = \frac{3 \times 1/3 + 1/3}{2} = 2/3 < 1.$$

Portanto o critério de Sassenfeld é satisfeito.

4.2 Método SOR

Veremos agora uma variante do método de Gauss-Seidel, denominado métodos SOR (ou também conhecida com o nome de relaxação sucessiva) que consiste em considerar um parâmetro real ω não nulo: No método SOR se escolhem as matrizes M e N como:

$$M = L + \frac{1}{\omega} \text{ e } N = (1 - \frac{1}{\omega})D + U$$

A utilização deste parâmetro permite obter uma convergência mais rápida, havendo um valor ω que é o parâmetro optimal, no sentido em que minimiza o raio espectral da matriz $C = M^{-1}N$

Portanto o método SOR é definido por:

$$x^{(k+1)} = (1 - \omega)x_1^{(k)} = (1 - \omega)x^{(k)} + \omega D^{-1}(b - Lx^{(k+1)} - Ux^{(k)}) \text{ com } k = 0, 1, 2, \dots$$

Onde $x^{(0)}$ é um vetor dado.

O método SOR pode ser expresso pela fórmula.

$$x_1^{(k+1)} = (1 - \omega)x_1^{(k)} + \frac{\omega}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right), i = 1, 2, 3, \dots, n$$

Notando que no caso em que $\omega = 1$ temos o método de Gauss-Seidel. A escolha $1 < \omega < 2$ caracteriza os métodos de sobre-relaxação, ao passo que os métodos de sub-relaxação são obtidos por valores $0 < \omega < 1$. A prática mostra que melhores resultados são obtidos na sobre-relação.

Exemplo: Seja o sistema linear $\begin{pmatrix} 5 & 1 & 1 \\ 3 & 4 & 1 \\ 3 & 3 & 6 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \\ 0 \end{pmatrix}$ cuja a solução exata

é: $x_1 = 1, x_2 = 1 \text{ e } x_3 = -1$

Aplicando o método SOR $x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \frac{\omega}{a_{ii}}(b_i - \sum_{j=1}^n a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)})$ ao sistema acima com $\omega = 0.9$, teremos:

$$x_1^{(k+1)} = (1 - \omega)x_1^{(k)} + \frac{\omega}{a_{11}}[b_1 - (a_{12}x_2^{(k)} + a_{13}x_3^{(k)})] = 0.1x_1^{(k)} + \frac{0.9}{5}[5 - (x_2^{(k)} + x_3^{(k)})]$$

$$x_1^{(k+1)} = 0.1x_1^{(k)} + 0.9 - 0.18x_2^{(k)} - 0.18x_3^{(k)}$$

$$x_2^{(k+1)} = (1 - \omega)x_2^{(k)} + \frac{\omega}{a_{22}}[b_2 - (a_{21}x_1^{(k+1)} + a_{23}x_3^{(k)})] = 0.1x_2^{(k)} + \frac{0.9}{4}[6 - (3x_1^{(k)} + x_3^{(k)})]$$

$$x_2^{(k+1)} = 0.1x_2^{(k)} + 1,35 - 0.675x_1^{(k+1)} - 0.225x_3^{(k)}$$

$$x_3^{(k+1)} = (1 - \omega)x_3^{(k)} + \frac{\omega}{a_{33}}[b_3 - (a_{31}x_1^{(k+1)} + a_{32}x_2^{(k+1)})]$$

$$x_3^{(k+1)} = 0.1x_3^{(k)} + \frac{0.9}{6}[0 - (3x_1^{(k+1)} + 3x_2^{(k+1)})] = 0.1x_3^{(k)} - 0.45x_1^{(k+1)} - 0.45x_2^{(k+1)}$$

Tomando $x^{(0)} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$

Para $k = 0$, vamos calcular $x^{(1)} = \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} \end{pmatrix}^T$

$$x_1^{(1)} = 0.1x_1^{(0)} + 0.9 - 0.18x_2^{(0)} - 0.18x_3^{(0)} = 0.1(0) + 0.9 - 0.18(0) - 0.18(0)$$

$$x_1^{(1)} = 0.9$$

$$x_2^{(1)} = 0.1x_1^{(0)} + 1.35 - 0.675x_1^{(1)} - 0.225x_3^{(0)} = 0.1(0) + 1,35 - 0,675(0,9) - 0,255(0)$$

$$x_2^{(1)} = 0,7425$$

$$x_3^{(1)} = 0.1x_3^{(0)} - 0,45x_1^{(1)} - 0,45x_2^{(1)} = 0,1(0) - 0,45(0,9) - 0,45(0,7425)$$

$$x_3^{(1)} = -0,7391$$

Portanto, $x^{(1)} = \begin{pmatrix} 0,9 & 0,7425 & -0,7391 \end{pmatrix}^T$

Para $k = 1$, vamos calcular $x^{(2)} = [x_1^{(2)} x_2^{(2)} x_3^{(2)}]^T$

$$x_1^{(2)} = 0,1x_1^{(1)} + 0,9 - 0,18x_2^{(1)} - 0,18x_3^{(1)} = 0,1(0,9) + 0,9 - 0,18(0,7425) - 0,18(-0,7391) = 0,9894$$

$$x_1^{(2)} = 0,9894$$

$$x_2^{(2)} = 0,1x_2^{(1)} + 1,35 - 0,675x_1^{(2)} - 0,225x_3^{(1)} = 0,1(0,7425) + 1,35 - 0,675(0,9894) -$$

$$0,225(-0,7391) = 0,9227$$

$$x_2^{(2)} = 0,9227$$

$$x_3^{(2)} = 0,1x_3^{(1)} - 0,45x_1^{(2)} - 0,45x_2^{(2)} = 0,1(-0,7391) - 0,45(0,9894) - 0,45(0,9227) = -0,9344$$

$$x_3^{(2)} = -0,9344$$

$$\text{Portanto, } x^{(2)} = \begin{bmatrix} 0,9894 & 0,9227 & -0,9344 \end{bmatrix}^T$$

Teorema 4.2.1 : (Condição Necessária) Para que haja convergência do método SOR, qualquer que seja a iterada inicial, é necessário que $0 < \omega < 2$.

Demonstração: A matriz de iteração do método SOR é $C = -M^{-1}N$

$$\begin{aligned} C &= \left(\frac{1}{\omega}D - L\right)^{-1}\left(\frac{1-\omega}{\omega}D + U\right) = \left[\frac{1}{\omega}D(I - \omega D^{-1}L)\right]^{-1}\left(\frac{1-\omega}{\omega}D + U\right) \\ &= (I - \omega D^{-1}L)^{-1}\omega D^{-1}\left(\frac{1-\omega}{\omega}D + U\right) \end{aligned}$$

ou,

$$C = (I - \omega D^{-1}L)^{-1}[(1 - \omega)I + \omega D^{-1}U]$$

se $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$ são os autovalores de C, então

$\det R = \lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$. Mas,

$$\det R = \det\{(I - \omega D^{-1}L)^{-1}[(1 - \omega)I + \omega D^{-1}U]\} = \det\{(I - \omega D^{-1}L)^{-1}\det[(1 - \omega)I + \omega D^{-1}U]\} = (1 - \omega)^n$$

pois, $I - \omega D^{-1}L$ é uma matriz triangular inferior com elementos iguais a 1 na diagonal e $(1 - \omega)I + \omega D^{-1}U$ é uma matriz triangular superior com elementos iguais a $1 - \omega$ na diagonal principal. Logo:

$$\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n = (1 - \omega)^n$$

Em particular, pelo menos um dos autovetores λ_1 de R deve satisfazer $|\lambda_1| \geq |1 - \omega|$.

Mas, se o método SOR converge, devemos ter também $\lambda < 1$ para todo autovetor λ de R. Logo

$$|1 - \omega| < 1$$

onde $0 < \omega < 2$ □

Teorema 4.2.2 (Condição Suficiente) *Se A é simétrica e positiva definida, então o método converge para qualquer $0 < \omega < 2$; em particular como vimos anteriormente, neste caso o método de Gauss-Seidel é convergente.*

Capítulo 5

Aplicação e Comparação dos Métodos Diretos e Iterativos

Neste capítulo vamos determinar o vetor solução do sistema abaixo, através dos métodos já estudados, primeiramente vamos aplicar os Métodos Diretos e posteriormente os Métodos Iterativos que foram abordados nesse trabalho. Fazendo uma comparação entre os mesmos.

Seja o sistema:

$$\begin{cases} 4x_1 + x_2 + x_3 + x_4 = 7 \\ 2x_1 - 8x_2 + x_3 - x_4 = -6 \\ x_1 + 2x_2 - 5x_3 + x_4 = -1 \\ x_1 + x_2 + x_3 - 4x_4 = -1 \end{cases} \sim \left(\begin{array}{cccc|c} 4 & 1 & 1 & 1 & 7 \\ 2 & -8 & 1 & -1 & -6 \\ 1 & 2 & -5 & 1 & -1 \\ 1 & 1 & 1 & -4 & -1 \end{array} \right)$$

5.1 Eliminação de Gauss

onde temos que:

$$m_{21} = 1/2 = 0,5$$

$$m_{31} = 1/4 = 0,25$$

$$m_{41} = 1/4 = 0,25$$

Daí, temos:

$$\left(\begin{array}{cccc|c} 4 & 1 & 1 & 1 & 7 \\ 0 & -8,5 & 0,5 & -1,5 & -9,5 \\ 0 & 1,75 & -5,25 & 0,75 & -2,75 \\ 0 & 0,75 & 0,75 & -4,25 & -2,75 \end{array} \right)$$

De onde vem:

$$m_{32} = \frac{1,75}{-8,5} = -0,2058$$

$$m_{42} = \frac{0,75}{-8,5} = -0,0882$$

logo, temos:

$$\left(\begin{array}{cccc|c} 4 & 1 & 1 & 1 & 7 \\ 0 & -8,5 & 0,5 & -1,5 & -9,5 \\ 0 & 0 & -5,0442 & 0,9558 & -4,7051 \\ 0 & 0 & 0,7941 & -4,3823 & -3,5879 \end{array} \right)$$

Daí, temos:

$$m_{43} = \frac{0,7941}{-5,0442} = -0,1574$$

Com isso, teremos então a matriz escalonada, que será resolvida por retro-substituição.

$$\left(\begin{array}{cccc|c} 4 & 1 & 1 & 1 & 7 \\ 0 & -8,5 & 0,5 & -1,5 & -9,5 \\ 0 & 0 & -5,0442 & 0,9558 & -4,7051 \\ 0 & 0 & 0 & -4,2318 & -4,3284 \end{array} \right)$$

Que é equivalente ao sistema abaixo:

$$\left\{ \begin{array}{l} 4x_1 + x_2 + x_3 + x_4 = 7 \\ -8,5x_2 + 0,5x_3 - 1,5x_4 = -9,5 \\ -5,0442x_3 + 0,9558x_4 = -4,7051 \\ -4,2318x_4 = -4,3284 \end{array} \right.$$

Onde:

$$x_4 = 1,0228$$

$$x_3 = 1,1265$$

$$x_2 = 0,9850$$

$$x_1 = 0,9850$$

$$\Rightarrow x = \left(0,9664 \quad 0,9850 \quad 1,1265 \quad 1,0228 \right)^T$$

Observe que trabalhamos com arredondamento (aproximações), logo o vetor solução será aproximado. Poderíamos melhorar a solução utilizando a técnica de refinamento de soluções.

5.2 Fatoração LU

Usando o sistema anterior.

$$\begin{cases} 4x_1 + x_2 + x_3 + x_4 = 7 \\ 2x_1 - 8x_2 + x_3 - x_4 = -6 \\ x_1 + 2x_2 - 5x_3 + x_4 = -1 \\ x_1 + x_2 + x_3 - 4x_4 = -1 \end{cases} \sim \left(\begin{array}{cccc|c} 4 & 1 & 1 & 1 & 7 \\ 2 & -8 & 1 & -1 & -6 \\ 1 & 2 & -5 & 1 & -1 \\ 1 & 1 & 1 & -4 & -1 \end{array} \right) = A$$

Cuja a solução exata é $x = [1 \ 1 \ 1 \ 1]^T$

Temos que decompor a matriz A dos coeficientes da seguinte forma: $A = L.U$ Onde: L = matriz triangular inferior, ou seja, $L = (M_{ij})$, com $m_{ii} = 1$, $1 \leq i \leq n$

U = matriz triangular superior, onde, $u = (a_{ij})$

Observe que a matriz A dos coeficientes, já foi escalonada pelo método da Eliminação Gaussiana. Logo, podemos obter as matrizes L e U como se segue:

$$A = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0,5 & 1 & 0 & 0 \\ 0,25 & -0,2058 & 1 & 0 \\ 0,25 & -0,0882 & -0,1574 & 1 \end{pmatrix}}_L \cdot \underbrace{\begin{pmatrix} 4 & 1 & 1 & 1 \\ 0 & -8,5 & 0,5 & -1,5 \\ 0 & 0 & -5,0442 & 0,9558 \\ 0 & 0 & 0 & -4,2318 \end{pmatrix}}_U$$

$Ax = b \Leftrightarrow (LU)x = b$. Seja $y = Ux$, então:

$Ly = b$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0,5 & 1 & 0 & 0 \\ 0,25 & -0,2058 & 1 & 0 \\ 0,25 & -0,0882 & -0,1574 & 1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 7 \\ -6 \\ -1 \\ -1 \end{pmatrix}$$

Resolvendo essa igualdade obteremos:

$$y_1 = 7$$

$$0,5y_1 + y_2 = -6 \Rightarrow y_2 = -9,5$$

$$0,25y_1 - 0,2058y_2 + y_3 = -1 \Rightarrow y_3 = -4,7051$$

$$0,25y_1 - 0,0882y_2 - 0,1574y_3 + y_4 = -1 \Rightarrow y_4 = -4,3284$$

$Ux = y$

$$\begin{pmatrix} 4 & 1 & 1 & 1 \\ 0 & -8,5 & 0,5 & -1,5 \\ 0 & 0 & -5,0442 & 0,9558 \\ 0 & 0 & 0 & -4,2318 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 7 \\ -9,5 \\ -4,7051 \\ -4,3284 \end{pmatrix}$$

$$\begin{cases} 4x_1 + x_2 + x_3 + x_4 = 7 \\ -8,5x_2 + 0,5x_3 - 1,5x_4 = -9,5 \\ -5,0442x_3 + 0,9558x_4 = -4,7051 \\ -4,2318x_4 = -4,3284 \end{cases}$$

Resolvendo o sistema, teremos:

$$x_1 = 0,9618$$

$$x_2 = 1,0034$$

$$x_3 = 1,1265$$

$$x_4 = 1,0228$$

Observe que a resolução, trabalhamos com aproximações. Logo o vetor solução também é uma aproximação da solução exata do sistema.

$$x = \left(0,9618 \quad 1,0034 \quad 1,1265 \quad 1,0228 \right)^T$$

Poderíamos melhorar este resultado aplicando a mesma técnica comentada na eliminação gaussiana (Técnica de Refinamento de Soluções).

5.3 Fatoração de Cholesky

Observe que a matriz A dos coeficientes não é simétrica, ou seja, não atende a condição do teorema 4.0.3. Isso significa que não podemos decompor a matriz A em $A = GG^T$ em que G é uma matriz triangular superior positiva. Portanto a Fatoração de Cholesky não é aplicável no sistema linear trabalhado.

Observação sobre os Métodos Diretos: O sistema adotado é de pequeno porte e apresenta densidade na sua matriz dos coeficientes (não apresenta coeficientes nulos) logo, os métodos diretos (Eliminação de Gauss e Fatoração LU) mostraram-se bastante satisfatórios, uma vez que conseguimos obter o vetor solução exato, não fosse os erros de arredondamento.

O sistema não pôde ser resolvido através da Fatoração de Cholesky uma vez que a matriz dos coeficientes não atende a condição de ser simétrica e definida positiva.

5.4 Método Iterativo de Gauss-Jacobi

Vamos determinar o vetor solução do sistema abaixo pelo Método de Jacobi, com o vetor solução inicial $x^{(0)} = \left(0 \quad 0 \quad 0 \quad 0 \right)^T$ e $\epsilon < 10^{-2}$

$$\begin{cases} 4x_1 + x_2 + x_3 + x_4 = 7 \\ 2x_1 - 8x_2 + x_3 - x_4 = -6 \\ x_1 + 2x_2 - 5x_3 + x_4 = -1 \\ x_1 + x_2 + x_3 - 4x_4 = -1 \end{cases}$$

Note que o sistema satisfaz o Critério das Linhas, o que nos assegura a convergência do sistema linear para qualquer que seja $x^{(0)}$.

As iterações podem ser obtidas pela fórmula

$$x_i^{k+1} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=0, j \neq i}^n a_{ij} x_j^{(k)} \right)$$

1ª iteração:

$$x_1^{(1)} = \frac{1}{4}(7 - 1.0 - 1.0 - 1.0) = 1,75$$

$$x_2^{(1)} = \frac{1}{-8}(-6 - 2.0 - 1.0 + 1.0) = 0,75$$

$$x_3^{(1)} = \frac{1}{-5}(-1 - 1.0 - 2.0 - 1.0) = 0,2$$

$$x_4^{(1)} = \frac{1}{-4}(-1 - 1.0 - 1.0 - 1.0) = 0,25$$

$$\therefore x^{(1)} = \left(1,75 \quad 0,75 \quad 0,2 \quad 0,25 \right)^T$$

2ª iteração:

$$x_1^{(2)} = \frac{1}{4}(7 - 1.0,75 - 1.0,2 - 1.0,25) = 1,45$$

$$x_2^{(2)} = \frac{1}{-8}(-6 - 2.1,75 - 1.0,2 + 1.0,25) = 1,1812$$

$$x_3^{(2)} = \frac{1}{-5}(-1 - 1.1,75 - 2.0,75 - 1.0,25) = 0,9$$

$$x_4^{(2)} = \frac{1}{-4}(-1 - 1.1,75 - 1.0,75 - 1.0,2) = 0,925$$

$$\therefore x^{(2)} = \left(1,45 \quad 1,1812 \quad 0,9 \quad 0,925 \right)^T$$

3ª iteração:

$$x_1^{(3)} = \frac{1}{4}(7 - 1.1,1812 - 1.0,9 - 1.0,925) = 0,9984$$

$$x_2^{(3)} = \frac{1}{-8}(-6 - 2.1,45 - 1.0,9 + 1.0,925) = 1,1093$$

$$x_3^{(3)} = \frac{1}{-5}(-1 - 1.1,45 - 2.1,1812 - 1.0,925) = 1,1474$$

$$x_4^{(3)} = \frac{1}{-4}(-1 - 1.1,45 - 1.1,1812 - 1.0,9) = 1,1328$$

$$\therefore x^{(3)} = \left(0,9984 \quad 1,1093 \quad 1,1474 \quad 1,1328 \right)^T$$

4ª iteração:

$$x_1^{(4)} = \frac{1}{4}(7 - 1.1, 1093 - 1.1, 1474 - 1.1, 1328) = 0,9026$$

$$x_2^{(4)} = \frac{1}{-8}(-6 - 2.0, 9984 - 1.1, 1474 + 1.1, 1328) = 1,0014$$

$$x_3^{(4)} = \frac{1}{-5}(-1 - 1.0, 9984 - 2.1, 1093 - 1.0, 1328) = 1,0699$$

$$x_4^{(4)} = \frac{1}{-4}(-1 - 1.0, 9984 - 1.1, 1093 - 1.0, 1474) = 1,0637$$

$$\therefore x^{(4)} = \left(0,9026 \quad 1,0014 \quad 1,0699 \quad 1,0637 \right)^T$$

5ª iteração:

$$x_1^{(5)} = \frac{1}{4}(7 - 1.1, 0014 - 1.1, 0699 - 1.1, 0637) = 0,9662$$

$$x_2^{(5)} = \frac{1}{-8}(-6 - 2.0, 9026 - 1.1, 0699 + 1.1, 0637) = 0,9764$$

$$x_3^{(5)} = \frac{1}{-5}(-1 - 1.0, 9026 - 2.1, 0014 - 1.1, 0637) = 0,9938$$

$$x_4^{(5)} = \frac{1}{-4}(-1 - 1.0, 9026 - 1.1, 0014 - 1.1, 0699) = 0,9934$$

$$\therefore x^{(5)} = \left(0,9662 \quad 0,9764 \quad 0,9938 \quad 0,9934 \right)^T$$

6ª iteração:

$$x_1^{(6)} = \frac{1}{4}(7 - 1.0, 9764 - 1.0, 9938 - 1.0, 9934) = 1,0091$$

$$x_2^{(6)} = \frac{1}{-8}(-6 - 2.0, 9662 - 1.0, 9938 + 1.0, 9934) = 0,9916$$

$$x_3^{(6)} = \frac{1}{-5}(-1 - 1.0, 9662 - 2.0, 9764 - 1.0, 9934) = 0,9824$$

$$x_4^{(6)} = \frac{1}{-4}(-1 - 1.0, 9662 - 1.0, 9764 - 1.0, 9938) = 0,9841$$

$$\therefore x^{(6)} = \left(1,0091 \quad 0,9916 \quad 0,9824 \quad 0,9841 \right)^T$$

7ª iteração:

$$x_1^{(7)} = \frac{1}{4}(7 - 1.0, 9916 - 1.0, 9824 - 1.0, 9841) = 1,0104$$

$$x_2^{(7)} = \frac{1}{-8}(-6 - 2.1, 0091 - 1.0, 9824 + 1.0, 9841) = 1,0020$$

$$x_3^{(7)} = \frac{1}{-5}(-1 - 1.1, 0091 - 2.0, 9916 - 1.0, 9841) = 0, 9952$$

$$x_4^{(7)} = \frac{1}{-4}(-1 - 1.1, 0091 - 1.0, 9916 - 1.0, 9938) = 0, 9957$$

$$\therefore x^{(7)} = \left(1, 0104 \quad 1, 0020 \quad 0, 9952 \quad 0, 9957 \right)^T$$

8ª iteração:

$$x_1^{(8)} = \frac{1}{4}(7 - 1.1, 0020 - 1.0, 9952 - 1.0, 9957) = 1, 0017$$

$$x_2^{(8)} = \frac{1}{-8}(-6 - 2.1, 0104 - 1.0, 9952 + 1.0, 9957) = 1, 0025$$

$$x_3^{(8)} = \frac{1}{-5}(-1 - 1.1, 0104 - 2.1, 0020 - 1.0, 9957) = 1, 0020$$

$$x_4^{(8)} = \frac{1}{-4}(-1 - 1.1, 0104 - 1.1, 0020 - 1.0, 9952) = 1, 0019$$

$$\therefore x^{(8)} = \left(1, 0017 \quad 1, 0025 \quad 1, 0020 \quad 1, 0019 \right)^T$$

Teste de Parada:

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

$$|x_1^{(8)} - x_1^{(7)}| = 0, 0087$$

$$|x_2^{(8)} - x_2^{(7)}| = 0, 0005$$

$$|x_3^{(8)} - x_3^{(7)}| = 0, 0068 \Rightarrow d_r^{(8)} = \frac{0, 0087}{1, 0025} = 0, 0086 < \epsilon = 10^{-2}$$

$$|x_4^{(8)} - x_4^{(7)}| = 0, 0062$$

Note que $d^{(10)} < \epsilon$, ou seja, com 8 iterações obtemos o vetor:

$$x^{(8)} = \left(1, 0017 \quad 1, 0025 \quad 1, 0020 \quad 1, 0019 \right)^T$$

Que é uma aproximação da solução exata e que está de acordo com a precisão desejada.

A tabela a seguir mostra os resultados dos vetores obtidos até a 10ª iteração.

Método de Gauss-Jacobi				
k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$
1	1,75	1,1875	1,025	1,2406
2	1,45	1,1812	0,9	0,925
3	0,9984	1,1093	1,1474	1,1328
4	0,9026	1,0014	1,0699	1,0637
5	0,9984	1,1093	1,1474	1,1328
6	1,0091	0,9916	0,9824	0,9841
7	1,0104	1,0020	0,9952	0,9957
8	1,0017	1,0025	1,0020	1,0019
9	0,9984	1,0004	1,0017	1,0015
10	0,9991	0,9996	1,0001	1,0001

5.5 Método Iterativo de Gauss-seidel

Iremos determinar o vetor solução do sistema abaixo, através do método iterativo de Gauss-Seidel.

$$\begin{cases} 4x_1 + x_2 + x_3 + x_4 = 7 \\ 2x_1 - 8x_2 + x_3 - x_4 = -6 \\ x_1 + 2x_2 - 5x_3 + x_4 = -1 \\ x_1 + x_2 + x_3 - 4x_4 = -1 \end{cases}$$

Seja o vetor solução inicial $x^{(0)} = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}^T$ e $\epsilon < 10^{-2}$

Este método tem como forma geral:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right)$$

1ª iteração:

$$x_1^{(1)} = \frac{1}{4}(7 - 1.0 - 1.0 - 1.0) = 1,75$$

$$x_2^{(1)} = \frac{1}{-8}(-6 - 2.1,75 - 1.0 + 1.0) = 1,1875$$

$$x_3^{(1)} = \frac{1}{-5}(-1 - 1.1,75 - 2.1,1875 - 1.0) = 1,025$$

$$x_4^{(1)} = \frac{1}{-4}(-1 - 1.1,75 - 1.1,1875 - 1.1,025) = 1,2406$$

$$\therefore x^{(1)} = \begin{pmatrix} 1,75 & 1,1875 & 1,025 & 1,2406 \end{pmatrix}^T$$

2ª iteração:

$$x_1^{(2)} = \frac{1}{4}(7 - 1.1, 1875 - 1.1, 025 - 1.1, 2406) = 0,8867$$

$$x_2^{(2)} = \frac{1}{-8}(-6 - 2.0, 8867 - 1.1, 025 + 1.1, 2406) = 0,9447$$

$$x_3^{(2)} = \frac{1}{-5}(-1 - 1.0, 8867 - 2.0, 9447 - 1.1, 2406) = 0,9994$$

$$x_4^{(2)} = \frac{1}{-4}(-1 - 1.0, 8867 - 1.0, 9447 - 1.0, 9994) = 0,9577$$

$$\therefore x^{(2)} = \left(0,8867 \quad 0,9447 \quad 0,9994 \quad 0,9577 \right)^T$$

3ª iteração:

$$x_1^{(3)} = \frac{1}{4}(7 - 1.0, 9447 - 1.0, 9994 - 1.0, 9577) = 1,0245$$

$$x_2^{(3)} = \frac{1}{-8}(-6 - 2.1, 0245 - 1.0, 9994 + 1.0, 9577) = 1,0113$$

$$x_3^{(3)} = \frac{1}{-5}(-1 - 1.1, 0245 - 2.1, 0113 - 1.0, 9577) = 1,0009$$

$$x_4^{(3)} = \frac{1}{-4}(-1 - 1.1, 0245 - 1.1, 0113 - 1.1, 0009) = 1,0091$$

$$\therefore x^{(3)} = \left(1,0245 \quad 1,0113 \quad 1,0009 \quad 1,0091 \right)^T$$

4ª iteração:

$$x_1^{(4)} = \frac{1}{4}(7 - 1.1, 0113 - 1.1, 0009 - 1.1, 0091) = 0,9946$$

$$x_2^{(4)} = \frac{1}{-8}(-6 - 2.0, 9946 - 1.1, 0009 + 1.1, 0091) = 0,9976$$

$$x_3^{(4)} = \frac{1}{-5}(-1 - 1.0, 9946 - 2.0, 9976 - 1.1, 0091) = 0,9997$$

$$x_4^{(4)} = \frac{1}{-4}(-1 - 1.0, 9946 - 1.0, 9976 - 1.0, 9997) = 0,9979$$

$$\therefore x^{(4)} = \left(0,9946 \quad 0,9976 \quad 0,9997 \quad 0,9979 \right)^T$$

5ª iteração:

$$x_1^{(5)} = \frac{1}{4}(7 - 1.0, 9976 - 1.0, 9997 - 1.0, 9979) = 1,0012$$

$$x_2^{(5)} = \frac{1}{-8}(-6 - 2.1, 0012 - 1.0, 9997 + 1.0, 9979) = 1,0005$$

$$x_3^{(5)} = \frac{1}{-5}(-1 - 1.1, 0012 - 2.1, 0005 - 1.0, 9979) = 1,0000$$

$$x_4^{(5)} = \frac{1}{-4}(-1 - 1.1, 0012 - 1.1, 0005 - 1.1, 0000) = 1, 0004$$

$$\therefore x^{(5)} = \left(1, 0012 \quad 1, 0005 \quad 1, 0000 \quad 1, 0004 \right)^T$$

Teste de Parada:

Aplicando o teste de parada, teremos:

$$|x_1^{(5)} - x_1^{(4)}| = 0, 0066$$

$$|x_2^{(5)} - x_2^{(4)}| = 0, 0029$$

$$|x_3^{(5)} - x_3^{(4)}| = 0, 0003 \Rightarrow d_r^{(5)} = \frac{0, 0066}{\max_{1 \leq i \leq 4} |x_i^{(5)}|} = 0, 00659 < \epsilon = 10^{-2}$$

$$|x_4^{(5)} - x_4^{(4)}| = 0, 0025$$

Note que o vetor obtido na 5ª iteração, já está de acordo com a precisão desejada.

A tabela a seguir mostra os resultados dos vetores obtidos até a 10ª iteração.

Método de Gauss-Seidel				
k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$
1	1,75	1,1875	1,025	1,2406
2	0,8867	0,9447	0,9994	0,9577
3	1,0245	1,0113	1,0009	1,0091
4	0,9946	0,9976	0,9997	0,9979
5	1,0012	1,0005	1,0000	1,0004
6	0,9997	0,9998	0,9999	0,9998
7	1,0001	1,0000	0,9999	1,0000
8	1,0000	0,9999	0,9999	0,9999
9	1,0000	1,0000	0,9999	0,9999
10	1,0000	1,0000	1,0000	1,0000

5.6 Método SOR ou Relaxação Sucessiva

Vamos determinar o vetor solução do sistema linear abaixo, através do Método SOR.

Vamos adotar o Método da Sub-relaxação, ou seja, $0 < \omega < 1$. Adotaremos $\omega = 0,9$

$$\begin{cases} 4x_1 + x_2 + x_3 + x_4 = 7 \\ 2x_1 - 8x_2 + x_3 - x_4 = -6 \\ x_1 + 2x_2 - 5x_3 + x_4 = -1 \\ x_1 + x_2 + x_3 - 4x_4 = -1 \end{cases}$$

Aplicando a fórmula do Método SOR:

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \frac{\omega}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right)$$

$$x_1^{(k+1)} = (1 - 0,9)x_1^{(k)} + \frac{0,9}{4}[7 - (x_2^{(k)} + x_3^{(k)} + x_4^{(k)})] = 0,1x_1^{(k)} + 0,225[7 - (x_2^{(k)} + x_3^{(k)} + x_4^{(k)})]$$

$$x_1^{(k+1)} = 0,1x_1^{(k)} + 1,575 - 0,225x_2^{(k)} - 0,225x_3^{(k)} - 0,225x_4^{(k)}$$

$$x_2^{(k+1)} = (1 - 0,9)x_2^{(k)} - \frac{0,9}{8}[-6 - (2x_1^{(k)} - x_3^{(k)} + x_4^{(k)})] = 0,1x_2^{(k)} - 0,1125[-6 - (2x_1^{(k)} - x_3^{(k)} + x_4^{(k)})]$$

$$x_2^{(k+1)} = 0,1x_2^{(k)} + 0,675 + 0,225x_1^{(k)} + 0,1125x_3^{(k)} - 0,1125x_4^{(k)}$$

$$x_3^{(k+1)} = (1 - 0,9)x_3^{(k)} - \frac{0,9}{5}[-1 - (x_1^{(k)} + 2x_2^{(k)} + x_4^{(k)})] = 0,1x_3^{(k)} - 0,18[-1 - (x_1^{(k)} + 2x_2^{(k)} + x_4^{(k)})]$$

$$x_3^{(k+1)} = 0,1x_3^{(k)} + 1,18 + 0,18x_1^{(k)} + 0,36x_2^{(k)} + 0,18x_4^{(k)}$$

$$x_4^{(k+1)} = (1 - 0,9)x_4^{(k)} - \frac{0,9}{4}[-1 - (x_1^{(k)} + x_2^{(k)} + x_3^{(k)})] = 0,1x_4^{(k)} - 0,225[-1 - (x_1^{(k)} + x_2^{(k)} + x_3^{(k)})]$$

$$x_4^{(k+1)} = 0,1x_4^{(k)} + 0,225 + 0,225x_1^{(k)} + 0,225x_2^{(k)} + 0,255x_3^{(k)}$$

Tomando $x^{(0)} = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}^T$

Para $k = 0$, vamos calcular $x^{(1)} = \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} & x_4^{(1)} \end{pmatrix}^T$

$$x_1^{(1)} = 0,1 \cdot 0 + 1,575 - 0,225 \cdot 0 - 0,225 \cdot 0 - 0,225 \cdot 0 = 1,575$$

$$x_2^{(1)} = 0,1 \cdot 0 + 0,675 + 0,225 \cdot 1,575 + 0,1125 \cdot 0 - 0,1125 \cdot 0 = 1,0293$$

$$x_3^{(1)} = 0,1 \cdot 0 + 1,18 + 0,18 \cdot 1,575 + 0,36 \cdot 1,0293 + 0,18 \cdot 0 = 0,8340$$

$$x_4^{(1)} = 0,1 \cdot 0 + 0,225 + 0,225 \cdot 1,575 + 0,225 \cdot 1,0293 + 0,255 \cdot 0,8340 = 0,9986$$

$$\therefore x^{(1)} = \begin{pmatrix} 1,575 & 1,0293 & 0,8340 & 0,9986 \end{pmatrix}^T$$

Para $k = 1$, vamos calcular $x^{(2)} = \begin{pmatrix} x_1^{(2)} & x_2^{(2)} & x_3^{(2)} & x_4^{(2)} \end{pmatrix}^T$

$$x_1^{(2)} = 0,1 \cdot 1,575 + 1,575 - 0,225 \cdot 1,0293 - 0,225 \cdot 0,8340 - 0,225 \cdot 0,9986 = 1,0885$$

$$x_2^{(2)} = 0,1 \cdot 1,0293 + 0,675 + 0,225 \cdot 1,0885 + 0,1125 \cdot 0,8340 - 0,1125 \cdot 0,9986 = 1,0043$$

$$x_3^{(2)} = 0,1 \cdot 0,8340 + 1,18 + 0,18 \cdot 1,0885 + 0,36 \cdot 1,0043 + 0,18 \cdot 0,9986 = 1,0006$$

$$x_4^{(2)} = 0,1.0,9986 + 0,225 + 0,225.1,0885 + 0,225.1,0043 + 0,255.1,0006 = 1,0208$$

$$\therefore x^{(2)} = \left(1,0885 \quad 1,0043 \quad 1,0006 \quad 1,0208 \right)^T$$

Para $k = 2$, vamos calcular $x^{(3)} = \left(x_1^{(3)} \quad x_2^{(3)} \quad x_3^{(3)} \quad x_4^{(3)} \right)^T$

$$x_1^{(3)} = 0,1.1,0885 + 1,575 - 0,255.1,0043 - 0,225.1,0006 - 0,255.1,0208 = 1,0030$$

$$x_2^{(3)} = 0,1.1,0043 + 0,675 + 0,255.1,0030 + 0,1125.1,0006 - 0,1125.1,0208 = 0,9988$$

$$x_3^{(3)} = 0,1.1,0006 + 0,18 + 0,18.1,0030 + 0,36.0,9988 + 0,18.1,0208 = 1,0039$$

$$x_4^{(3)} = 0,1.1,0208 + 0,225 + 0,225.1,0030 + 0,225.0,9988 + 0,255.1,0039 = 1,0033$$

$$\therefore x^{(3)} = \left(1,0030 \quad 0,9988 \quad 1,0039 \quad 1,0033 \right)^T$$

Para $k = 3$, vamos calcular $x^{(4)} = \left(x_1^{(4)} \quad x_2^{(4)} \quad x_3^{(4)} \quad x_4^{(4)} \right)^T$

$$x_1^{(4)} = 0,1.1,0030 + 1,575 - 0,225.0,9988 - 0,225.1,0039 - 0,225.1,0033 = 0,9989$$

$$x_2^{(4)} = 0,1.0,9988 + 0,675 + 0,225.0,9989 + 0,1125.1,0039 - 0,1125.1,0033 = 0,9997$$

$$x_3^{(4)} = 0,1.1,0039 + 0,18 + 0,18.0,9989 + 0,36.0,9997 + 0,18.1,0033 = 1,0006$$

$$x_4^{(4)} = 0,1.1,0033 + 0,225 + 0,225.0,9989 + 0,225.0,9997 + 0,225.1,0006 = 1,0001$$

$$\therefore x^{(4)} = \left(0,9989 \quad 0,9987 \quad 1,0006 \quad 1,0001 \right)^T$$

Teste de parada:

Aplicando o teste de parada, teremos:

$$|x_1^{(4)} - x_1^{(3)}| = 0,0041$$

$$|x_2^{(4)} - x_2^{(3)}| = 0,0009$$

$$|x_3^{(4)} - x_3^{(3)}| = 0,0033 \Rightarrow d_r^{(4)} = \frac{0,0041}{\max_{1 \leq i \leq 4} |x_i^{(4)}|} = 0,0040 < \epsilon = 10^{-2}$$

$$|x_4^{(4)} - x_4^{(3)}| = 0,0032$$

A tabela a seguir mostra os resultados dos vetores obtidos até a 10ª iteração.

Método SOR com $\omega = 0,9$				
k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$
1	1,575	1,0293	0,8340	0,9986
2	1,0885	1,0043	1,0006	1,0208
3	1,0030	0,9988	1,0039	1,0033
4	0,9989	0,9987	1,0006	1,0001
5	0,9998	0,9999	1,0000	0,9999
6	1,0000	0,9999	0,9999	0,9999
7	1,0000	1,0000	1,0000	1,0000
8	1,0000	1,0000	1,0000	1,0000
9	1,0000	1,0000	1,0000	1,0000
10	1,0000	1,0000	1,0000	1,0000

Observe que o Método SOR ou Relaxação Sucessiva mostrou-se bastante eficiente na resolução do sistema linear pois com apenas 4 iterações chegamos a precisão desejada e na 7^a iteração chegamos a solução exata do sistema. Utilizamos o método da sub-relação no qual o parâmetro ω assume um valor tal que $0 < \omega < 1$.

Método de Gauss-Jacobi					Método de Gauss-Seidel			
k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$
1	1,75	1,1875	1,025	1,2406	1,75	1,1875	1,025	1,2406
2	1,45	1,1812	0,9	0,925	0,8867	0,9447	0,9994	0,9577
3	0,9984	1,1093	1,1474	1,1328	1,0245	1,0113	1,0009	1,0091
4	0,9026	1,0014	1,0699	1,0637	0,9946	0,9976	0,9997	0,9979
5	0,9984	1,1093	1,1474	1,1328	1,0012	1,0005	1,0000	1,0004
6	1,0091	0,9916	0,9824	0,9841	0,9997	0,9998	0,9999	0,9998
7	1,0104	1,0020	0,9952	0,9957	1,0001	1,0000	0,9999	1,0000
8	1,0017	1,0025	1,0020	1,0019	1,0000	0,9999	0,9999	0,9999
9	0,9984	1,0004	1,0017	1,0015	1,0000	1,0000	0,9999	0,9999
10	0,9991	0,9996	1,0001	1,0001	1,0000	1,0000	1,0000	1,0000

Método SOR com $\omega = 0,9$				
k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$
1	1,575	1,0293	0,8340	0,9986
2	1,0885	1,0043	1,0006	1,0208
3	1,0030	0,9988	1,0039	1,0033
4	0,9989	0,9987	1,0006	1,0001
5	0,9998	0,9999	1,0000	0,9999
6	1,0000	0,9999	0,9999	0,9999
7	1,0000	1,0000	1,0000	1,000
8	1,0000	1,0000	1,0000	1,000
9	1,0000	1,0000	1,0000	1,000
10	1,0000	1,0000	1,0000	1,000

Observação: Os métodos iterativos também mostraram-se eficientes, uma vez que a matriz dos coeficientes do sistema é estritamente diagonalmente dominante, isto é, em cada linha o elemento da diagonal principal é maior em módulo que a soma dos outros elementos, portanto obedece o critério das linhas o que garante a convergência do sistema. Até a 10ª iteração não conseguimos chegar na solução exata do sistema através do Método de Gauss-Jacobi, no entanto quando utilizamos os Métodos de Gauss-Seidel e SOR por sub-relaxação com $\omega = 0,9$ chegamos na solução exata na 10ª e 7ª iterações respectivamente.

Neste exemplo assim como na maioria dos sistemas lineares principalmente os de grande porte o Método SOR é preferível ao Método de Gauss-Seidel por acelerar a convergência, mais isso depende muito do valor atribuído ao parâmetro ω . O Método SOR supera amplamente o Método de Gauss-Jacobi. Na maioria das vezes utilizando os Métodos de Gauss-Seidel e SOR chega-se na solução do sistema calculando um número de iterações duas vezes menor do que a calculada utilizando o método de Gauss-Jacobi.

Considerações Finais

Neste trabalho de conclusão de curso, foram apresentados as matrizes que é sem dúvida uma das estruturas algébricas mais utilizadas, pois suas mais variadas formulações matemáticas permitem a simplificação não só da parte teórica, mas também das próprias aplicações. As resoluções de sistemas lineares estão nesse caso, ou seja, na definição e resolução desses sistemas, a forma matricial está sempre presente, seja na esquematização dos Métodos Diretos, seja no estudo da convergência dos Métodos Iterativos.

Como parte de maior importância, neste trabalho foram apresentados os Métodos Diretos e Iterativos para a resolução de sistemas lineares. Estes métodos são amplamente utilizados desde problemas mais simples do nosso cotidiano (por exemplo, para se ter uma noção de quanto se gasta ou se ganha em diversas atividades comerciais, no planejamento e gerenciamento de empresas) até os problemas de maior complexidade como é o caso dos grandes números de sistemas que resultam da discretização das equações diferenciais que apresentam como características muitos elementos nulos nas matrizes dos coeficientes (esparsidade).

Verificamos que a escolha do método a ser utilizado depende das peculiaridades do problema.

Vimos que os Métodos Diretos são processos finitos e, portanto, teoricamente, obtêm a solução exata de qualquer sistema não singular de equações a menos de erros de arredondamentos. Já os Métodos Iterativos tem convergência assegurada apenas sob determinadas condições.

Os Métodos Iterativos se mostram muito eficiente em sistemas de grande porte onde as matrizes dos coeficientes apresenta esparsidade, pois esses métodos tem como principal vantagem não alterar a estrutura da matriz dos coeficientes. Já os Métodos Diretos quando aplicados a esses sistemas provocam o preenchimento da matriz dos coeficientes, ou seja, durante o processo de eliminação poderão surgir elementos não nulos em posições que originalmente eram nulos exigindo então técnicas especiais para a escolha do pivô para reduzir este preenchimento, contudo existem situações nas quais pode ser impossível aplicar o Método Direto.

Quanto aos erros de arredondamento vimos que os Métodos Diretos apresentam sérios problemas de arredondamento. Adotando as técnicas de pivoteamento encontramos uma forma de amenizar esses problemas.

Os Métodos Iterativos apresentam menos erros de arredondamento, sendo que a convergência, uma vez assegurada, independe da aproximação inicial para a solução do sistema. Sendo assim somente os erros cometidos na última iteração afetam a solução, pois os erros cometidos nas iterações anteriores não levarão à divergência do processo nem à convergência a um outro vetor que não a solução do sistema.

Referências Bibliográficas

- [1] ARENALES, Selma, DAREZZO, Artur. **Cálculo Numérico**: Aprendizagem com apoio de software, Editora Thomson, São Paulo/SP 2008;
- [2] BARROSO, Leônidas Conceição. **Cálculo Numérico com Aplicações**: Editora Harbra, 2ª edição, S. Paulo/SP 1987;
- [3] BURDAM, Richard L., FAIRES, J. Douglas. **Análise Numérica**. Editora Thomson, 7ª Edição, 2001;
- [4] DA SILVA, Rogério Brasil **Métodos Diretos e Iterativos na Solução de Sistemas de Equações Lineares**: Disponível em <http://www.unifap.br/?pages=biblioteca>, 2011;
- [5] IEZZI, Gelson; HAZZAN, Samuel. **Fundamentos de Matemática Elementar**. vol. 4. Editora Atual, 7ª Edição, São Paulo/SP. 2010;
- [6] FREITAS, Sergio Roberto. **Métodos Numéricos**. Departamento de Computação e Estatística do Centro de Ciência Exatas e Tecnologia, UFMS. 2000;
- [7] RUGGIERO, M. A. G. e Lopes, V. L. R. **Cálculo Numérico**. Aspectos Teóricos e Computacionais. Editora McGraw-Hill, São Paulo/SP. 1988;