

UNIVERSIDADE FEDERAL DA PARAÍBA  
CENTRO DE CIÊNCIAS E TECNOLOGIA  
DEPARTAMENTO DE SISTEMAS E COMPUTAÇÃO

ESTUDO E IMPLEMENTAÇÃO DE MÉTODOS ITERATIVOS  
PARA SOLUÇÃO DE SISTEMAS DE EQUAÇÕES LINEAR  
RES INTEGRADOS A UMA BIBLIOTECA NUMÉRICA

MARIA DAS NEVES ALMEIDA SOARES

CAMPINA GRANDE - PARAÍBA

Outubro/1983



S676e Soares, Maria das Neves Almeida  
Estudo e implementacao de metodos iterativos para solucao de sistemas de equacoes lineares integrados a uma biblioteca numerica / Maria das Neves Almeida Soares. - Campina Grande, 1983.  
51 f. : il.

Dissertacao (Mestrado em Ciencias) - Universidade Federal da Paraiba, Centro de Ciencias e Tecnologia.

1. Biblioteca de Sub-Rotinas (Informatica) 2. Biblioteca Numerica 3. Metodos Iterativos 4. Dissertacao I. Queiroz, Bruno Correia da Nobrega, M.Sc. II. Universidade Federal da Paraiba - Campina Grande (PB) III. Título

CDU 004.428(043)

ESTUDO E IMPLEMENTAÇÃO DE MÉTODOS ITERATIVOS PARA  
SOLUÇÃO DE SISTEMAS DE EQUAÇÕES LINEARES INTEGRADOS  
A UMA BIBLIOTECA NUMÉRICA

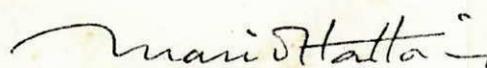
MARIA DAS NEVES ALMEIDA SOARES

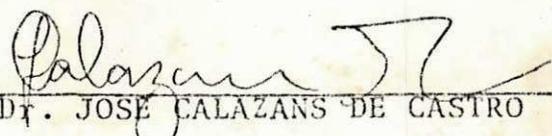
Tese submetida ao Corpo Docente do Programa de Pós-Graduação em Sistemas e Computação, Área de Concentração: Ciência da Computação, do Centro de Ciências e Tecnologia da Universidade Federal da Paraíba, como parte dos requisitos necessários para a obtenção do Grau de Mestre em Ciências (M.Sc).

Aprovada por:

COMISSÃO EXAMINADORA

  
Prof. M.Sc. BRUNO CORREIA DA NOBREGA QUEIROZ  
- Presidente -

  
Prof. M.Sc. MARIO TOYOTARO HATTORI  
Examinador

  
Prof. Dr. JOSE CALAZANS DE CASTRO  
Examinador

À minha filha Danielle como um incentivo  
para que se disponha a grandes êxitos em  
seus estudos.

À minha mãe Célia minha gratidão por  
sua ajuda sempre indispensável em  
minhas realizações.

## AGRADECIMENTOS

Ao Professor Bruno Correia da Nóbrega Queiroz pela sua atenção, estímulo e valiosa orientação.

Ao Professor Mário T. Hattori pelas constantes revisões deste texto, e sugestões que foram decisivas e fundamentais na conclusão deste trabalho.

A todos os membros do DSC que de alguma forma cooperaram nesta realização.

## RESUMO

Este trabalho consta de um estudo dos métodos iterativos para a resolução de sistemas de equações lineares visando a implementação desses métodos em uma biblioteca de algoritmos numéricos em desenvolvimento no DSC/UFPA.

Inicialmente é apresentada uma visão geral de todos os métodos iterativos selecionados e em seguida um estudo detalhado do fator de relaxação no método de Sobre-Relaxação Sucessiva (SOR).

Finalmente são apresentadas as subrotinas e os testes seguindo as técnicas usadas no projeto e implantação da biblioteca, bem como as conclusões relativas a comparações entre os métodos diretos e iterativos e entre os métodos iterativos entre si.

## ABSTRACT

The present work contains an analysis of the iterative methods for solving linear systems of equations in view of their implantation in a library of numerical algorithms in the Department of Computer Science - UFPb.

Thus a preliminary survey of all selected iterative methods is presented, followed by an exhaustive research of the relaxation factor in the Successive Over-Relaxation Method (SOR).

Finally subroutines and tests are presented according to the techniques used in the design and implantation of the library, as well as conclusions drawn from a comparative study between direct and iterative methods and between iterative methods themselves.

## ÍNDICE

CAPÍTULO I	- INTRODUÇÃO	01
CAPÍTULO II	- MÉTODOS ITERATIVOS	05
2.1	- Introdução	05
2.2	- Método Iterativo Linear Estacionário de Grau 1	05
2.3	- Convergência	06
2.4	- Consistência	07
2.5	- Taxa de Convergência	08
2.6	- Revisão de Alguns Métodos Iterativos	08
2.6.1	- Método de Jacobi	08
2.6.2	- Método de JOR	09
2.6.3	- Método de Gauss-Seidel	09
2.6.4	- Método SOR	11
2.7	- A Influência das Propriedades da Matriz A na Convergência dos Métodos Iterativos	15
2.7.1	- Definições Básicas	15
2.7.1.1	- Matriz de Diagonal Dominante	15
2.7.1.2	- Matriz de Diagonal Estritamente Dominante	15
2.7.1.3	- Matriz Irredutível	15
2.7.1.4	- Matriz Positiva Definida	16
2.7.1.5	- Matriz Consistentemente Ordenada	16
2.7.2	- Critérios de Convergência	16
CAPÍTULO III	- OTIMIZAÇÃO DO FATOR DE RELAXAÇÃO NO MÉTODO SOR	19
3.1	- Introdução	19
3.2	- Determinação do Fator Ótimo de Relaxação	20
3.3	- Comportamento de $S(Y_w)$	28

CAPÍTULO IV	- IMPLEMENTAÇÃO, TESTES COMPARATIVOS E CONCLUSÕES	34
4.1	- Implementação	34
4.2	- Testes	40
4.3	- Conclusões	48
REFERÊNCIAS BIBLIOGRÁFICAS		50

## CAPÍTULO I

### INTRODUÇÃO

Neste trabalho nos propomos a implementar alguns métodos para inserir numa biblioteca educacional de algoritmos numéricos e apresentar uma revisão dos métodos iterativos para a solução de sistemas de equações lineares. Objetivamos despertar o interesse pelo assunto, evidenciar sua importância e em termos práticos esclarecer estudantes ou profissionais de qualquer área de interesse que necessitem utilizar esta matéria como um meio para a resolução de problemas formulados matematicamente dentro de seu ramo de atividade.

Resolver sistemas de equações lineares cuja matriz dos coeficientes é esparsa de ordem elevada e estruturada é muitas vezes um dos passos da solução de um problema maior. Exemplos típicos são os problemas que envolvem a resolução de sistemas de equações diferenciais parciais dos tipos parabólico, elíptico e hiperbólico.

Devido às dificuldades e na maioria dos casos à impossibilidade de se obter uma solução analítica, tais problemas são resolvidos por métodos numéricos o que leva à resolução de um sistema de equações lineares, em geral dotado das características acima citadas.

Como exemplos de áreas que conduzem a problemas desta natureza podem ser citadas as de difusão de neutrons, escoamento de fluidos, elasticidade, transmissão de calor em sólidos, previsão do tempo, além de problemas de engenharia [14].

A resolução de sistemas de grande porte e esparsos é praticamente impossível sem o uso de computadores. Usualmente são utilizados métodos iterativos que não destroem a estrutura da matriz dos coeficientes, ao contrário dos métodos diretos que introduzem elementos não nulos nas posições que inicialmente continham elementos nulos. Métodos iterativos se adaptam melhor a tais sistemas quando se visa a minimização do tempo de processamento e do espaço de armazenamento requeridos no computador. Contudo deve-se chamar a atenção para o fato de que não há regras rígidas e que existem até mesmo controvérsias em escolher um método direto ou iterativo [9]. Uma justificativa para a escolha de métodos iterativos na solução de sistemas de equações lineares cuja matriz dos coeficientes é esparsa de ordem elevada e estruturada é encontrada em [9]. Segundo o autor, a solução deste tipo de sistemas frequentemente causa problemas de espaço de armazenamento. Entretanto a aplicação dos métodos iterativos pode contornar esta dificuldade, porque estes métodos são econômicos na utilização do espaço de memória, quando comparados com os métodos diretos. Isto se deve à simplicidade e uniformidade das operações a serem realizadas.

A característica que mais pesa na escolha do tipo de método é a esparsidade já que em se usando métodos diretos são realizados cálculos desnecessários ao passo que em se aplicando métodos iterativos o número de cálculos é reduzido.

O que foi exposto acima evidencia que os métodos iterativos para a solução de sistemas de equações lineares têm aplicações variadas. Estas aplicações, constituem uma forte motivação para se estudar este assunto.

A estrutura deste trabalho estabelece critérios de escolha, na prática, de um método que possa ser aplicado com sucesso na solução de um sistema evitando dificuldades de não convergência para a solução desejada. A aplicação do método de Sobre-Relaxação Sucessiva (SOR) dentro das condições mais eficientes possíveis é outro ponto a ser focalizado no trabalho. Além disso foram desenvolvidas sub-rotinas para verificar na prática as conclusões teóricas. Para alcançar esta metas o texto inicialmente revisa os fundamentos da teoria dos métodos iterativos. Será feito o estudo de um método iterativo linear estacionário de grau 1 que se caracteriza por fornecer a cada iteração uma aproximação do vetor solução satisfazendo à equação

$$u^{(m+1)} = Gu^{(m)} + k. \quad (1.1)$$

Serão destacados os conceitos de consistência, convergência, taxa de convergência, e estabelecidas condições sob as quais o método converge para a solução única do sistema.

Na análise do método genérico definido pela equação (1.1) não há preocupação quanto à forma da matriz  $G$  e do vetor  $k$ . Escolhendo-se formas particulares desses parâmetros obtém-se os métodos

dos de Jacobi, SOR, Gauss-Seidel e Sobre-Relaxação Simultânea (JOR). Cada um desses métodos será revisto e todas as conclusões, propriedades e resultados para o caso geral serão estendidos a cada um deles fazendo-se comparações. As condições de convergência dos quatro métodos que variam de acordo com as propriedades da matriz dos coeficientes são consideradas em seguida. A atenção será voltada para o caso das matrizes dos coeficientes ser real. Além disso para limitar o escopo deste trabalho não serão vistos todos os casos restringindo-se aos sistemas onde a matriz tem pelo menos uma dentre as propriedades abaixo, o que garante a convergência de algum método: matriz com diagonal estritamente dominante, matriz irredutível com diagonal dominante, matriz positiva definida, matriz positiva definida e consistentemente ordenada.

Desde que para o método SOR  $G = Y_w$ , é uma função de  $W$ , é de se esperar que este fator tenha influência na velocidade de convergência. O capítulo III revisa o SOR estabelecendo uma equação para o cálculo do fator  $W$ ótimo, no caso particular das matrizes dos coeficientes positivas definidas e consistentemente ordenadas. Esta classe de matrizes é importante pois surge quando se aplica o método das diferenças finitas para resolver certas equações diferenciais parciais elípticas [14]. Uma análise da variação do raio espectral de  $Y_w$ ,  $S(Y_w)$  como função de  $W$  é feita neste capítulo.

Um quarto capítulo apresenta resultados numéricos e um anexo apresenta a listagem de sub-programas desenvolvidos em FORTRAN.

## CAPÍTULO II

### MÉTODOS ITERATIVOS

#### 2.1. Introdução

Este capítulo estuda um método iterativo estacionário genérico de grau 1 e define os métodos de Jacobi, JOR, Gauss-Seidel e SOR (particularizando os parâmetros da equação do método geral). São estabelecidos conceitos, propriedades e teoremas relativos à consistência e convergência dos métodos.

#### 2.2. Método Iterativo Linear Estacionário de Grau 1

Seja um sistema de equações lineares

$$Au = b, \quad (2.1)$$

de solução única, cuja matriz dos coeficientes  $A$  é de ordem  $n$ . Um método iterativo para resolver este sistema é aquele que selecionando uma aproximação inicial da verdadeira solução determina uma sequência  $u^{(1)}, u^{(2)}, \dots, u^{(m)}$  que dentro de certas condi

ções convergirã para a soluçãõ.

Um método iterativo é linear estacionário de grau 1, se cada novo termo da sequência de números é obtido de uma função linear de  $A$  e  $b$  e do termo anterior. Assim,

$$u^{(m+1)} = \phi(u^{(m)}, A, b). \quad (2.2)$$

Trataremos do método iterativo genérico definido por

$$u^{(m+1)} = Gu^{(m)} + k \quad (2.3)$$

onde  $G$  é uma matriz que depende de  $A$  e  $k$  é um vetor que depende de  $A$  e de  $b$ .

### 2.3. Convergência

O método iterativo definido por (2.3) é convergente se e somente se para qualquer solução inicial  $u^{(0)}$ , a sequência gerada converge para algum  $\bar{u}$ .

O teorema a seguir é importante no estudo da convergência pois estabelece uma condição que os autovalores da matriz  $G$  devem satisfazer para haver convergência.

#### Definição 2.1

O raio espectral  $S(A)$  de uma matriz  $A$  é o maior dentre os módulos dos autovalores de  $A$ ,

$$S(A) = \max_{1 \leq i \leq n} |\lambda_i(A)|$$

Teorema (2.1)

O método iterativo definido pela equação (2.3) é convergente se e somente se o raio espectral da matriz  $G$ ,  $S(G)$ , satisfaz à condição  $S(G) < 1$ .

O teorema (2.1) estabelece uma condição necessária e suficiente para a convergência do método. Entretanto ele não garante a convergência para a solução única do sistema. Para se determinar quando isto ocorre faz-se necessária a introdução do conceito de consistência.

2.4. Consistência

O método iterativo definido por (2.3) é consistente com (2.1) se e somente se toda solução de (2.1) for solução da equação

$$(I-G)u = k \quad (2.4)$$

O teorema seguinte estabelece uma condição sob a qual o método converge para a única solução do sistema (2.1) qualquer que seja a aproximação inicial.

Teorema (2.2)

Se o sistema (2.1) tem solução e o método iterativo é consistente com (2.1) e convergente, então  $A$  é não singular e a convergência se dá para a única solução do sistema qualquer que seja a aproximação inicial [14].

Se um método iterativo é consistente com o sistema (2.1) e se em alguma iteração a solução  $\bar{u}$  deste sistema for obtida, em

todas as iterações subsequentes o mesmo vetor  $\bar{u}$  será repetidamente encontrado. Este fato evidencia também a importância do conceito de consistência.

## 2.5. Taxa de Convergência

Até esta altura tratou-se de condições sob as quais o método iterativo definido por (2.3) converge para a solução única do sistema  $Au = b$ , sem preocupação com a velocidade de convergência.

A taxa de convergência é um conceito que surge da necessidade de se analisar essa velocidade e que se define por

$$R(G) = -\log S(G). \quad (2.5)$$

Da definição da taxa de convergência conclui-se que a velocidade de convergência será tanto maior quanto menor for  $S(G)$ .

## 2.6. Revisão de Alguns Métodos Iterativos

Na seção anterior tratamos de um método iterativo genérico definido pela equação (2.3) sem especificar a forma da matriz  $G$  e do vetor  $k$ . A escolha destes parâmetros define métodos particulares tais como o de Jacobi, Gauss-Seidel, JOR e SOR.

### 2.6.1. O Método de Jacobi

Seja o sistema (2.1), onde  $A = [a_{ij}]$  e o vetor  $b = [b_i]$ .

Se na equação (2.3)  $G$  é definida por  $G=B$ , onde

$$B = [b_{ij}] = \begin{cases} -a_{ij}/a_{ii}, & i \neq j \\ 0, & i = j \end{cases}$$

e  $k = K = [b_i/a_{ii}]$ , obtém-se o método de Jacobi.

Uma vez que é interessante fazer sempre referência ao método como um caso particular do método geral faz-se necessário estabelecer uma correspondência entre alguns elementos da equação geral com a equação do método particular em questão. Desta forma será útil definir:

$$D = \text{diag } A$$

$$C = D - A$$

Conclui-se que:  $B = D^{-1}C$  e  $K = D^{-1}b$ . Além disso, tem-se:

$$(I-B) = I - D^{-1}C = D^{-1}(D-C) = D^{-1}A$$

### 2.6.2. Método JOR

O método iterativo JOR é uma variante do método de Jacobi onde se escolhe um parâmetro real  $W$  e a sequência é formada pela equação

$$u^{(m+1)} = B_w u^{(m)} + WK \quad (2.6)$$

onde  $B_w = WB + (1-W)I$ . Se  $W=1$ , o método é de Jacobi.

Como no caso anterior faz-se necessário estabelecer uma correspondência entre elementos das equações (2.3) e (2.6)

$$(I-B_w) = I - (WB + (1-W)I) = W(I-B) = WD^{-1}A.$$

### 2.6.3. Método de Gauss-Seidel

Como mais um caso particular do método definido por (2.3), tem-se o método de Gauss-Seidel, definido pela equação

$$u_i^{(m+1)} = \sum_{j=1}^{i-1} b_{ij} u_j^{(m+1)} + \sum_{j=i+1}^n b_{ij} u_j^{(m)} + k_i \quad (2.7)$$

onde  $b_{ij}$  é um elemento de  $B$ , e  $k_i$  é um elemento de  $K$  ( $B$  e  $K$  definidos na seção 2.6.1), e  $u_i^{(m+1)}$  e  $u_i^{(m)}$ , indicam os  $i$ -ésimos componentes dos vetores  $u^{(m+1)}$  e  $u^{(m)}$  respectivamente.

Ao contrário dos métodos de Jacobi e JCR como se pode observar, a equação do método foi apresentada acima em função de um componente genérico da aproximação  $u^{(m+1)}$ . Para homogeneidade e padronização das considerações, faz-se necessário que se apresente como anteriormente a equação do método na forma matricial.

Sejam

$$U = \begin{cases} 0 & , i \geq j \\ b_{ij} & , i < j \end{cases} \quad \text{e} \quad L = \begin{cases} 0 & , i \leq j \\ b_{ij} & , i > j \end{cases}$$

De acordo com as definições acima, vemos que  $U$  e  $L$  são matrizes estritamente triangulares e tais que  $B = U+L$ . Assim, a sequência será formada pela equação

$$u^{(m+1)} = Lu^{(m+1)} + Uu^{(m)} + K. \quad (2.8)$$

Desde que  $(I-L)$  é triangular inferior e seus elementos da diagonal são iguais a 1, então  $\det(I-L) = 1$  e portanto é inversível. Assim é que a equação do método é dada por

$$u^{(m+1)} = Yu^{(m)} + (I-L)^{-1}K \quad (2.9)$$

onde  $Y = (I-L)^{-1}U$ . Veja-se agora a correspondência entre os elementos das equações (2.3) e (2.9).

$$G = Y$$

$$(I-G) = (I-Y) = (I-L)^{-1} D^{-1}A$$

## 2.6.4. Método SOR

Este método é uma variante do Gauss-Seidel onde se escolhe um parâmetro real  $W$  e a sequência é formada por:

$$u_i^{(m+1)} = W \left[ \sum_{j=1}^{i-1} b_{ij} u_j^{(m+1)} + \sum_{j=i+1}^n b_{ij} u_j^{(m)} + k_i \right] + (1-W) u_i^{(m)} \quad (2.10)$$

Quando  $W=1$ , tem-se o método de Gauss-Seidel.

Como no caso do método de Gauss-Seidel é necessário que se apresente a equação do método na forma matricial para padronização das considerações.

De acordo com a equação (2.10) pode-se afirmar que:

$$u^{(m+1)} = W [Lu^{(m+1)} + Uu^{(m)} + K] + (1-W)u^{(m)}$$

Daí,  $u^{(m+1)} = WLu^{(m+1)} + WUu^{(m)} + (1+W)u^{(m)} + WK$  e tem-se  $(I-WL)u^{(m+1)} = (WU + (1-W)I)u^{(m)} + WK$ . Mas  $\det(I-WL) \neq 0$  e portanto  $(I-WL)$  é inversível. Então

$$u^{(m+1)} = (I-WL)^{-1} (WU + (1-W)I) u^{(m)} + (I-WL)^{-1} WK.$$

Fazendo-se

$$Y_w = (I-WL)^{-1} (WU + (1-W)I), \text{ tem-se:}$$

$$u^{(m+1)} = Y_w u^{(m)} = (I-WL)^{-1} WK \quad (2.11)$$

como equação do método SOR.

Veja-se a correspondência entre os elementos das equações (2.3) e (2.11)

$$G = Y_w$$

$$(I-G) = (I-Y_w) = W(I-WL)^{-1} D^{-1} A$$

Definidos estes quatro métodos básicos pode-se concluir pela própria construção das matrizes que uma condição necessária para a aplicação de qualquer um dos métodos é que todos os ele

mentos da diagonal da matriz  $A$  dos coeficientes sejam diferentes de zero.

No que diz respeito à convergência, o teorema (2.1) permite se concluir que cada um dos métodos em questão é convergente se e somente se o raio espectral da sua respectiva matriz correspondente à matriz  $G$  for menor que 1. Em outras palavras tem-se baseado no teorema (2.1):

- a) O método de Jacobi converge se e somente se  $S(B) < 1$
- b) O método de Gauss-Seidel converge se e somente se  $S(Y) < 1$
- c) O método JOR converge se e somente se  $S(B_w) < 1$
- d) O método SOR converge se e somente se  $S(Y_w) < 1$ .

Por outro lado, conforme se observa para os quatro métodos em questão é possível expressar  $k$  por  $Mb$  e  $(I-G)$  por  $MA$ , onde  $M$  é uma matriz. Assim considerando-se o sistema (2.1) tem-se que  $(I-G)A^{-1}b = MAA^{-1}b = Mb = k$  e, portanto,  $A^{-1}b$  é também solução de (2.4). Isto mostra que o método definido pela equação (2.3) é consistente com o sistema (2.1) de acordo com a definição de consistência.

Com base no exposto acima, pode-se afirmar que os métodos de Jacobi, Gauss-Seidel, JOR e SOR são consistentes com o sistema (2.1) de solução única.

Nas equações (2.6) e (2.11) que definem os métodos JOR e SOR aparece um parâmetro  $w$  mais do que nos métodos de Jacobi e Gauss-Seidel. É de se esperar que este parâmetro que é o fator de relaxação  $w$ , tenha influência na convergência do método. Esta

influência é estudada a seguir.

Teorema (2.3)

O método JOR não converge se  $W = 0$ .

Teorema (2.4)

O método SOR não converge se  $W$  não pertence ao intervalo  $(0,2)$ .

Segue um quadro comparativo dos parâmetros envolvidos na equação (2.3) e nas equações particulares referentes aos métodos de Jacobi, JOR, Gauss-Seidel e SOR.

GERAL	JACOBI	GAUSS-SEIDEL	JOR	SOR
$u^{(m+1)} = Gu^{(m)} + k$	$u^{(m+1)} = Bu^{(m)} + K$	$u^{(m+1)} = Yu^{(m)} + (I-L)^{-1}K$	$u^{(m+1)} = B_w u^{(m)} + WK$	$u^{(m+1)} = Y_w u^{(m)} + (I-WL)^{-1}K$
$G$	$B = D^{-1}C$	$Y = (I-L)^{-1}U$	$B_w = WD^{-1}C + (1-W)I$	$Y_w = (I-WL)^{-1}(WU + (1-W)I)$
$k$	$K = D^{-1}b$	$(I-L)^{-1}D^{-1}b$	$WK = WD^{-1}b$	$(I-WL)^{-1}WD^{-1}b$
$(I-G)$	$(I-B) = D^{-1}A$	$(I-L)^{-1}D^{-1}A$	$(I-B_w) = WD^{-1}A$	$W(I-WL)^{-1}D^{-1}A$

QUADRO 1

Comparativo dos parâmetros dos métodos: Geral,  
Jacobi, Gauss-Seidel, JOR e SOR.

## 2.7. A Influência das Propriedades da Matriz A na Convergência dos Métodos Iterativos

### 2.7.1. Definições Básicas

#### 2.7.1.1. Matriz de diagonal dominante

#### 2.7.1.2. Matriz de diagonal estritamente dominante

Uma matriz  $A$  de ordem  $n$ , tem diagonal dominante, se e somente se

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

e

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

para um valor de  $i$ , no mínimo. Se por outro lado tem-se

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

a matriz é de diagonal estritamente dominante.

#### 2.7.1.3. Matriz Irredutível

Uma matriz  $A$  de ordem  $n$  é irredutível se e somente se  $n=1$  ou se quando  $n>1$  e  $a_{ij}=0$  com  $i \neq j$ ,  $1 \leq i \leq n$ ,  $1 \leq j \leq n$  existe uma sequência  $i_1, i_2, \dots, i_s$  tal que  $a_{i_1 i_1}, a_{i_1 i_2}, a_{i_2 i_3}, \dots, a_{i_s i_s}$  são todos diferentes de zero.

#### 2.7.1.4. Matriz Positiva Definida

Uma matriz  $A$ , real é positiva definida se e somente se é simétrica e todos os seus autovalores são positivos.

#### 2.7.1.5. Matriz Consistentemente Ordenada

Uma matriz de ordem  $n$  é consistentemente ordenada, se e somente se, existem  $t$  subconjuntos disjuntos de  $W = \{1, 2, \dots, n\}$  tais que sua união é igual a  $W$  e:  $j \in S_{k+1}$  quando  $i \in S_k$  e  $j > i$ , enquanto que  $j \in S_{k-1}$  se  $j < i$ , desde que  $i \neq j$  e  $a_{ij} \neq 0$  ou  $a_{ji} \neq 0$ .

Os teoremas a seguir estabelecem critérios de convergência para os quatro métodos, com base nas propriedades da matriz  $A$  dos coeficientes do sistema (2.1) determinado.

### 2.7.2. Critérios de Convergência

#### Teorema (2.5)

Se o método de Jacobi converge, então o método JOR converge quando  $0 < W < 1$ .

#### Teorema (2.6)

Se  $A$  é uma matriz de ordem  $n$  com diagonal estritamente dominante, então os métodos de Jacobi, Gauss-Seidel e o método JOR convergem se  $0 < W < 1$ .

#### Teorema (2.7)

Se  $A$  é uma matriz de ordem  $n$  irredutível com diagonal dominante então convergem os métodos de Jacobi, Gauss-Seidel e os métodos JOR e SOR se  $0 < W < 1$ .

Teorema (2.8)

Se  $A$  é uma matriz real positiva definida convergem os métodos de Gauss-Seidel e o SOR se  $0 < \omega < 2$ .

Teorema (2.9)

Se  $A$  é uma matriz real positiva definida e consistentemente ordenada, convergem os métodos de Jacobi, Gauss-Seidel, JOR se  $0 < \omega < 1$  e SOR se  $0 < \omega < 2$ .

As demonstrações dos teoremas acima podem ser encontradas em [14] ou [15].

O quadro 2 apresenta um resumo para facilitar a aplicação dos métodos com base nas propriedades da matriz dos coeficientes do sistema.

	MATRIZ COM DIAGONAL ESTRITAMENTE DOMINANTE.	MATRIZ IRREDUTÍVEL E COM DIAGONAL DOMINANTE.	MATRIZ REAL POSITIVA DEFINIDA.	MATRIZ REAL POSITIVA DEF. E CONSISTENTEMENTE ORDENADA.
JACOBI	Converge	Converge	Não Analisado	Converge
JOR	Converge se $0 < W \leq 1$	Converge se $0 < W \leq 1$	Não Analisado	Converge se $0 < W \leq 1$
GAUSS-SEIDEL	Converge	Converge	Converge	Converge
SOR	Não Analisado	Converge se $0 < W \leq 1$	Converge se $0 < W < 2$	Converge se $0 < W \leq 2$

QUADRO 2

A convergência dos métodos de Jacobi, Gauss-Seidel, JOR e SOR, de acordo com as propriedades da matriz dos coeficientes do sistema.

## CAPÍTULO III

### OTIMIZAÇÃO DO FATOR DE RELAXAÇÃO NO MÉTODO SOR

#### 3.1. Introdução

Sabemos que a convergência do método iterativo definido pela equação (2.1) depende do valor do raio espectral de  $G$  como ficou estabelecido no Capítulo II. Além disto, vimos que esta convergência será tanto mais rápida quanto menor for  $S(G)$ . Por estas razões podemos concluir que sendo a matriz  $G = Y_w$  e logo  $S(G) = S(Y_w)$  no caso do método SOR uma função de  $w$ , a escolha do fator ótimo de relaxação,  $w_{\text{ótimo}}$ , é feita procurando-se  $w$  que minimize  $S(Y_w)$ , maximizando a taxa de convergência  $R(Y_w)$ .

Neste trabalho, a determinação de  $w_{\text{ótimo}}$  será restrita ao caso das matrizes dos coeficientes positivas definidas e consistentemente ordenadas.

O teorema (3.1) é devido a David M. Young [14] e estabelece uma equação para o cálculo do fator ótimo no caso particular que nos propusemos a analisar. Apresentamos uma demonstração alternativa.

### 3.2. Determinação do fator de relaxação ótimo.

Teorema (3.1) Se  $A$  é positiva definida e consistentemente ordenada, o valor ótimo de  $W$  para a aplicação do SOR ao sistema (2.1) é dado pela equação

$$W_{\text{ótimo}} = \frac{2}{1 + (1 - (S(B))^2)^{1/2}}, \quad (3.1)$$

onde  $B$  é a matriz do método de Jacobi.

Prova

1ª parte

Se  $A$  é positiva definida então é hermitiana por definição. Desde que  $A$  é real e hermitiana, então  $A$  é simétrica e daí  $B = I - D^{-1}A$  também é simétrica. Sendo  $B$  real e simétrica é hermitiana e portanto tem todos os seus autovalores reais.

Por outro lado, todos os elementos da diagonal de  $A$  são positivos e portanto não nulos. Desde que  $A$  é consistentemente ordenada então [14]

$$(S(B))^2 = S(Y).$$

Mas  $S(Y) < 1$  pois o método de Gauss-Seidel converge. Dai  $S(B) < 1$ .

2ª parte

Se  $A$  é consistentemente ordenada com elementos não nulos na diagonal, todos os autovalores de  $Y_w$  podem ser obtidos resolvendo-se uma das equações.

$$(\lambda + W - 1)^2 = W^2 \mu^2 \lambda, \quad (3.2)$$

$$\lambda + W - 1 = W \mu \lambda^{1/2} \quad (3.3)$$

para cada autovalor  $\mu$  de  $B$  [14].

Escolhendo-se um valor de  $W$  no intervalo  $(0,2)$  e fixando-se um autovalor  $\mu$  de  $B$ , cada equação fornecerá duas raízes que são os autovalores de  $Y_w$ .

Se  $r(W,\mu)$  é o maior dentre os módulos dessas raízes,  $S(Y_w) = \max \{r(W,\mu)\}$ , fazendo-se  $\mu$  percorrer o conjunto dos autovalores de  $B$ .

Fazendo-se  $\lambda = x^2$ , a equação (3.3) se reduz a

$$x^2 - W\mu x + (W-1) = 0 \quad (3.4)$$

Resolvendo a equação quadrática (3.4) encontra-se:

$$x = \frac{W\mu \pm (W^2\mu^2 - 4(W-1))^{1/2}}{2}$$

e

$$\lambda = \frac{(W\mu \pm (W^2\mu^2 - 4(W-1))^{1/2})^2}{4}$$

onde os dois valores de  $\lambda$  são os autovalores de  $Y_w$  obtidos para um certo  $W$  e para um valor fixo de  $\mu$ .

Sejam

$$\lambda_1(W,\mu) = \left( \frac{W\mu + (W^2\mu^2 - 4(W-1))^{1/2}}{2} \right)^2$$

e,

$$\lambda_2(W,\mu) = \left( \frac{W\mu - (W^2\mu^2 - 4(W-1))^{1/2}}{2} \right)^2$$

O objetivo agora, é mostrar que

$$r(W,\mu) = \left| \left( \frac{W|\mu| + (W^2\mu^2 - 4(W-1))^{1/2}}{2} \right)^2 \right| \quad (3.5)$$

Se  $\lambda_1(W, \mu)$  e  $\lambda_2(W, \mu)$  são raízes reais, neste caso  $(W^2\mu^2 - 4(W-1))^{1/2} > 0$ .

Se  $\mu > 0$ , desde que  $W > 0$  então  $\lambda_1(W, \mu)$  é a raiz de maior valor absoluto e, portanto,

$$\begin{aligned} r(W, \mu) = \lambda_1(W, \mu) &= \left( \frac{W\mu + (W^2\mu^2 - 4(W-1))^{1/2}}{2} \right)^2 = \\ &= \left( \frac{W|\mu| + (W^2\mu^2 - 4(W-1))^{1/2}}{2} \right)^2 \end{aligned}$$

Se  $\mu < 0$ ,  $\lambda_2(W, \mu)$  é a raiz de maior módulo pois temos que:

$$\begin{aligned} \lambda_2(W, \mu) &= \left( \frac{W\mu - (W^2\mu^2 - 4(W-1))^{1/2}}{2} \right)^2 = \\ &= \left[ - \left( \frac{W(-\mu) + (W^2\mu^2 - 4(W-1))^{1/2}}{2} \right)^2 \right] = \\ &= \left( \frac{W|\mu| + (W^2\mu^2 - 4(W-1))^{1/2}}{2} \right)^2 \end{aligned}$$

Então se  $\lambda_1(W, \mu)$  e  $\lambda_2(W, \mu)$  são raízes reais, o maior dentre os módulos destas raízes é

$$r(W, \mu) = \left| \left( \frac{W|\mu| + (W^2\mu^2 - 4(W-1))^{1/2}}{2} \right)^2 \right|$$

Observemos que se  $\mu = 0$ , temos  $\lambda_1(W, \mu) = \lambda_2(W, \mu) = (1-W)$  e, portanto,  $r(W, \mu) = |1-W|$ .

Por outro lado  $\left| \left( \frac{W|\mu| + (W^2\mu^2 - 4(W-1))^{1/2}}{2} \right)^2 \right| = \left| \left( \frac{(-4(W-1))^{1/2}}{2} \right)^2 \right| = |1-W|$  e portanto a relação (3.5) também é sa feita.

Sejam  $\lambda_1(W, \mu)$  e  $\lambda_2(W, \mu)$  pares de complexos conjugados.

Então  $\lambda_1 = a + bi$  e  $\lambda_2 = a - bi$

$|\lambda_1| = |a + bi| = |a - bi| = |\lambda_2|$ , e portanto

$$\begin{aligned} |\lambda_1| &= |\lambda_2| = \left| \left( \frac{W\mu + (W^2\mu^2 - 4(W-1))^{1/2}}{2} \right)^2 \right| = \\ &= \left| \frac{W\mu}{2} + \frac{(-W^2\mu^2 + 4(W-1))^{1/2}}{2} i \right| = \\ &= \left( \frac{W\mu}{2} \right)^2 + \frac{(-W^2\mu^2 + 4(W-1))}{2^2} = \\ &= \left( \frac{W|\mu|}{2} \right)^2 + \frac{(-W^2\mu^2 + 4(W-1))}{2^2} = \\ &= \left| \left( \frac{W|\mu| + (W^2\mu^2 - 4(W-1))^{1/2}}{2} \right)^2 \right| \end{aligned}$$

Como os módulos de ambas as raízes são iguais,  $r(W, \mu)$  satisfaz à relação (3.5).

Mostraremos agora, que  $r(W, \mu)$  cresce quando  $|\mu|$  cresce.

Se  $W \leq 1$  então  $-4(W-1) \geq 0$  e sendo todo o numerador da expressão positivo é claro que  $r(W, \mu) \leq r(W, \mu')$  quando  $|\mu| \leq |\mu'|$ .  
Se por outro lado  $W > 1$ , então  $-4(W-1) < 0$ .

1) Se  $\mu^2 > \frac{4(W-1)}{W^2}$ , então  $W^2\mu^2 - 4(W-1) > 0$ .

Quando  $|\mu|$  cresce, então  $W^2\mu^2 - 4(W-1)$  cresce e como no caso anterior  $r(W, \mu) \leq r(W, \mu')$ , quando  $|\mu| \leq |\mu'|$ .

2) Se  $\mu^2 \leq \frac{4(W-1)}{W^2}$ , então  $W^2\mu^2 - 4(W-1) \leq 0$

Seja  $m = W^2\mu^2 - 4(W-1) \leq 0$

$$r(W, \mu) = \left| \frac{W|\mu| + m^{1/2}}{2} \right|^2 = \left| \frac{W|\mu| + (-m)^{1/2}i}{2} \right|^2 =$$

$$= \frac{W^2|\mu|^2}{4} + \frac{(-m)}{4}$$

$$\text{Então } r(W, \mu) = \frac{W^2|\mu|^2}{4} + \frac{4(W-1) - W^2\mu^2}{4} =$$

$$= \frac{W^2|\mu|^2 + 4(W-1) - W^2|\mu|^2}{4} = (W-1)$$

assim  $r(W, \mu) = (W-1)$ . Se  $|\mu| \leq |\mu'|$ , então  $r(W, \mu) = (W-1) = r(W, \mu')$  e portanto,  $r(W, \mu) \leq r(W, \mu')$ .

Assim,  $r(W, \mu) \leq r(W, \mu')$  quando  $|\mu| \leq |\mu'|$  para qualquer  $W \in (0, 2)$ . Daí, pode-se concluir que para um certo valor fixo de  $W \in (0, 2)$ , desde que temos  $|\mu| \leq S(B)$ , então  $r(W, \mu) \leq r(W, S(B))$ , e portanto  $r(W, S(B)) = \max \{r(W, \mu)\}$  quando  $\mu$  percorre o conjunto dos autovalores de  $B$ .

$$\text{Assim, } r(W, S(B)) = S(Y_w) \quad (3.6)$$

Com esta conclusão, nos resta agora determinar o valor de  $W$  que torna mínimo  $S(Y_w)$ , ou seja, que torna o maior dentre os módulos das raízes da equação (3.3), fazendo-se  $\mu = S(B)$ , o menor possível.

Determinação do valor ótimo de  $W$ .

1) Seja  $S(B) \neq 0$

Tomemos a equação:

$$(\lambda + W - 1)^2 = W^2 (S(B))^2 \lambda \quad (3.7)$$

A equação (3.7) nos leva a

$$\lambda^2 + (2(W-1) - W^2(S(B))^2)\lambda + (W-1)^2 = 0 \quad (3.8)$$

Mostraremos que existe um único valor de  $W$ , para o qual a equação (3.8) tem duas raízes iguais. Neste caso, o discriminante desta equação de 2º grau é nulo, ou seja:

$$(2(W-1) - W^2(S(B))^2)^2 - 4(W-1)^2 = 0$$

Daí, tem-se que:

$$W^2 S(B)^2 (W^2 (S(B))^2 - 4(W-1)) = 0 \quad (3.9)$$

$$\text{e portanto } W^2 (S(B))^2 - 4W + 4 = 0. \quad (3.10)$$

Resolvendo a equação (3.10), em  $W$  teremos que:

$$W = \frac{4 \pm (16 - 4(4(S(B))^2)^{1/2}}{2 \cdot (S(B))^2}, \text{ e daí}$$

$$W = \frac{2 \pm 2(1 - (S(B))^2)^{1/2}}{(S(B))^2}.$$

Desde que  $(S(B))^2 < 1$  e  $W < 2$ , a única solução será

$$W = \frac{2 - 2(1 - (S(B))^2)^{1/2}}{(S(B))^2} = \frac{2}{1 + (1 - (S(B))^2)^{1/2}}.$$

Chamaremos este valor de  $W_b$  e teremos:

$$W_b = \frac{2}{1 + (1 - (S(B))^2)^{1/2}} \quad (3.11)$$

É fácil de verificar que  $W_b > 1$ .

Tomaremos agora a equação

$f(W) = W^2 (S(B))^2 (W^2 (S(B))^2 - 4(W-1))$ , onde o segundo mem  
bro é o discriminante (3.9).

$$\text{Seja } g(W) = (W^2(S(B))^2 - 4(W-1))$$

$$g'(W) = 2W(W(B))^2 - 4 = 2(W(S(B))^2 - 2). \text{ Assim,}$$

$$g'(W) = 2(W(S(B))^2 - 2) \quad (3.12)$$

Se  $0 < W < 2$  e  $0 < (S(B))^2 < 1$ , então  $0 < W(S(B))^2 < 2$  e portanto  $W(S(B))^2 - 2 < 0$ , de onde se conclui que  $g'(W) < 0$  e portanto a função  $g(W)$  é decrescente para  $0 < W < 2$ .

Quando  $W = W_b$ ,  $g(W_b) = 0$  e concluímos então que:

Se  $W < W_b$ , então  $g(W) > 0$

Se  $W > W_b$ , então  $g(W) < 0$

e conseqüentemente:  $f(W) > 0$  se  $W < W_b$ , e

$f(W) < 0$  se  $W > W_b$ .

Portanto, concluímos que o discriminante (3.9) é positivo se  $W < W_b$  e negativo se  $W > W_b$ .

Tomemos  $W$  no intervalo  $(0, W_b)$ . Desde que o discriminante é positivo, a equação (3.8) tem duas raízes reais, onde

$$r(W, S(B)) = S(Y_w) = \frac{(WS(B) + (W^2(S(B))^2 - 4(W-1))^{1/2})^{1/2}}{2}$$

Provaremos agora que  $S(Y_w)$  é decrescente no intervalo considerado.

$$\frac{d S(Y_w)}{dw} = S(Y_w)^{1/2} \frac{(S(B) \cdot ((W \cdot S(B))^2 - 4(W-1))^{1/2} + (S(B))^2 W - 2)}{(W^2(S(B))^2 - 4(W-1))^{1/2}}$$

Analisando  $S(Y_w)$  concluímos que  $[S(Y_w)]^{1/2} > 0$ . Além disso, como as raízes são reais, o discriminante é positivo, e portanto o denominador da fração é positivo.

O numerador da fração, satisfaz:

$S(B) (W^2 (W(B))^2 - 4(W-1))^{1/2} > 0$  e  $(S(B))^2 W - 2 < 0$ , já que  $W < 2$  e  $(W(B))^2 < 1$ . Por outro lado,  $((S(B))^2 \cdot W - 2)^2 = W^2 (W(B))^4 - 4W(S(B))^2 + 4$  e,  $(S(B) (W^2 (S(B))^2 - 4(W-1))^{1/2})^2 = W^2 (S(B))^4 - 4W(S(B))^2 + 4S(B)$ . Então,

$$((S(B))^2 W - 2)^2 - (S(B) (W^2 (S(B))^2 - 4(W-1))^{1/2})^2 = (4(1-S(B))) > 0$$

$$\therefore ((S(B))^2 W - 2)^2 > (S(B) (W^2 (S(B))^2 - 4(W-1))^{1/2})^2$$

$$\therefore |(S(B))^2 W - 2| > |(S(B) (W^2 (S(B))^2 - 4(W-1))^{1/2})|$$

Podemos então concluir que o numerador da fração é negativo, e portanto  $\frac{d S(Y_w)}{dW} < 0$  o que leva a concluir que  $S(Y_w)$  é decrescente no intervalo  $(0, W_b)$ , ou em outras palavras, que  $S(Y_{wb}) \leq S(Y_w)$  para todo  $W \leq W_b$ .

Consideremos agora os valores de  $W$  no intervalo  $(W_b, 2)$ . Desde que, para valores de  $W$  neste intervalo o discriminante (3.9) é negativo, a equação (3.8) tem duas raízes complexas conjugadas cujo produto é  $(W-1)^2$ .

Sejam  $p + qi$  e  $p - qi$  estas raízes

$$(p+qi)(p-pi) = (W-1)^2$$

$$\therefore (p^2 + q^2) = (W-1)^2$$

$$\therefore (p^2 + q^2)^{1/2} = ((W-1)^2)^{1/2}$$

$$\therefore |p + qi| = |p - qi| = (p^2 + q^2)^{1/2} = ((W-1)^2)^{1/2} = (W-1)$$

Se  $W > W_b$  então  $W > 1$  e  $((W-1)^2)^{1/2} = (W-1)$ . Então, se  $W > W_b$  as raízes têm módulo  $(W-1)$  e como são iguais os seus módulos, con

clui-se que  $r(W, S(B)) = W-1$ .  $S(Y_w) = W-1$ . Sendo  $S(Y_w) = W-1$  para  $W \in (W_b, 2)$ , é claro que  $S(Y_w)$  é crescente neste intervalo e portanto  $S(Y_w) \geq S(Y_{wb})$ , se  $W \geq W_b$ . Então, concluímos que  $S(Y_{wb}) \leq S(Y_w)$  se  $W \in (0, 2)$ , e portanto  $w_b$  minimiza  $S(Y_w)$ , sendo pois o valor ótimo de  $W$ .

2) Seja  $S(B) = 0$ .

Neste caso  $W_b = 1$ .

Por outro lado, se  $S(B) = 0$ , então  $S(Y) = S(Y_1) = 0$  [14]. Então  $S(Y_{wb}) = 0$ .

Dêsde que  $S(Y_w) \geq 0$ , podemos afirmar que  $S(Y_w) \geq S(Y_{wb})$  e portanto  $W_b = 1$  minimiza  $S(Y_w)$  sendo o valor ótimo de  $W$ .

### 3.3. Comportamento de $S(Y_w)$

Já se viu que  $W_b$  minimiza  $S(Y_w)$  nas condições do (teorema 3.1). Assim sendo,  $S(Y_w)$  é crescente no intervalo  $[W_b, 2)$  e decrescente no intervalo  $(0, W_b]$ .

Será agora feita uma análise da taxa de variação de  $S(Y_w)$  nestes dois intervalos.

#### 1. Intervalo $[W_b, 2)$

Neste intervalo  $S(Y_w)$  cresce à medida que  $W$  se afasta do valor  $W_b$ . Viu-se que para estes valores o discriminante de equação (3.8) é negativo e todas as suas raízes são complexas e de módulo  $(W-1)$ . Assim, a taxa de crescimento de  $S(Y_w)$  será dada por:

$$\frac{d(S(Y_w))}{dW} = \frac{d(W-1)}{dW} = 1$$

Conclui-se que  $S(Y_w)$  cresce linearmente.

Quando  $W$ , a partir de  $W_b$  sofre um acréscimo igual a  $k$ ,  $S(Y_w)$  é acrescido de  $k$ .

A taxa de convergência neste caso, decresce de

$$\frac{(R(Y_{w+b}) - R(Y_w))}{R(Y_{w+b})} \times 100\% .$$

## 2. Intervalo $(0, W_b]$

À medida que  $W$  se aproxima de  $W_b$  neste intervalo,  $S(Y_w)$  cresce.

Para estes valores de  $W$  o discriminante da equação (3.8) é positivo e todas as raízes são reais.

Viu-se que:

$$S(Y_w) = \frac{(W(S(B)) + [W^2(S(B))^2 - 4(W-1)]^{1/2})}{2} .$$

Assim, a taxa de crescimento de  $S(Y_w)$  será dada por:

$$\frac{dS(Y_w)}{dW} = 2 S(Y_w)^{1/2} \times \frac{1}{2} d\left(\frac{WS(B) + (W^2(S(B))^2 - 4(W-1))^{1/2}}{dW}\right)$$

$$\frac{dS(Y_w)}{dW} \times \frac{1}{S(Y_w)^{1/2}} = \frac{d[W(S(B)) + W^2(S(B))^2 - 4(W-1)]^{1/2}}{dW} =$$

$$= S(B) + \frac{1}{2} (W^2(S(B))^2 - 4(W-1))^{-1/2} (2W(S(B))^2 - 4) =$$

$$= S(B) + \frac{W(S(B))^2 - 2}{(W^2(S(B))^2 - 4(W-1))^{1/2}} =$$

$$= S(B) - \frac{[2 - W((S(B))^2)]}{(W^2(S(B))^2 - 4(W-1))^{1/2}} =$$

$$= \frac{S(B)(W^2(S(B))^2 - 4(W-1))^{1/2} - [2 - W((S(B))^2)]}{(W^2(S(B))^2 - 4(W-1))^{1/2}}$$

Multiplicando o numerador e denominador da expressão por  $S(B)[(W^2(S(B))^2 - 4(W-1))^{1/2} + (2 - W(S(B)))]$  tem-se:

$$\frac{dS(Y_w)}{dW} \cdot \frac{1}{(S(Y_w))^{1/2}} = \frac{-4[1 - (S(B))^2]}{(W^2(S(B))^2 - 4(W-1))^{1/2} \cdot S(B)[(W^2(S(B))^2 - 4(W-1))^{1/2} + 2 - W(S(B))]}$$

Será feita a seguir uma análise da expressão acima:

Seja a função

$$f(W)^{1/2} = [W^2(S(B))^2 - 4(W-1)]^{1/2}$$

Desde que  $f(W)$  é contínua tem-se que:  $\lim_{W \rightarrow W_b} [W^2(S(B))^2 - 4(W-1)]^{1/2} = 0$  pois a expressão é o discriminante de (3.8). Assim, o denominador da fração tende a zero quando  $W \rightarrow W_b$  e daí

$$\lim_{W \rightarrow W_b} \frac{1}{S(Y_w)^{1/2}} \cdot \frac{dS(Y_w)}{dW} = -\infty$$

$$\lim_{W \rightarrow W_b} \frac{dS(Y_w)}{dW} = -\infty$$

Conclui-se que quando  $W$  se aproxima de  $W_b$ , o raio espectral de  $Y_w$  decresce numa taxa extraordinariamente grande, e portanto um pequeno decréscimo de  $W_b$  acarreta um grande decréscimo de  $S(Y_w)$ .

Assim, incrementar  $W_b$  de uma pequena quantidade, causa um menor decréscimo em  $S(Y_w)$  e conseqüentemente no número de iterações necessário para que se obtenha a convergência do método, do que diminuir  $W_b$  de uma quantidade equivalente.

Os gráficos a seguir,  $W$  versus  $S(Y_w)$  e  $W$  versus  $R(S(Y_w))$ , foram traçados com base na matriz

$$\begin{bmatrix} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & 4 \end{bmatrix}$$

Estes dão uma idéia melhor do que foi exposto acima.

A TABELA 1 fornece os valores de  $W$ ,  $S(Y_w)$  e  $R(S(Y_w))$  utilizados no traçado dos gráficos 1 e 2.

TABELA 1

$W$	$S(Y_w)$	$R(S(Y_w))$
0.1	0.949	0.0526
0.2	0.895	0.1114
0.3	0.837	0.1776
0.4	0.776	0.2533
0.5	0.711	0.3414
0.6	0.640	0.4567
0.7	0.562	0.5753
0.8	0.476	0.7423
0.9	0.376	0.9784
1.0	0.250	1.3800
1.071 ( $W_b$ )	0.071	2.6339
1.1	0.100	2.3000
1.2	0.200	1.6000
1.3	0.300	1.2000
1.4	0.400	0.9160
1.5	0.500	0.6931
1.6	0.600	0.5108
1.7	0.700	0.3566
1.8	0.800	0.2231

$$S(Y_w) = (W-1) \quad , \quad W \geq W_b$$

$$S(Y_w) = \left( \frac{W(S(B)) + [W^2(S(B))^2 - 4(W-1)]^{1/2}}{2} \right)^2, \quad W < W_b$$

$$R(S(Y_w)) = -\log S(Y_w)$$

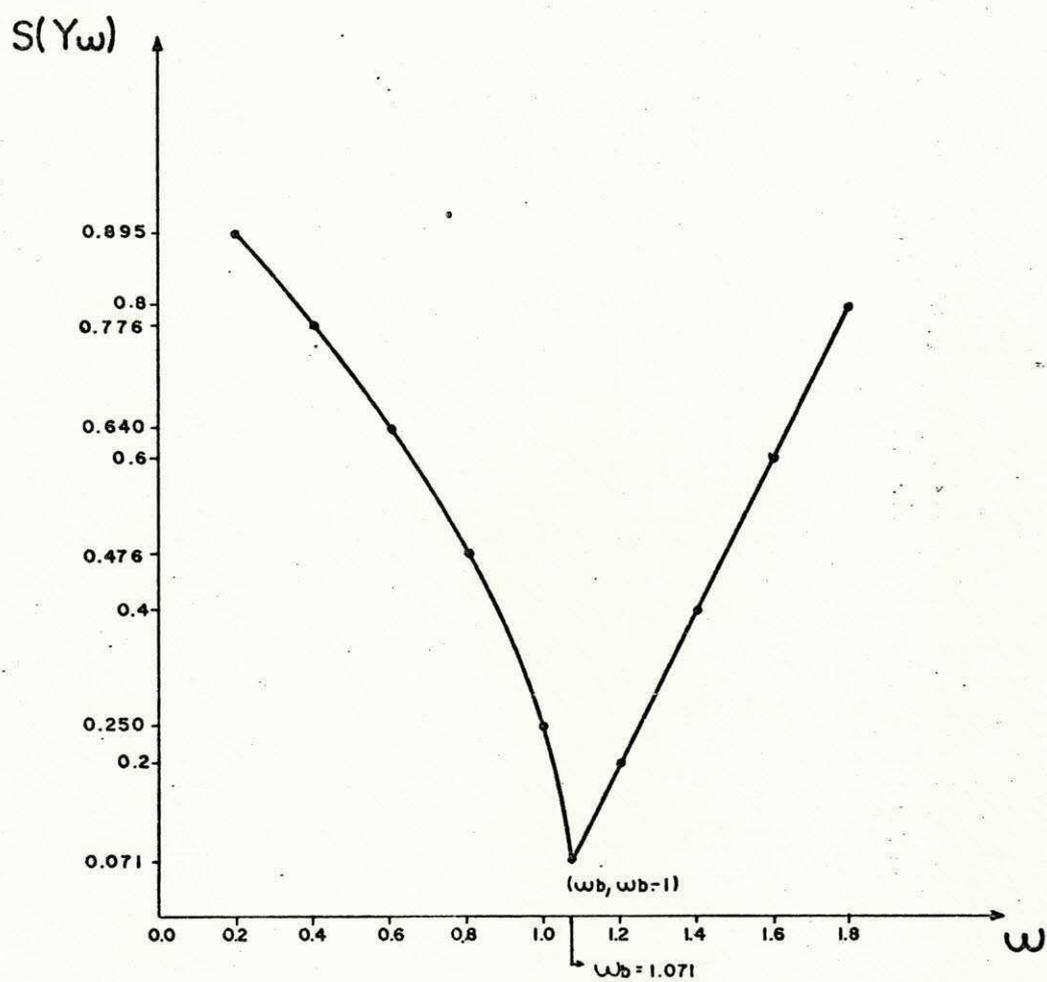


GRÁFICO I

DEMONSTRATIVO DA VARIAÇÃO DE  $S(Y_\omega)$  COM  $\omega$ , PARA  $0,2 \leq \omega \leq 1,80$ , PARA O MÉTODO SOR

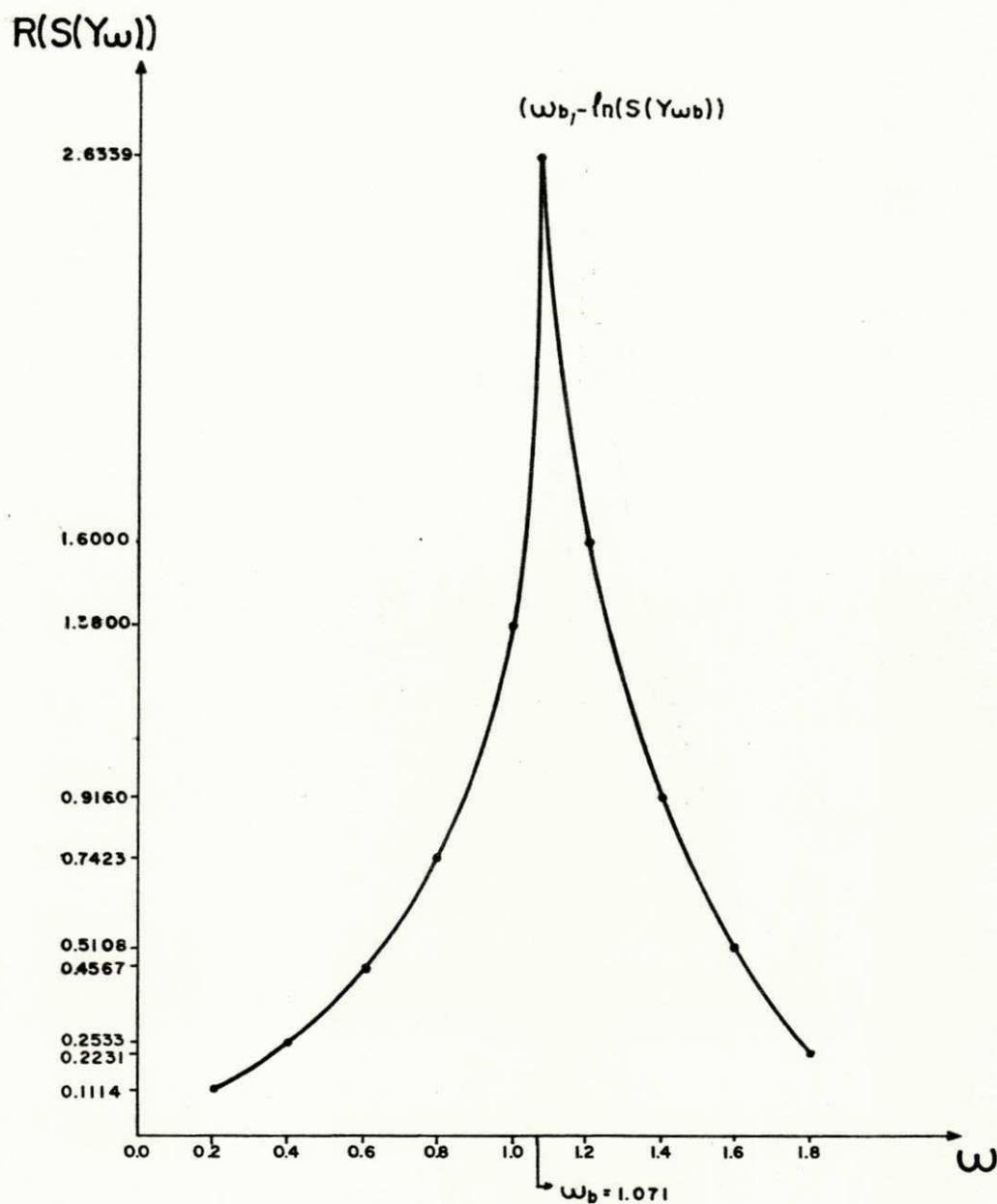


GRÁFICO 2

DEMONSTRATIVO DA VARIAÇÃO DA TAXA DE  
CONVERGÊNCIA DO MÉTODO SOR COM  $\omega$ ,  
PARA  $0.2 \leq \omega \leq 1.8$

## CAPÍTULO IV

### IMPLEMENTAÇÃO , TESTES COMPARATIVOS E CONCLUSÕES

#### 4.1. Implementação

Com o objetivo de poupar esforços em se desenvolver rotinas já elaboradas bem como complementar este trabalho, neste capítulo apresentamos as rotinas de aplicação dos métodos iterativos tratados. Além destas, constam também algumas rotinas de serviço cujo desenvolvimento foi indispensável. Dentre estas rotinas de apoio, algumas se destinam a testar certas propriedades de matrizes, tendo certa relevância a subrotina DANI7 e a subrotina POSID. Estas rotinas são importantes por serem elas destinadas a verificar se uma matriz tem as propriedades que são condições necessárias à matriz dos coeficientes para se aplicar o método SOR na solução do sistema com o fator de relaxação ótimo definido pela equação 3.1.

## 1. SUBROTINAS DOS MÉTODOS

### 1.1. Objetivo

Estas rotinas se destinam à determinação da solução única de um sistema de equações lineares pelos métodos de Jacobi, JOR, Gauss-Seidel e SOR. Os nomes destas subrotinas são respectivamente JACOB, JOR, GSEID e SOR.

### 1.2. Formato de Chamada

Cada uma das subrotinas obedece a um dentre os dois tipos de formato de chamada.

CALL NOME (NDIM,ND,AM,VC,NIT,IND,SOLF,APR,T) (4.1)

CALL NOME (NDIM,ND,AM,VC,NIT,IND,SOLF,APR,T,W) (4.2)

### 1.3. Descrição dos Parâmetros

NDIM - Indica o número de linhas da matriz como é declarado no comando de especificação de dimensão do programa de chamada.

ND - Variável inteira cujo valor é igual à dimensão do sistema a ser resolvido.

AM - Matriz real de ordem ND que representa a matriz dos coeficientes do sistema.

VC - Vetor constante do sistema, de dimensão ND.

NIT - Variável inteira que indica o número de iterações.

IND - Variável inteira que indica se o método converge. O seu valor é inicialmente zero e assumirá valor um se não houver convergência.

SOLF - Vetor de dimensão ND que representa a solução do sistema.

APR - Vetor de dimensão ND que representa uma aproximação da solução obtida em uma iteração.

T - Tolerância na aproximação da solução do sistema.

W - Fator de relaxação para o método JOR.

## 2. SUBROTINAS DE SERVIÇO

### 2.1. Subrotinas de Teste de Propriedades

#### 2.1.1. Objetivo

Estas rotinas determinam as propriedades de uma matriz, dentre aquelas visadas no estabelecimento dos critérios de convergência. Dentre elas destacam-se como mais relevantes:

DANI - Testa a irredutibilidade.

POSID - Testa se uma matriz é positiva definida, através de seus autovalores.

DANI7 - Verifica se uma matriz é consistentemente ordenada.

#### 2.1.2. Formato de Chamada

As subrotinas obedecem ao seguinte formato de chamada:

CALL NOME (NDIM, ND, AM, IP)

### 2.1.3. Descrição dos Parâmetros

NDIM - Indica o número de linhas da matriz como é declarado no comando de especificação de dimensão do programa de chamada.

ND - Variável inteira cujo valor é igual à dimensão da matriz que se testa a propriedade.

AM - Matriz real de ordem ND que será testada.

IP - Variável inteira cujo valor indica se a matriz tem a propriedade em questão.

## 2.2. Subrotinas Gerais de Serviços

### 2.2.1. Objetivo

Oferecer apoio às demais subrotinas.

TRASF - Transformar quando possível a matriz dos coeficientes de um sistema em uma matriz de diagonal dominante através da troca de linhas, efetuando as correspondentes trocas no vetor independente.

RELAX - Determina o fator de relaxação ótimo para o método SOR.

### 2.2.2. Formato de Chamada

```
CALL TRASF (NDIM,ND,AM,VC,IND)
```

```
CALL RELAX (NDIM,ND,B,W)
```

### 2.2.3. Descrição dos Parâmetros

NDIM - Indica o número de linhas da matriz como é declarado no

comando de especificação de dimensão do programa de chamada.

ND - Variável inteira cujo valor é igual à dimensão do sistema.

AM - Matriz real de ordem ND que representa a matriz dos coeficientes do sistema.

VC - Vetor real constante do sistema de dimensão ND.

IND - Variável inteira que indica se uma troca de linhas é possível.

B - Matriz do método de Jacobi associada à matriz dos coeficientes do sistema.

W - Fator de relaxação.

## 4.2. Testes

Nesta seção são apresentados alguns resultados numéricos, confirmando as previsões.

Estes resultados apresentam-se descritos a seguir.

4.2.1. A taxa de convergência para o SOR é máxima para  $W_{\text{óti}}$  mo, calculado de acordo com o que foi estabelecido no Capítulo III.

Considere-se as quatro matrizes positivas definidas e consistentemente ordenadas apresentadas a seguir.

Matriz 1

$$\begin{bmatrix} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & 4 \end{bmatrix}$$

Matriz 2

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

Matriz 3

$$\begin{bmatrix} 4 & -1 & 0 & 0 & -1 & 0 \\ -1 & 4 & -1 & 0 & 0 & -1 \\ 0 & -1 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & -1 & 0 \\ -1 & 0 & 0 & -1 & 4 & -1 \\ 0 & -1 & 0 & 0 & -1 & 4 \end{bmatrix}$$

Matriz 4

$$\begin{bmatrix} 4 & 0 & 1 & 1 \\ 0 & 4 & 0 & 1 \\ 1 & 0 & 4 & 0 \\ 1 & 1 & 0 & 4 \end{bmatrix}$$

De acordo com o que foi visto, o método SOR é convergente para qualquer sistema que tenha matrizes dos coeficientes com as propriedades acima citadas. Além disso é possível se determinar um fator ótimo de relaxação pela equação (3.1).

Aplicando o SOR aos sistemas 1, 2, 3 e 4 que têm as quatro matrizes em questão como matrizes dos coeficientes e os vetores constantes abaixo:

VETOR 1

$$\begin{bmatrix} 0 \\ 0 \\ 10^3 \\ 10^3 \end{bmatrix}$$

VETOR 2

$$\begin{bmatrix} 6 \times 10^3 \\ 0 \\ 0 \end{bmatrix}$$

VETOR 3

$$\begin{bmatrix} -1 \\ 5 \\ 3 \\ -2 \\ 8 \\ 0 \end{bmatrix}$$

VETOR 4

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

Constata-se a convergência, sendo apresentados no quadro a seguir o fator de relaxação ótimo para cada caso.

SISTEMA	$W_{\text{ótimo}}$
1	1.071797
2	1.072627
3	1.09340
4	1.044640

QUADRO 3

Resolvendo-se em seguida os sistemas com fatores de relaxação diferentes do ótimo, fazendo-se  $W$  variar de 0.6 a 1.8 com um incremento de 0.2, pode-se comprovar através dos valores das taxas de convergência calculadas em cada caso que esta convergência ocorre mais rapidamente quando se usa o  $W$  ótimo.

Verifica-se ainda que a taxa de convergência diminui à medida que  $W$  se afasta do valor ótimo tanto para a esquerda quanto para a direita, como ilustra o quadro 4.

O número de iterações necessárias à convergência em cada caso para cada valor de  $W$ , é apresentado no quadro 4 e comprovadamente diminui à medida que se utiliza um fator  $W$  cada vez mais próximo do valor ótimo.

W	SISTEMA 1		SISTEMA 2		SISTEMA 3		SISTEMA 4	
	$S(Y_w)$	$R(S(Y_w))$	$S(Y_w)$	$R(S(Y_w))$	$S(Y_w)$	$R(S(Y_w))$	$S(Y_w)$	$R(S(Y_w))$
0.1	0,9487	0,0526	0,9489	0,0524	0,9546	0,0464	0,9392	0,0627
0.2	0,8945	0,1114	0,8951	0,1108	0,9064	0,0982	0,8757	0,1327
0.3	0,8372	0,1776	0,8380	0,1767	0,8550	0,1565	0,8092	0,2117
0.4	0,7762	0,2533	0,7772	0,2521	0,7980	0,2231	0,7391	0,3023
0.5	0,7107	0,3414	0,7120	0,3397	0,7405	0,3003	0,6649	0,4080
0.6	0,6400	0,4567	0,6414	0,4439	0,6757	0,3919	0,5857	0,5348
0.7	0,5624	0,5753	0,5642	0,5723	0,6041	0,5039	0,5003	0,6926
0.8	0,4759	0,7423	0,4780	0,7383	0,5236	0,6470	0,4063	0,9008
0.9	0,3758	0,9784	0,3781	0,9727	0,4298	0,8442	0,2991	1,2069
1.0	0,2500	1,3800	0,2525	1,3763	0,3125	1,1631	0,1636	1,8101
WÓTIMO	0,071797	2,6339	0,72627	2,6224	0,0934	2,3708	0,04464	3,1091
1.1	0.1	2,3000	2,3000	2,3000	2,3000	2,3000	2,3000	2,3000
1.2	0.2	1,6000	1,6000	1,6000	1,6000	1,6000	1,6000	1,6000
1.3	0.3	1,2000	1,2000	1,2000	1,2000	1,2000	1,2000	1,2000
1.4	0,4	0,9160	0,9160	0,9160	0,9160	0,9160	0,9160	0,9160
1.5	0,5	0,6931	0,6931	0,6931	0,6931	0,6931	0,6931	0,6931
1.6	0,6	0,5108	0,5108	0,5108	0,5108	0,5108	0,5108	0,5108
1.7	0,7	0,3566	0,3566	0,3566	0,3566	0,3566	0,3566	0,3566
1.8	0,8	0,2231	0,2231	0,2231	0,2231	0,2231	0,2231	0,2231

$$S(Y_w) = \frac{W S(B) + (W^2 S(B)^2 - 4(W-1))^{1/2}}{2} \quad \text{se } W < \text{Wóximo}, S(Y_w) = (W-1) \text{ se } W \geq \text{Wóximo}.$$

	NÚMERO DE ITERAÇÕES			
	SISTEMA 1	SISTEMA 2	SISTEMA 3	SISTEMA 4
0.6	28	34	18	12
0.8	18	22	12	9
1.0	11	13	8	6
Wótimo	8	9	6	5
1.2	10	11	7	6
1.4	16	18	10	10
1.6	28	31	17	16
1.8	>50	>50	38	35

QUADRO 5

Neste quadro é apresentado o número de iterações necessárias para a convergência do método SOR, considerando-se uma variação do fator de relaxação de 0.6 a 1.8, para os sistemas 1, 2, 3 e 4. A tolerância utilizada é de  $0.1 \times 10^{-2}$ .

4.2.2. Aplicação dos métodos de Jacobi, Gauss-Seidel, JOR e SOR.

Consideremos os quatro sistemas de equações lineares em questão no ítem anterior.

O quadro 6 apresenta os resultados obtidos quando se aplica os métodos de Jacobi, JOR, SOR e Gauss-Seidel aos quadro sistemas, com uma tolerância de  $0.1 \times 10^{-2}$ .

MÉTODO	NÚMERO DE ITERAÇÕES			
	SISTEMA 1	SISTEMA 2	SISTEMA 3	SISTEMA 4
JACOBI	18	24	14	9
JOR	17	33	19	11
G.-SEIDEL	11	13	8	6
SOR	8	9	6	5

QUADRO 6

Conforme se pode observar, nestes casos onde as matrizes dos coeficientes são positivas definidas e consistentemente ordenadas, a solução pelo método SOR com o  $W$  ótimo, é encontrada com um número de iterações inferior ao número de iterações correspondentes às soluções pelos outros três métodos.

Pode-se notar também que, o método JOR não funcionou como uma melhoria do método de Jacobi exceto para o primeiro sistema. Isto deu-se devido a má escolha do fator  $W$  que foi feita aleatoriamente. A escolha conveniente deste fator que não foi tratado neste texto deverá diminuir o número de iterações necessárias para atingir a convergência quando se compara os métodos JOR e de Jacobi.

#### 4.2.3. Aplicação da Subrotina TRASF.

Consideremos os seguintes sistemas, representados por sua matriz dos coeficientes e seu vetor independente.

$$A_1 = \begin{bmatrix} 1 & 10 & 0 & 0 \\ 0 & 1 & 0 & 10 \\ 10 & -1 & 0 & 0 \\ 0 & 0 & 2 & -1 \end{bmatrix} \quad b_1 = \begin{bmatrix} 2 \\ 10 \\ 10 \\ 19 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 1 & 10 & 1 & 0 & 0 & 0 \\ 2 & 0 & 20 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 30 & 3 \\ 10 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 2 & -2 & 20 \\ 0 & 0 & 0 & 10 & -1 & 0 \end{bmatrix} \quad b_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 5 \\ 5 \\ 5 \end{bmatrix}$$

Conforme pode-se observar nenhum dos métodos iterativos pode ser aplicado para solucionar estes sistemas já que nas diagonais das matrizes  $A_1$  e  $A_2$  existem elementos nulos.

Aplicando-se a subrotina TRASF aos sistemas, consegue-se transformá-los em sistemas equivalentes cujas matrizes dos coeficientes têm diagonal estritamente dominante, conforme apresentamos abaixo. Além disso estas matrizes passam a ter certos conjuntos de propriedades que permitam a escolha de um método iterativo para a solução do sistema.

$$A_1 = \begin{bmatrix} 10 & -1 & 0 & 0 \\ 1 & 10 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 10 \end{bmatrix} \quad b_1 = \begin{bmatrix} 10 \\ 2 \\ 19 \\ 10 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 10 & 1 & 0 & 0 & 0 & -1 \\ 1 & 10 & 1 & 0 & 0 & 0 \\ 2 & 0 & 20 & 1 & 0 & 0 \\ 0 & 0 & 0 & 10 & -1 & 0 \\ 0 & 3 & 0 & 0 & 30 & 3 \\ 0 & 0 & 0 & 2 & -2 & 20 \end{bmatrix} \quad b_2 = \begin{bmatrix} 5 \\ 1 \\ 1 \\ 5 \\ 0 \\ 5 \end{bmatrix}$$

Matrizes  $A_1$  e  $A_2$  e vetores  $b_1$  e  $b_2$  com as linhas permutadas pela subrotina TRASF.

Apresentamos a seguir, a solução dos sistemas pelo método de Jacobi.

$$u_1 = \begin{bmatrix} 1.009900 \\ 0.09901005 \\ 9.549504 \\ 0.9900991 \end{bmatrix}$$

$$u_2 = \begin{bmatrix} 0.5146590 \\ 0.05116820 \\ -0.02634139 \\ 0.4975107 \\ -0.02489278 \\ 0.1977597 \end{bmatrix}$$

### 4.3. Conclusões

Após o estudo, implementação e testes dos algoritmos apresentados neste trabalho verificamos as vantagens do uso dos métodos iterativos nos sistemas de equações onde a matriz dos coeficientes é esparsa, estruturada e de grande porte. Não se pode contudo esperar superioridade absoluta desses métodos sobre os métodos diretos a não ser nos casos acima citados.

Podemos também concluir que apesar das restrições de convergência os métodos iterativos podem ser largamente utilizados visto que é grande o número de sistemas onde a matriz dos coeficientes apresenta as propriedades necessárias para a utilização dos mesmos. Um exemplo disto é que a resolução de equações diferenciais elípticas, parabólicas ou hiperbólicas pelo método numérico das diferenças finitas sempre conduz a sistemas de equações lineares onde é possível e vantajoso o uso dos métodos iterativos.

Verificou-se também que entre os métodos iterativos não existe supremacia absoluta. Isto depende das propriedades da matriz dos coeficientes do sistema. Há casos por exemplo em que ao contrário do que se poderia esperar o método de Jacobi oferece maior rapidez de convergência do que o método de Gauss-Seidel. Comparando-se o método de Gauss-Seidel com o SOR encontra-se casos onde este pode não ser mais eficiente que o primeiro qualquer que seja o fator de relaxação escolhido [11]. Além disso em casos onde esta maior eficiência pode ocorrer não se pode garanti-la a não ser com a utilização do fator  $W$  que minimiza o raio espectral de  $Y_w$ . A solução de equações parciais elípticas são casos típicos onde a utilização do SOR é sempre mais eficiente do que a do Gauss-Seidel quando se usa  $W$  calculado pela equa

ção (3.1).

Nenhuma referência sobre a escolha adequada do fator pa  
ra o método JOR consta neste texto. Sugerimos como assunto para  
outro trabalho o estudo da utilização do método JOR com a melhor  
eficiência possível.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ALBRECHT, P., Análise Numérica, Livros Técnicos e Científicos Editora S.A., Rio de Janeiro, 1973.
2. BOLDRINI, J.L., Costa, S.I., RIBEIRO, V.L., WETZLER, H.G., Álgebra Linear, Harba, Campinas, 1978.
3. CULLEN, C.G., Matrices and Linear Transformations, Addison - Wesley Publishing Company, Inc., Pittsburgh, 1967.
4. FADDEEVA, V.N., Computational Methods of Linear Algebra, Dover Publications, Inc. New York, 1959.
5. GERVREY M., Journal de Mathématiques Pures et Appliquées. Sixième Serie - Tome neuvième, 1913
6. GERVREY M., Idem - Tome Dixième, 1914
7. HOUSEHOLDER, ALSTON S., The theory of Matrices in Numerical Analysis, Dover Publications, Inc. New York, 1975.
8. MURDOCH, D.C., Álgebra Linear, Livros Técnicos e Científicos Editora Ltda., Rio de Janeiro, 1972.
9. RALSTON; A., A First Course in Numerical Analysis, Mc Graw-Hill Book Company, New York, 1965.
10. SANTOS, V.R.B., Curso de Cálculo Numérico, Ao Livro Técnico S.A., Rio de Janeiro, 1972.
11. STRANG, G., Linear Álgebra and Its Applications, Academic Press, New York, 1976.

12. VARGA, R.S., *Matrixe Iterative Analysis*, Prentice Hall, Englewood Cliffs, N. Jersey, 1962.
13. WILKINSON, J.H., *The Algelraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.
14. YOUNG, D.M., *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.
15. YOUNG, D.M., GREGORY, R.T., *A Survey of Numerical Mathematics*, Vol II, Addison-Wesley Publishing Company, 1972.