



Universidade Federal de Uberlândia
Instituto de Matemática e Estatística

Coordenação do Curso de Matemática

**MÉTODO DE DIFERENÇAS FINITAS
PARA PROBLEMAS DE VALOR DE
CONTORNO**

Edinilson Ferreira Vilela

Uberlândia-MG

2024

Edinilson Ferreira Vilela

**MÉTODO DE DIFERENÇAS FINITAS
PARA PROBLEMAS DE VALOR DE
CONTORNO**

Monografia apresentada ao Curso de Graduação em Matemática da Universidade Federal de Uberlândia, como parte dos requisitos para a obtenção de título de **LICENCIATURA EM MATEMÁTICA**.

Área de concentração: Matemática

Linha de pesquisa: Análise Numérica

Orientador: Santos Alberto Enriquez Remigio

Uberlândia-MG

2024



**Universidade Federal de Uberlândia
Instituto de Matemática**

Coordenação do Curso de Matemática

A banca examinadora, conforme abaixo assinado, certifica a adequação deste trabalho de conclusão de curso para obtenção do grau de Licenciatura em Matemática.

Uberlândia, _____ de _____ de 20_____

BANCA EXAMINADORA

Santos Alberto Enriquez Remigio

Josuel Kruppa Rogenski

Rafael Alves Figueiredo

**Uberlândia-MG
2024**

AGRADECIMENTOS

Dedico aqui meu agradecimento aos meus familiares, meu pai, minha mãe, meus irmãos e irmãs, por sempre me apoiarem e incentivarem, e também à todos os meus amigos.

Ao meu orientador, Santos Alberto Enriquez Remigio, que sempre esteve muito disposto a me ajudar e ensinar, e com sua ajuda possibilitou que esse trabalho fosse possível de ser realizado.

Ao professor Alonso Sepúlveda, o qual foi meu orientador de iniciação científica durante o início do curso e sempre me incentivou a continuar em frente.

À toda a coordenação do curso que sempre foi muito dispostas a ajudar a resolver qualquer dúvida ou problema.

Aos os diversos docentes do Instituto de Matemática e Estatística da UFU que sempre me ajudaram. E ao aos vários amigos e colegas os quais tive a oportunidade conhecer na Universidade Federal de Uberlândia.

RESUMO

Os problemas de valor de contorno aparecem em diversas áreas das ciências e engenharias, modelando fenômenos do cotidiano, como a distribuição de temperatura, a deformação de estruturas e o comportamento de campos eletromagnéticos.

Matematicamente, esses problemas são descritos por equações diferenciais ordinárias ou parciais, as quais frequentemente não possuem solução analítica. Isso torna necessário o uso de métodos numéricos para obter aproximações das soluções. Neste trabalho, o método das diferenças finitas foi escolhido como abordagem principal.

O objetivo principal deste trabalho é introduzir o método das diferenças finitas, com foco em seus fundamentos matemáticos e na aplicação à solução de problemas de valor de contorno de segunda ordem em uma e duas dimensões, sujeitos a condições de contorno do tipo Dirichlet. Também são discutidos métodos iterativos estacionários clássicos (Jacobi, Gauss-Seidel e o método Sobre-relaxação Sucessiva) para a resolução dos sistemas de equações lineares resultantes da aplicação do método das diferenças finitas. Em particular, demonstra-se que o uso de diferenças finitas centradas de segunda ordem para aproximar derivadas de segunda ordem garante que a solução numérica convirja com a mesma ordem de precisão.

Adicionalmente, é realizada a análise da convergência das soluções dos sistemas lineares resolvidos pelos métodos iterativos, com foco na equação de Poisson associada ao método das diferenças finitas. Resultados numéricos são apresentados para verificar e validar a teoria discutida, demonstrando a eficácia da abordagem.

Palavras-chave: método das diferenças finitas, equações diferenciais, problema de valor de contorno, equação de Poisson, métodos iterativos.

ABSTRACT

Boundary value problems arise in various fields of science and engineering, modeling everyday phenomena such as temperature distribution, structural deformation, and the behavior of electromagnetic fields.

Mathematically, these problems are described by ordinary or partial differential equations, which often lack analytical solutions. This makes necessary the use of numerical methods to obtain approximate solutions. In this work, the finite difference method was chosen as the main approach.

The main objective of this study is to introduce the finite difference method, focusing on its mathematical foundations and its application to solving second-order boundary value problems in one and two dimensions, subject to Dirichlet boundary conditions. Classical stationary iterative methods (Jacobi, Gauss-Seidel, and the Successive Over-relaxation method) for solving the linear systems resulting from the finite difference method are also discussed. In particular, it is shown that using second-order centered finite differences to approximate second-order derivatives ensures that the numerical solution converges with the same order of accuracy.

Additionally, an analysis of the convergence of the solutions of linear systems solved by iterative methods is conducted, with a focus on the Poisson equation associated with the finite difference method. Numerical results are presented to verify and validate the discussed theory, demonstrating the effectiveness of the approach.

Keywords: finite difference method, differential equations, boundary value problem, Poisson's equation, iterative methods.

SUMÁRIO

1	Introdução	1
1.1	Equações Diferenciais	1
1.2	Condições de Contorno	3
1.3	Método das Diferenças Finitas	4
2	Problema unidimensional	7
2.1	Diferenças finitas	7
2.2	Aplicação do método	12
2.3	Erro Local e Global	14
2.4	Consistência	15
2.5	Estabilidade	16
2.6	Convergência	18
3	Problema bidimensional	19
3.1	Aplicação do método	19
3.2	Ordenação dos pontos	23
3.3	Análise do método	26
3.4	Princípio do máximo	28
4	Métodos iterativos	36
4.1	Conceitos preliminares	36
4.2	Métodos iterativos estacionários	39
4.3	Método de Jacobi	41
4.4	Método de Gauss-Seidel	43
4.5	Método de Sobre-relaxação Sucessiva	45
4.6	Matrizes consistentemente ordenadas	48
5	Resultados numéricos	55
5.1	Conceitos Preliminares	55
5.2	Problema unidimensional	56
5.3	Problema de bidimensional	60
6	Conclusões	63
	Referências Bibliográficas	65

1. INTRODUÇÃO

Os problemas de valor de contorno aparecem em diversas áreas da ciência e engenharia, representando fenômenos como a distribuição de temperatura, a deformação de estruturas e o comportamento de campos eletromagnéticos. Esses problemas são modelados por equações diferenciais, ordinárias ou parciais, para os quais nem sempre há solução analítica e daí a necessidade de métodos numéricos que permitam encontrar aproximações para a solução.

Apresentaremos neste capítulo uma breve introdução às equações diferenciais, também serão apresentados exemplos simples de problemas de valor de contorno, assim como uma rápida introdução ao método das diferenças finitas. Logo, apresentaremos os objetivos do trabalho.

1.1 EQUAÇÕES DIFERENCIAIS

Introduziremos aqui os conceitos básicos sobre equações diferenciais, aqui começaremos definindo as equações diferenciais ordinárias e expandiremos depois para as equações diferenciais parciais.

Uma equação diferencial ordinária de ordem k é uma equação que relaciona uma função desconhecida $u : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ com suas derivadas. Uma equação diferencial ordinária é dado pela seguinte forma geral:

$$\frac{d^k u}{dx^k} = f\left(x, u, \frac{du}{dx}, \dots, \frac{d^{k-1}u}{dx^{k-1}}\right),$$

onde f é uma função dada.

Como exemplos de equações diferenciais ordinárias, que podem ser encontrados em [2], temos o modelo populacional de Malthus, dado pela equação

$$\frac{dP}{dt} = kP,$$

que modela um crescimento de uma população $P = P(t)$ que cresce de forma proporcional, com uma constante de proporcionalidade k , ao tamanho da população.

Outro exemplo, na mecânica clássica, é a segunda lei de Newton, que é representada pela equação

$$m \frac{d^2 x}{dt^2} = F$$

onde $x = x(t)$ é a posição de uma partícula em um instante t , m é a massa da partícula e F a força resultante atuando sobre a partícula.

Ainda na mecânica clássica, o movimento de um pêndulo simples é governado pela equação

$$\frac{d^2\theta}{dt^2} + \frac{g}{l} \sin \theta = 0$$

onde $\theta = \theta(t)$ é o ângulo de abertura do pêndulo, g a aceleração da gravidade e l o comprimento do pêndulo.

Uma Equação Diferencial Parcial é uma equação que envolve uma função de várias variáveis que não se conhece, e as derivadas parciais dessa função [5].

Seja $u : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, chama-se Equação Diferencial Parcial de Ordem k uma expressão do tipo:

$$F(D_1^k u, \dots, D_n^k u, D_1^{k-1} u, \dots, D_n u, u) = 0$$

onde F é uma função dada.

O operador laplaciano ∇^2 é definido em termos das derivadas espaciais ($\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $\frac{\partial}{\partial z}$, por exemplo), mas é amplamente utilizado em problemas que envolvem tanto variáveis espaciais quanto temporais, como as Equações Diferenciais Parciais que descrevem fenômenos físicos. Assim, o laplaciano ∇^2 , em coordenadas cartesianas é definido por:

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2},$$

para u definido em uma região do plano $\Omega \subset \mathbb{R}^2$, e

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2},$$

para u definido em uma região do espaço $\Omega \subset \mathbb{R}^3$.

Como exemplos de equações diferenciais parciais, temos de acordo com [5], a equação do calor em um meio homogêneo, que é dada por

$$\frac{\partial u}{\partial t} = k \nabla^2 u,$$

onde k é a constante de condução térmica do meio e $u = u(x, y, z, t)$ a distribuição de temperatura.

A propagação de uma onda é descrita por uma equação diferencial parcial, a equação da onda,

$$\frac{\partial^2 u}{\partial t^2} = c^2 \nabla^2 u,$$

onde $u(x, y, z, t)$ representa a a perturbação gerada da onda, e c é a velocidade de propagação da onda.

Outros exemplo é a equação de Laplace,

$$\nabla^2 u = 0,$$

que pode ser usada, por exemplo, para modelar a distribuição de temperatura em um sistema estacionário $u = u(x, y, z)$. E a sua versão não homogênea, a equação de Poisson,

$$\nabla^2 u = f$$

que pode ser usada, por exemplo, para modelar o potencial de uma quantidade, neste caso $u = u(x, y, z)$ é o potencial, e f uma função dada.

1.2 CONDIÇÕES DE CONTORNO

Para se resolver um problema de Equações Diferenciais Parciais é preciso de impor condições auxiliares, podendo elas serem condições iniciais ou condições de contorno.

As condições de contorno podem ser do tipo Dirichlet, onde os valores de u são especificados na fronteira $\partial\Omega$, de Neumann, onde as derivadas normais de u são especificadas na fronteira $\partial\Omega$ e de Robin, onde são especificados na fronteira $\partial\Omega$ uma combinação dos valores de u e sua derivada normal. Podemos expressar essas condições como:

1. Condição de contorno de tipo Dirichlet:

$$u \Big|_{\partial\Omega} = g,$$

2. Condição de contorno de tipo Neumann:

$$\frac{\partial u}{\partial \mathbf{n}} \Big|_{\partial\Omega} = g,$$

3. Condição de contorno de tipo Robin:

$$\left(pu + q \frac{\partial u}{\partial \mathbf{n}} \right) \Big|_{\partial\Omega} = g,$$

onde p e q são funções dadas, g é uma função conhecida na borda do domínio $\partial\Omega$ e $\frac{\partial}{\partial \mathbf{n}}$ denota a derivada na direção normal da fronteira.

Quando resolve-se uma equação de Poisson sujeita a condições de contorno do tipo Neumann, surge um conceito denominado de condição de compatibilidade.

A condição de compatibilidade garante que as condições de contorno fornecidas sejam consistentes com a equação diferencial, permitindo a existência de uma solução para o problema.

Conforme [13], para a Equação de Poisson $\nabla^2 u = f$ definida em uma região do plano $\Omega \subset \mathbb{R}^2$, é necessário que satisfaça a seguinte condição, integrando ambos os lados da equação de Poisson sobre o domínio Ω , obtemos

$$\iint_{\Omega} \nabla^2 u \, dA = \iint_{\Omega} f \, dA,$$

reescrevendo o laplaciano como o divergente do gradiente, ou seja, $\nabla^2 u = \nabla \cdot (\nabla u)$, podemos aplicar o teorema do divergente ao lado esquerdo da equação, obtendo assim, uma integral para o fluxo da borda do domínio:

$$\iint_{\Omega} \nabla^2 u \, dA = \iint_{\Omega} \nabla \cdot (\nabla u) \, dA = \oint_{\partial\Omega} \nabla u \cdot \mathbf{n} \, dr.$$

Como $\frac{\partial u}{\partial \mathbf{n}} = \nabla u \cdot \mathbf{n}$, onde \mathbf{n} é o vetor normal a borda, então

$$\oint_{\partial\Omega} \nabla u \cdot \mathbf{n} \, dr = \oint_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} \, dr.$$

Logo a condição de contorno do tipo Neumann para a equação de Poisson $\nabla^2 u = f$ em uma região Ω do plano deve satisfazer a seguinte condição de compatibilidade:

$$\oint_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} \, dr = \iint_{\Omega} f \, dA.$$

Resolver problemas de valor de contorno analiticamente nem sempre é possível, especialmente quando se lida com domínios complexos ou condições de contorno irregulares. Nessas situações, métodos numéricos, como o método das diferenças finitas, tornam-se ferramentas essenciais. Este texto foca na aplicação do método das diferenças finitas para resolver problemas de valor de contorno unidimensional de segunda ordem sujeito a condições Dirichlet, e o problema de valor de contorno bidimensional de segunda ordem sujeito à condição de Dirichlet.

1.3 MÉTODO DAS DIFERENÇAS FINITAS

O método das diferenças finitas é amplamente utilizado devido à sua simplicidade e facilidade de implementação. Ele transforma equações diferenciais em sistemas de equações não lineares ou lineares, dependendo da equação diferencial, que podem ser resolvidos usando métodos numéricos. No entanto, a precisão das soluções depende da espessura da malha: grades mais finas fornecem soluções mais precisas, mas aumentam o custo computacional.

Além disso, a aplicação do método em domínios irregulares ou com condições de contorno complexas pode exigir adaptações específicas, como a utilização de malhas não uniformes ou métodos de interpolação. Mesmo assim, o método das diferenças finitas continua a ser uma técnica poderosa e versátil para a solução numérica de problemas de valor de contorno em diversos campos da ciência e engenharia.

O método das diferenças finitas consiste em escolher pontos no domínio da solução procurada e substituir a derivada que aparece na equação por uma diferença finita. A seguir apresentaremos exemplos, sem entrar em detalhes do método de diferenças finitas. Esses detalhes serão apresentados no respectivo capítulo.

Por exemplo, para o problema de valor de contorno unidimensional descrito pela equação

diferencial ordinária de segunda ordem:

$$\frac{d^2u}{dx^2} = f(x),$$

com condições de contorno do tipo Dirichlet:

$$u(a) = \alpha \text{ e } u(b) = \beta,$$

onde α e β são valores reais conhecidos.

Para utilizar o método de diferenças finitas, é necessário estabelecer uma malha computacional que consiste em um conjunto de pontos onde será calculada a solução aproximada da nossa função. Aqui, vamos definir uma malha uniforme, ou seja, uma malha onde todos os pontos estão distribuídos uniformemente.

Seja $[a, b]$ o domínio do nosso problema, particionamos o domínio $[a, b]$ em n partes iguais, desta forma, o espaçamento entre os pontos da malha será $h = \frac{b-a}{n}$, e os pontos da malha serão dados por

$$x_i = a + ih, \text{ para } i = 0, 1, 2, \dots, n.$$

Uma vez definida a malha computacional do nosso problema, devemos substituir a derivada segunda que aparece na equação por uma diferença finita. Por exemplo, por uma fórmula de diferenças finitas centradas de segunda ordem em x_i , então, a expressão discreta da equação contínua é dada por:

$$\left. \frac{d^2u}{dx^2} \right|_{x_i} \approx \frac{u(x_{i-1}) - 2u(x_i) + u(x_{i+1}))}{h^2}.$$

Aplicando a fórmula nos pontos internos da malha x_i para $i = 1, 2, \dots, n-1$, considerando U_i como a aproximação do termo $u(x_i)$, obtemos um sistemas de equações linear da forma

$$\frac{U_{i-1} - 2U_i + U_{i+1}}{h^2} = f_i$$

cuja solução nos dá o valor aproximado da função $u(x)$ avaliado nos pontos x_i .

Para o problema de Poisson bidimensional em um domínio retangular:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y)$$

com $\Omega = [a, b] \times [c, d]$, sujeito a condição de Dirichlet $u|_{\partial\Omega} = g(x, y)$, o método das diferenças finitas seguem o mesmo passo.

Primeiro escolhemos a malha computacional discretizando o domínio do nosso problema, usando os espaçamentos $h = \frac{b-a}{m}$ e $k = \frac{d-c}{n}$, obtemos um conjunto de pontos (x_i, y_j) da forma:

$$x_i = a + ih \text{ para } i = 0, 1, 2, \dots, m,$$

$$y_j = c + jk \text{ para } j = 0, 1, 2, \dots, n.$$

Em seguida avaliamos as derivadas parciais nos pontos internos da malha (x_i, y_j) para $i = 1, 2, \dots, m - 1$ e $j = 1, 2, \dots, m - 1$. Considerando $U_{i,j} \approx u(x_i, y_j)$, obtemos um sistema linear da forma

$$\frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{h^2} + \frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{k^2} = f_{i,j},$$

cujas soluções nos dá o valor da aproximada função $u(x_i, y_j)$ nos pontos (x_i, y_j) da malha.

O objetivo principal deste trabalho é apresentar aspectos matemáticos do uso do método das diferenças finitas na solução de problemas de valor de contorno em uma e duas dimensões, sujeitos a condições de contorno do tipo Dirichlet.

Em particular, aborda-se a convergência de segunda ordem, quando são utilizadas diferenças finitas centradas de segunda ordem, bem como a análise da convergência das soluções dos sistemas lineares pelos métodos clássicos aplicados à equação de Poisson, associada ao método das diferenças finitas.

Neste capítulo, apresentou-se uma breve revisão dos conceitos de equações diferenciais ordinárias e parciais, com foco nos problemas de valor de contorno de segunda ordem, além de uma introdução ao método das diferenças finitas.

No Capítulo 2, discute-se a aplicação do método das diferenças finitas a um problema de valor de contorno unidimensional. Neste capítulo, são introduzidos os conceitos de consistência, estabilidade e convergência, incluindo a demonstração de que o método converge.

O Capítulo 3 aplica o método das diferenças finitas de segunda ordem à equação de Poisson. São discutidas algumas possíveis ordenações dos pontos, a estrutura do sistema linear associado e as propriedades que garantem a convergência de segunda ordem do método.

No Capítulo 4, são apresentados os métodos iterativos clássicos para a solução do sistema linear associado à equação de Poisson, como os métodos de Jacobi, Gauss-Seidel e SOR. Além disso, são exploradas as propriedades de convergência desses métodos na solução numérica.

O Capítulo 5 aborda conceitos necessários para a análise de convergência do método das diferenças finitas. Exemplos de resolução de problemas de valor de contorno unidimensionais e bidimensionais são apresentados, confirmando que o método das diferenças finitas centradas de segunda ordem apresenta a precisão teórica relatada na literatura.

Por fim, no Capítulo 6, são apresentadas as conclusões do trabalho e sugestões para pesquisas futuras.

2. PROBLEMA UNIDIMENSIONAL

Apresentaremos aqui um problema de valor de contorno unidimensional associado à uma equação diferencial ordinária de segunda ordem. Este problema servirá de base para o problema de valor de contorno bidimensional que trabalharemos no capítulo 3.

Consideremos o seguinte problema de valor de contorno:

$$\begin{cases} u''(x) = f(x), & \text{para } x \in (a, b), \\ u(a) = \alpha \quad \text{e} \quad u(b) = \beta. \end{cases}$$

onde $u(x)$ é uma função desconhecida definida em um domínio $[a, b]$, e α e β são valores dados.

Veremos nesse capítulo a aplicação do método de diferenças finitas para o problema unidimensional proposto, faremos uma análise do erro do esquema de discretização obtido pelo método de diferenças finitas, e estudaremos os conceitos de consistência, estabilidade e convergência do método de diferenças finitas.

Para o desenvolvimento desse capítulo nos basearemos principalmente nas referências [3, 9, 14, 15].

2.1 DIFERENÇAS FINITAS

O método das diferenças finitas tem como parte central a aproximação das derivadas por um esquema finito de diferenças. Ao longo desse texto, vamos considerar u como sendo uma função suave, isto é, uma função que seja suficientemente diferenciável no domínio Ω do nosso problema.

Para obter os esquemas de diferenças finitas que usaremos ao longo do texto, vamos inicialmente considerar a derivada da função u :

$$u'(x) = \lim_{h \rightarrow 0} \frac{u(x+h) - u(x)}{h}.$$

Suponhamos que conheçamos o valor de $u(x)$ nos pontos do domínio $[a, b]$, e queiramos calcular uma aproximação para $u'(x)$. A primeira escolha a se pensar para calcular tal aproximação poderia ser:

$$u'(x) \approx \frac{u(x+h) - u(x)}{h}, \quad (2.1)$$

para um h pequeno. Assim, quanto mais próximo de 0 for o valor de h , espera-se, pela definição

de limite, que a aproximação da derivada seja melhor.

Na equação 2.1, usamos dois pontos que estão separados por uma distância h . Podemos, de maneira análoga, considerar a seguinte aproximação para a derivada.

$$u'(x) \approx \frac{u(x) - u(x - h)}{h}, \quad (2.2)$$

que considera um ponto $x - h$, que se situa anteriormente à x com uma distância h .

Uma outra possibilidade para calcular os pontos seria realizar uma média entre as duas aproximações, ou seja,

$$\frac{1}{2} \left(\frac{u(x + h) - u(x)}{h} + \frac{u(x) - u(x - h)}{h} \right),$$

obtendo assim, a aproximação

$$u'(x) \approx \frac{u(x + h) - u(x - h)}{2h}, \quad (2.3)$$

que possui utiliza os pontos em $x - h$ e $x + h$, com uma distância $2h$.

A partir de 2.1, 2.2 e 2.3, podemos definir os seguintes esquemas de diferenças finitas para aproximar a primeira derivada $u'(x)$.

Definição 2.1.1. Chama-se *diferença finita sucessiva* para a primeira derivada $u'(x)$ o esquema $\delta_h^+ u(x)$ definido por

$$\delta_h^+ u(x) = \frac{u(x + h) - u(x)}{h}.$$

Definição 2.1.2. Chama-se *diferença finita regressiva* para a primeira derivada $u'(x)$ o esquema $\delta_h^- u(x)$ definido por

$$\delta_h^- u(x) = \frac{u(x) - u(x - h)}{h}.$$

Definição 2.1.3. Chama-se *diferença finita centrada* para a primeira derivada $u'(x)$ o esquema $\delta_h^* u(x)$ definido por

$$\delta_h^* u(x) = \frac{u(x + h) - u(x - h)}{2h}.$$

Onde δ_h^+ , δ_h^- , δ_h^* são chamados operadores de diferença finita.

Para aproximarmos a derivada segunda $u''(x)$, podemos considerar $u''(x) = (u'(x))'$, da forma

$$u''(x) = \lim_{h \rightarrow 0} \frac{u'(x + h) - u'(x)}{h},$$

fazendo de maneira análoga à deriva primeira, podemos aproximar usando a diferença sucessiva,

$$u''(x) \approx \frac{u'(x + h) - u'(x)}{h},$$

considerando a aproximação regressiva para $u'(x + h)$ e $u'(x)$, temos

$$u''(x) \approx \frac{1}{h} \left(\frac{u(x + h) - u(x)}{h} - \frac{u(x) - u(x - h)}{h} \right),$$

portanto, obtemos a seguinte aproximação para a derivada segunda $u''(x)$:

$$u''(x) \approx \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}. \quad (2.4)$$

Assim, podemos definir, a partir de 2.4, o operador de diferenças finitas centradas para a aproximação da derivada segunda $u''(x)$.

Definição 2.1.4. Chama-se *diferença finita centrada* para a segunda derivada $u''(x)$ o esquema $\delta_h^2 u(x)$ definido por

$$\delta_h^2 u(x) = \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}.$$

Esse método para obter as diferenças finitas não nos diz, porém, o quão boa é a nossa aproximação. Para podermos analisar o erro de nossa aproximação, faremos uso do Teorema de Taylor, que nos permitirá exibir a ordem do erro da aproximação.

Teorema 2.1.5 (Teorema de Taylor). *Seja $f : [a, b] \rightarrow \mathbb{R}$ uma função de classe C^k . Se $f^{(k)}$ é diferenciável no intervalo aberto (a, b) , então para todo $x \in [a, b]$, existe um número $\xi = \xi(x)$ no intervalo (a, b) tal que:*

$$f(x) = f(a) + (x-a)f'(x) + \cdots + \frac{(x-a)^k}{k!} f^{(k)}(a) + \frac{(x-a)^{k+1}}{(k+1)!} f^{(k+1)}(\xi).$$

Para descrever a ordem do erro do método das diferenças finitas, introduziremos o conceito de O-grande, conforme [9].

Definição 2.1.6. Se $f(h)$ e $g(h)$ são duas funções de h , então dizemos que

$$f(h) = O(g(h)), \text{ quando } h \rightarrow 0,$$

se existem duas constantes $C > 0$ e h_0 , tal que

$$\left| \frac{f(h)}{g(h)} \right| < C, \text{ para todo } h < h_0.$$

De forma equivalente, podemos dizer que $f(h) = O(g(h))$, se $|f(h)| < C|g(h)|$ para todo h suficientemente pequeno. Assim, se tivermos uma função $f(h)$, tal que, existe uma constante $C > 0$, de forma que

$$|f(h)| < Ch^p,$$

para $h > 0$, então dizemos que a $f(h)$ é de ordem p , e indicamos por $f(h) = O(h^p)$.

Para podermos obter o erro da diferença sucessiva usando o Teorema de Taylor, vamos expandir o $u(x+h)$ em torno do ponto x , assim,

$$u(x+h) = u(x) + hu'(x) + \frac{h^2}{2}u''(\xi).$$

subtraindo o termo $u(x)$ de ambos os lados, e dividindo por h obtemos:

$$\frac{u(x+h) - u(x)}{h} = u'(x) + \frac{u''(\xi)}{2}h,$$

dessa forma concluímos que

$$\delta_h^+ u(x) - u'(x) = O(h), \quad (2.5)$$

e portanto o erro da diferença sucessiva δ_h^+ é de primeira ordem.

Analogamente podemos obter o método de diferença regressiva expandindo $u(x-h)$ em torno do ponto x :

$$u(x-h) = u(x) - hu'(x) + \frac{h^2}{2}u''(\xi).$$

Subtraindo o termo $u(x)$ de ambos os lados, e dividindo por h obtemos:

$$\frac{u(x) - u(x-h)}{h} = u'(x) - \frac{u''(\xi)}{2}h$$

dessa forma concluímos que

$$\delta_h^- u(x) - u'(x) = O(h), \quad (2.6)$$

e portanto o erro da diferença regressiva δ_h^- é também de primeira ordem.

Para obtermos a diferença centrada devemos usar a expansão de Taylor dos termos $u(x+h)$ e $u(x-h)$ em torno do ponto x . Para $u(x+h)$, temos que

$$u(x+h) = u(x) + hu'(x) + \frac{h^2}{2}u''(x) + \frac{h^3}{3!}u'''(\xi_1),$$

para $\xi_1 \in (x, x+h)$. Analogamente, para $u(x-h)$, temos que

$$u(x-h) = u(x) - hu'(x) + \frac{h^2}{2}u''(x) - \frac{h^3}{3!}u'''(\xi_2).$$

para $\xi_2 \in (x-h, x)$. Subtraindo o termo $u(x-h)$ de $u(x+h)$, obtemos

$$u(x+h) - u(x-h) = 2hu'(x) + \frac{h^3}{3!}[u'''(\xi_1) + u'''(\xi_2)].$$

Pelo teorema do valor intermediário, existe um ξ tal que

$$u'''(\xi) = \frac{u'''(\xi_1) + u'''(\xi_2)}{2},$$

logo, temos que

$$u(x+h) - u(x-h) = 2hu'(x) + 2\frac{h^3}{3!}u'''(\xi),$$

para $\xi \in (x-h, x+h)$. Dividindo ambos os membros da equação por $2h$, obtemos

$$\frac{u(x+h) - u(x-h)}{2h} = u'(x) + \frac{u'''(\xi)}{3!}h^2,$$

e portanto

$$\delta_h^* - u'(x) = O(h^2). \quad (2.7)$$

Assim, o erro de aproximação da diferença finita centrada δ_h^* é de segunda ordem.

Para obtermos a diferença centrada de segunda ordem para a derivada $u''(x)$, utilizamos o Teorema de Taylor para expandir os termos $u(x+h)$ e $u(x-h)$ em torno do ponto x . Para $u(x+h)$, temos que

$$u(x+h) = u(x) + hu'(x) + \frac{h^2}{2}u''(x) + \frac{h^3}{3!}u'''(x) + \frac{h^4}{4!}u^{(4)}(\xi),$$

para $\xi_1 \in (x, x+h)$. Analogamente, para $u(x-h)$, temos que

$$u(x-h) = u(x) - hu'(x) + \frac{h^2}{2}u''(x) - \frac{h^3}{3!}u'''(x) + \frac{h^4}{4!}u^{(4)}(\xi).$$

para $\xi_2 \in (x-h, x)$, Somando os termos $u(x-h)$ e $u(x+h)$, obtemos

$$u(x-h) + u(x+h) = 2u(x) + \frac{2h^2}{2}u''(x) + \frac{h^4}{4!}[u^{(4)}(\xi_1) + u^{(4)}(\xi_2)],$$

logo, pelo teorema do valor intermediário, existe um ξ tal que

$$u'''(\xi) = \frac{u'''(\xi_1) + u'''(\xi_2)}{2},$$

e portanto,

$$u(x-h) + u(x+h) = 2u(x) + \frac{2h^2}{2}u''(x) + \frac{2h^4}{4!}u^{(4)}(\xi).$$

Dividindo ambos os lados da equação por h^2 , obtemos

$$\frac{u(x-h) - 2u(x) + u(x+h)}{h^2} = u''(x) + \frac{u^{(4)}(\xi)}{12}h^2,$$

concluimos então que:

$$\delta_h^2 u(x) - u''(x) = O(h^2), \quad (2.8)$$

portanto, o erro da diferença finita centrada δ_h^2 é de segunda ordem.

Assim, a partir de 2.5, 2.6, 2.7 e 2.8, obtivemos os seguintes esquemas de diferenças finitas:

$$\begin{aligned} u'(x) &= \frac{u(x+h) - u(x)}{h} + O(h), \\ u'(x) &= \frac{u(x) - u(x-h)}{h} + O(h), \\ u'(x) &= \frac{u(x+h) - u(x-h)}{2h} + O(h^2), \\ u''(x) &= \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} + O(h^2). \end{aligned}$$

Vemos que as diferenças finitas sucessivas e regressivas são de primeira ordem, enquanto

que ambas as diferenças centradas são de segunda ordem.

Neste texto focaremos no uso das diferenças finitas centradas, que possuem maior ordem de precisão. Elas serão usadas para aproximar a equação diferencial do nosso problema, resultando, desta forma, em um sistema de equações lineares. A solução do sistema fornecerá uma aproximação para a solução da equação diferencial.

2.2 APLICAÇÃO DO MÉTODO

Para aplicarmos o método de diferenças finitas, devemos, inicialmente, definir uma malha computacional, que consiste em uma discretização do domínio $[a, b]$ do problema de valor de contorno. Para estudarmos a aplicação do método, vamos usar o nosso problema de exemplo proposto no início desse capítulo:

$$\begin{cases} u''(x) = f(x), & \text{para } x \in (a, b), \\ u(a) = \alpha, & \text{e } u(b) = \beta. \end{cases}$$

onde $u(x)$ é uma função desconhecida definida em um domínio $[a, b]$, e α e β são valores dados.

Para definirmos a malha que discretizará o domínio, iremos escolher um conjunto de pontos $x_0, x_1, x_2, \dots, x_n$, de forma que:

$$a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b.$$

Podemos escolher esse conjunto de pontos de várias maneiras possíveis. A maneira mais simples é escolhermos os pontos de forma que eles estejam distribuídos uniformemente ao longo do domínio do problema.

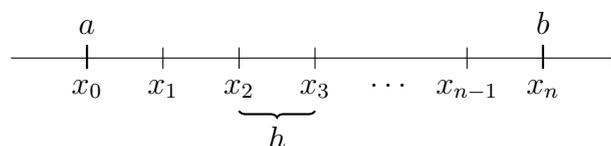
Particionamos o domínio $[a, b]$ do problema em n partes iguais, dessa forma, obtemos um conjunto de $n+1$ pontos, onde a distância entre dois pontos consecutivos será dado por $h = \frac{b-a}{n}$, e portanto, o primeiro ponto da malha será $x_0 = a$, seguido por $x_1 = a + h$, $x_2 = a + 2h$, e assim sucessivamente, até o ponto $x_n = b$.

Logo, teremos uma malha de pontos homogênea conforme a Figura 2.1, onde os pontos da malha são definidos como:

$$x_i = a + ih, \text{ para } i = 0, 1, 2, \dots, n,$$

com espaçamento uniforme $h = \frac{b-a}{n}$.

Figura 2.1: Malha uniformes para o domínio $[a, b]$.



Fonte: O autor.

A solução desse sistema linear nos dará uma aproximação \mathbf{U} , para a função u nos pontos x_i da malha. Para a resolução desse sistema, usaremos métodos iterativos estacionário, conforme veremos no Capítulo 4.

2.3 ERRO LOCAL E GLOBAL

O erro de truncamento local se trata da diferença entre a equação diferencial original e a aproximação obtida pelo método de diferença finita. O erro de truncamento local é calculado para cada ponto x_j do domínio da equação [9, 15].

Seja L um operador diferencial em u associado com uma equação diferencial, por exemplo, dado a equação $au''(x) + bu'(x) + cu(x) = f(x)$, temos que o operador diferencial será:

$$L = aD^2 + bD + c,$$

onde D é o operador derivada, assim, temos que $Lu = f$. Consideremos L_h como sendo um operador de diferença finita correspondente à L , por exemplo,

$$L_h = a\delta_h^2 + b\delta_h^* + c.$$

A partir disso, podemos definir a noção de erro de truncamento local em um ponto x_i , que mede a diferença entre a equação diferencial original e o esquema de diferenças finitas usado para aproximar a equação.

Definição 2.3.1. Dados um operador diferencial L e um operador de diferenças finitas L_h . Chama-se *erro de truncamento local* no ponto x_i , a diferença:

$$\tau_h u(x_i) = L_h u(x_i) - Lu(x_i),$$

Como o erro de truncamento local é definido pontualmente, em cada ponto x_i , podemos ter um erro de truncamento local $\tau_h u(x_i)$ diferente em cada ponto, dessa forma podemos definir o vetor $\boldsymbol{\tau}_h$, como sendo

$$\boldsymbol{\tau}_h = [\tau_h u(x_1), \tau_h u(x_2), \dots, \tau_h u(x_n)]^t.$$

Exemplo 2.3.2. Consideremos o seguinte problema de valor de contorno com condições:

$$\begin{cases} u''(x) = f(x), & \text{para } x \in (a, b), \\ u(a) = \alpha \quad \text{e} \quad u(b) = \beta. \end{cases}$$

Podemos aproximar a derivada segunda pela diferença finita centrada de segunda ordem $\delta^2 U_i = \frac{U_{i+1} - 2U_i + U_{i-1}}{h^2}$. Assim, temos que o erro de truncamento local no ponto x_i é dado pela diferença entre a aproximação e a derivada no ponto x_i .

$$\tau_h u(x_i) = \delta_h^2 u(x_i) - D^2 u(x_i).$$

Conforme obtido na seção sobre as diferenças finitas, o termo $\delta_h^2 u(x_i)$, satisfaz

$$\frac{u(x_{i-1}) - 2u(x_i) + u(x_{i+1}))}{h^2} = u''(x_i) + \frac{h^2}{12} u^{(4)}(\xi_i),$$

para $\xi_i \in (x_i - h, x_i + h)$. Logo o erro de truncamento local no ponto x_i é

$$\tau_h u(x_i) = \frac{h^2}{12} u^{(4)}(\xi_i).$$

para todo ponto interno x_i da malha, e portanto, $\boldsymbol{\tau}_h = O(h^2)$.

Para podermos analisar o erro entre a solução exata do problema, e a solução aproximada obtida a partir do sistema linear, precisaremos definir o conceito de erro global, que consiste na diferença entre a solução exata \mathbf{u} e a solução aproximada \mathbf{U} , onde:

$$\mathbf{u} = [u(x_1), u(x_2), \dots, u(x_{n-1})]^t, \quad \text{e} \quad \mathbf{U} = [U_1, U_2, \dots, U_{n-1}]^t$$

Definição 2.3.3. Seja \mathbf{u} o vetor composto por $u(x_i)$, e \mathbf{U} o vetor contendo as aproximações U_i , então chama-se *erro global*, a diferença:

$$\mathbf{e}_h = \mathbf{U} - \mathbf{u}.$$

Seja A_h a matriz dos coeficientes obtida a partir do método de diferenças finitas usando um espaçamento h , \mathbf{f}^* o vetor obtido satisfazendo a discretização $A_h \mathbf{U} = \mathbf{f}^*$, que contem os termos $f(x_i)$ e os termos obtidos a partir das condições de contorno.

Como a matriz A_h contem os termos obtidos a partir da discretização L_h , a menos dos valores da borda, e $L\mathbf{u} = \mathbf{f}$, então $L_h \mathbf{u} - L\mathbf{u} = A_h \mathbf{u} - \mathbf{f}^*$. e portanto $A\mathbf{u} = \mathbf{f}^* + \boldsymbol{\tau}_h$, logo,

$$A\mathbf{e}_h = A(\mathbf{U} - \mathbf{u}) = A\mathbf{U} - A\mathbf{u} = \mathbf{f}^* - (\mathbf{f}^* + \boldsymbol{\tau}_h) = -\boldsymbol{\tau}_h.$$

Portanto, o erro global \mathbf{e} e o erro local $\boldsymbol{\tau}$, se relacionam por

$$\mathbf{e}_h = -A^{-1} \boldsymbol{\tau}_h.$$

Logo, para uma norma $\|\cdot\|$, o erro global satisfaz a seguinte relação:

$$\|\mathbf{e}_h\| = \|A_h^{-1} \boldsymbol{\tau}_h\| \leq \|A_h^{-1}\| \|\boldsymbol{\tau}_h\|. \tag{2.10}$$

2.4 CONSISTÊNCIA

Um método de diferenças finitas é dito ser consistente com a equação diferencial, se a discretização do problema aproxima a equação original do problema quando o espaçamento h do problema tende à zero, ou seja,

$$\lim_{h \rightarrow 0} L_h \mathbf{u} = L\mathbf{u}.$$

Definição 2.4.1. Um método de diferenças finitas é consistente se:

$$\|\boldsymbol{\tau}_h\| \rightarrow 0, \text{ quando } h \rightarrow 0.$$

No geral se $L_h u(x_i) - Lu(x_i) = O(h^p)$ para todos os pontos x_i , então $\|\boldsymbol{\tau}_h\| \leq Ch^p$ para $p > 0$ e $C > 0$ constante. Dessa forma, se $h \rightarrow 0$, então como $\|\boldsymbol{\tau}_h\|$ está limitada por Ch^p , então $\|\boldsymbol{\tau}_h\| \rightarrow 0$, e portanto o método é consistente.

No exemplo 2.3.2 vimos que a discretização do problema $u''(x) = f(x)$ usando diferença finita centrada de segunda ordem satisfaz:

$$\delta_h^2 u(x_i) - u''(x_i) = O(h^2),$$

e portanto, o erro de truncamento local $\tau_h u(x_i) = O(h^2)$, logo $\|\boldsymbol{\tau}_h\| \rightarrow 0$ quando $h \rightarrow 0$ e portanto o problema é consistente.

2.5 ESTABILIDADE

Seja A_h a matriz dos coeficientes do método de diferenças finitas obtido a partir da discretização do problema usando um intervalo h . Conforme foi visto anteriormente, como $A_h \mathbf{e}_h = -\boldsymbol{\tau}_h$ então temos que a norma erro global do satisfaz

$$\|\mathbf{e}_h\| = \|A_h^{-1} \boldsymbol{\tau}_h\| \leq \|A_h^{-1}\| \|\boldsymbol{\tau}_h\|,$$

o que nos dá uma cota superior para o erro geral.

Se o método for consistente então quando $h \rightarrow 0$ teremos que $\boldsymbol{\tau}_h \rightarrow 0$, para que o erro global seja limitado devemos garantir que exista uma constante $C > 0$ tal que $\|A_h^{-1}\| \leq C$, dessa forma:

$$\|\mathbf{e}_h\| \leq C \|\boldsymbol{\tau}_h\|.$$

A partir disso fazemos a seguinte definição.

Definição 2.5.1. Um método de diferenças finitas é estável se A_h é invertível e

$$\|A_h^{-1}\| \leq C, \text{ para todo } h < h_0,$$

onde C e h_0 são constantes independentes de $h > 0$.

Dado uma matriz A simétrica, a norma-2 da matriz A é dada pelo seu raio espectral:

$$\|A\|_2 = \rho(A) = \max_j |\lambda_j|,$$

onde λ_i são os autovalores da matriz A . Nesse caso, temos que:

$$\|A^{-1}\|_2 = \rho(A^{-1}) = \max_j |(\lambda_j)^{-1}| = \frac{1}{\min_j |\lambda_j|}. \quad (2.11)$$

Assim, para mostrar a estabilidade, basta mostrar que o termo na equação 2.11 é limitada por uma constante.

No problema exemplo anterior, obtivemos a matriz A_h de dimensão $n - 1$ por $n - 1$ do nosso problema exemplo, de acordo com [9], a matriz do problema possui a estrutura de uma matriz Toeplitz. Pode-se verificar que para $h = \frac{1}{n}$, os autovalores da matriz A_h são dados pela expressões:

$$\lambda_j = \frac{2}{h^2} [\cos(\pi j h) - 1], \text{ para } j = 1, 2, \dots, n - 1.$$

Enquanto que os autovetores \mathbf{u}_j associados aos autovalores λ_j possuem uma estrutura da forma

$$\mathbf{u}_j = \begin{bmatrix} \text{sen}(\pi j h) \\ \text{sen}(2\pi j h) \\ \text{sen}(3\pi j h) \\ \vdots \\ \text{sen}((n - 1)\pi j h) \end{bmatrix}.$$

Como $0 < j < n$ e $h = \frac{1}{n}$, então $0 < \pi j h < \pi$. Dessa forma, podemos concluir que para $h = \frac{1}{n}$, teremos a seguinte ordenação de termos:

$$1 > \cos(\pi h) > \cos(\pi 2h) > \dots > \cos(\pi j h) > \dots > \cos(\pi(n - 1)h) > 0.$$

Subtraindo 1 de todos os termos da sequência, e multiplicando por $\frac{2}{h^2}$, concluímos que

$$0 > \lambda_1 > \lambda_2 > \dots > \lambda_{n-1} > -1,$$

e portanto,

$$0 < |\lambda_1| < |\lambda_2| < \dots < |\lambda_{n-1}| < 1.$$

Assim, o menor autovalor em módulo é $\lambda_1 = \frac{2}{h^2}(\cos(\pi h) - 1)$.

Fazendo a expansão em série de Taylor do termo $\cos(\pi h)$:

$$\cos(\pi h) = 1 - \frac{\pi^2}{2}h^2 + \frac{\pi^4}{4!}h^4 - \frac{\pi^6}{6!}h^6 + \dots,$$

temos que:

$$\lambda_1 = \frac{2}{h^2} \left(1 - \frac{\pi^2}{2}h^2 + O(h^4) - 1 \right) = -\pi^2 + O(h^2),$$

e portanto, ao aplicar a equação 2.11, obtemos que

$$\|A_h^{-1}\|_2 = \frac{1}{\min_j |\lambda_j|} = \frac{1}{|\lambda_1|} \rightarrow \frac{1}{\pi^2}, \text{ quando } h \rightarrow 0.$$

Pela definição de limite, isto implica que, para todo $\epsilon > 0$, existe um h_0 tal que

$$\left| \|A_h^{-1}\|_2 - \frac{1}{\pi^2} \right| < \epsilon \text{ para todo } h < h_0,$$

como $\frac{1}{h^2} > \frac{1}{\pi^2}$, então

$$\|A_h^{-1}\|_2 < \frac{1}{\pi^2} + \epsilon.$$

Assim, considerando $C > \frac{1}{\pi^2} + \epsilon$, concluímos que

$$\|A_h^{-1}\|_2 < C \text{ para todo } h < h_0,$$

e portanto o método de diferenças finitas é estável segundo a norma $\|\cdot\|_2$ de matriz. Como as normas $\|\cdot\|_1$, $\|\cdot\|_2$ e $\|\cdot\|_\infty$ de matriz são equivalentes, então existem constantes p e q tal que

$$\|A_h^{-1}\|_1 \leq p \|A_h^{-1}\|_2 \leq pC, \quad \text{e} \quad \|A_h^{-1}\|_\infty \leq q \|A_h^{-1}\|_2 \leq qC.$$

Dessa forma, A_h também é estável segundo as normas $\|\cdot\|_1$ e $\|\cdot\|_\infty$ de matriz.

2.6 CONVERGÊNCIA

Um método de diferenças finitas é dito ser convergente quando a solução aproximada gerada pelo método de diferenças se aproxima da solução real do problema, ou seja o erro geral da solução deve diminuir quanto menor for o tamanho do espaçamento escolhido.

Definição 2.6.1. Um método de diferenças finitas é convergente se:

$$\lim_{h \rightarrow 0} \|\mathbf{e}_h\| = 0.$$

Na prática é muito difícil mostrar que um problema é convergente apenas usando a definição 2.6.1, para isso será usado um teorema que relaciona convergência com consistência e estabilidade.

Teorema 2.6.2 (Teorema da Equivalência de Lax). *Um esquema de diferenças finita é convergente se, e somente se, ele é consistente e estável.*

Esse teorema nos permite demonstrar a convergência de um esquema de diferenças finitas, através da consistência e da estabilidade.

Consideremos o problema exemplo, foi mostrado que o problema é consistente e estável, logo pelo teorema ele também é convergente.

De fato, vimos que o erro de truncamento local do problema é dado por $\tau_h = O(h^2)$ e que $\|A_h\|_2 < \frac{1}{\pi^2}$. Portanto para o erro global do problema temos que:

$$\|\mathbf{e}_h\|_2 = \|A_h^{-1} \tau_h\|_2 \leq \|A_h^{-1}\|_2 \|\tau_h\|_2 \approx \frac{1}{\pi^2} \|\tau_h\|_2 = O(h^2).$$

Logo, $\mathbf{e}_h \rightarrow 0$ quando $h \rightarrow 0$.

3. PROBLEMA BIDIMENSIONAL

No presente capítulo apresentaremos a aplicação do método das diferenças finitas centradas de segunda ordem à equação de Poisson, em coordenadas cartesianas definidas em uma região retangular, sujeita a condições de contorno do tipo Dirichlet.

Consideremos o problema de valor de contorno de tipo Dirichlet definido por:

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y) & , \text{ para } (x, y) \in \Omega, \\ u(x, y) = g(x, y) & , \text{ para } (x, y) \in \partial\Omega. \end{cases} \quad (3.1)$$

onde $f(x, y)$ e $g(x, y)$ são funções dadas, e $\Omega \in \mathbb{R}^2$ é o domínio do problema.

Nesse texto focaremos no caso onde está definido em um domínio retangular $\Omega = [a, b] \times [c, d]$

Mostraremos que o uso do método de diferenças finitas centrada de segunda ordem nos dá uma aproximação de segunda ordem para a solução do problema. Usaremos os conceitos de erro de truncamento local, erro global e o princípio do máximo para analisar o erro de aproximação do problema.

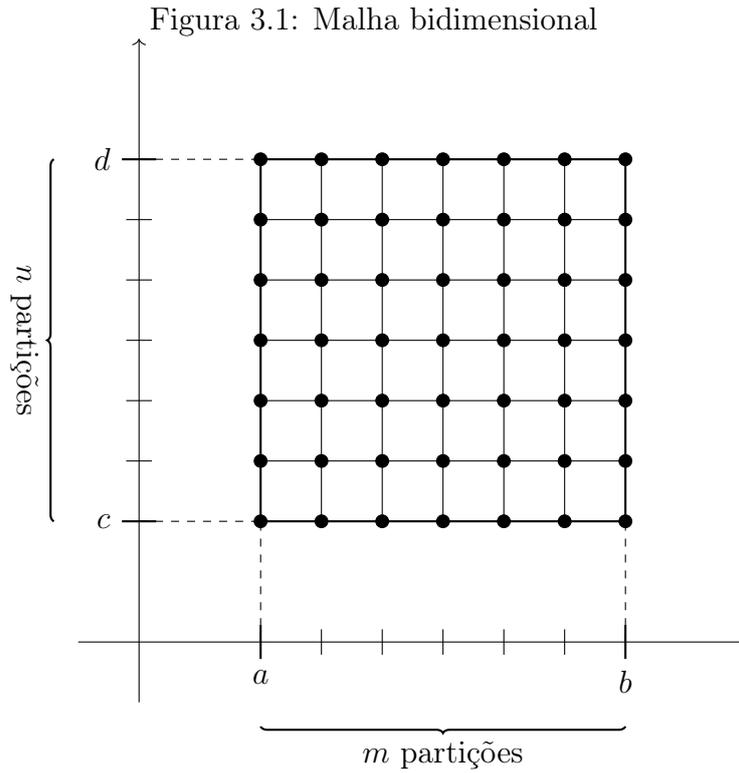
As referências bibliográficas usadas para o desenvolvimento este capítulo são [6], [7], [9], [10], [13] e [15].

3.1 APLICAÇÃO DO MÉTODO

Para aplicarmos o método de diferenças finitas no problema de valor de contorno bidimensional, precisamos, primeiramente, definir uma malha computacional, onde calcularemos a solução aproximada do nosso problema. No nosso problema iremos considerar apenas o uso de malhas uniformes, isto é onde os pontos estão igualmente espaçados.

Para a construção da nossa malha, dividiremos os intervalos $[a, b]$ e $[c, d]$ em m e n partes, respectivamente, conforme a Figura 3.1, obtendo um conjunto de $(m + 1) \cdot (n + 1)$ pontos, incluindo pontos contidos na borda do domínio. Assim, podemos definir os pontos (x_i, y_j) da seguinte forma:

$$\begin{aligned} x_i &= a + ih, \text{ com } i = 0, 1, 2, \dots, m, \text{ e } h = \frac{b - a}{m}, \\ y_j &= c + jk, \text{ com } j = 0, 1, 2, \dots, n, \text{ e } k = \frac{d - c}{n}. \end{aligned}$$



Fonte: O autor.

Após a criação da malha computacional, podemos aplicar o método das diferenças finitas substituindo as derivadas, em cada ponto interno da malha, por um esquema de diferença finita. No presente trabalho, para aproximar as derivadas parciais de segunda ordem $\frac{\partial^2 u}{\partial x^2}$ e $\frac{\partial^2 u}{\partial y^2}$, usaremos as diferenças finitas centradas de segunda ordem, para derivadas parciais, que são da forma:

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u(x-h, y) - 2u(x, y) + u(x+h, y)}{h^2},$$

$$\frac{\partial^2 u}{\partial y^2} \approx \frac{u(x, y-k) - 2u(x, y) + u(x, y+k)}{k^2}.$$

Podemos obter a ordem do erro de maneira análoga ao Capítulo 2, usando a expansão em polinômio de Taylor em uma variável para $u(x \pm h, y)$ e $u(x, y \pm k)$, temos que

$$u(x \pm h, y) = u(x, y) \pm h \frac{\partial u}{\partial x}(x, y) + \frac{h^2}{2} \frac{\partial^2 u}{\partial x^2}(x, y) \pm \frac{h^3}{6} \frac{\partial^3 u}{\partial x^3} + \frac{h^4}{24} \frac{\partial^4 u}{\partial x^4}(\xi, \eta),$$

$$u(x, y \pm k) = u(x, y) \pm k \frac{\partial u}{\partial y}(x, y) + \frac{k^2}{2} \frac{\partial^2 u}{\partial y^2}(x, y) \pm \frac{k^3}{6} \frac{\partial^3 u}{\partial y^3} + \frac{k^4}{24} \frac{\partial^4 u}{\partial y^4}(\xi, \eta).$$

Obtemos, de maneira análoga ao capítulo 2, as seguintes expressões

$$\frac{u(x_{i-1}, y_j) - 2u(x_i, y_j) + u(x_{i+1}, y_j))}{h^2} = \frac{\partial^2 u}{\partial x^2}(x_i, y_j) + \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4}(\xi_i, \eta_j), \quad (3.2)$$

$$\frac{u(x_i, y_{j-1}) - 2u(x_i, y_j) + u(x_i, y_{j+1}))}{k^2} = \frac{\partial^2 u}{\partial y^2}(x_i, y_j) + \frac{k^2}{12} \frac{\partial^4 u}{\partial y^4}(\xi_i, \eta_j). \quad (3.3)$$

Considerando os operadores diferenciais δ_h^2 com espaçamento h , e δ_k^2 com espaçamento k , assumiremos, nesse texto, que o δ_h^2 se trata do operador aplicado na primeira variável x , enquanto que δ_k^2 se trata do operador aplicado na segunda variável y , ou seja,

$$\begin{aligned}\delta_h^2 u(x, y) &= \frac{u(x-h, y) - 2u(x, y) + u(x+h, y)}{h^2}, \\ \delta_k^2 u(x, y) &= \frac{u(x, y-k) - 2u(x, y) + u(x, y+k)}{k^2}.\end{aligned}$$

Portanto a partir de 2.1 e 2.1, temos que

$$\begin{aligned}\delta_h^2 u(x_i, y_j) - \frac{\partial^2 u}{\partial x^2}(x_i, y_j) &= \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4}(\xi_i, \eta_j), \\ \delta_k^2 u(x_i, y_j) - \frac{\partial^2 u}{\partial y^2}(x_i, y_j) &= \frac{k^2}{12} \frac{\partial^4 u}{\partial y^4}(\xi_i, \eta_j).\end{aligned}$$

Logo, as aproximações das derivadas parciais são ambas de segunda ordem. Somando ambas as equações, concluímos que

$$\delta_h^2 u(x_i, y_j) + \delta_k^2 u(x_i, y_j) = f(x_i, y_j) + O(k^2) + O(h^2).$$

Considerando $U_{i,j} \approx u(x_i, y_j)$ a solução aproximada do problema no ponto (x_i, y_j) desconsiderando os termos $O(k^2)$ e $O(h^2)$, temos então que a solução aproximada do problema é dada pela equação

$$\frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{h^2} + \frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{k^2} = f(x_i, y_j), \quad (3.4)$$

para $i = 1, 2, \dots, m-1$ e $j = 1, 2, \dots, n-1$.

Reordenando os termos da equação 3.4, podemos escrever a equação como

$$\frac{1}{k^2} U_{i,j-1} + \frac{1}{h^2} U_{i-1,j} - \left(\frac{2}{h^2} + \frac{2}{k^2} \right) U_{i,j} + \frac{1}{h^2} U_{i+1,j} + \frac{1}{k^2} U_{i,j+1} = f(x_i, y_j). \quad (3.5)$$

Considerando $f_{i,j} = f(x_i, y_j)$ e α , β e γ como

$$\alpha = - \left(\frac{2}{h^2} + \frac{2}{k^2} \right), \quad \beta = \frac{1}{h^2}, \quad \gamma = \frac{1}{k^2},$$

podemos reescrever a equação 3.5 como

$$\gamma U_{i,j-1} + \beta U_{i-1,j} - \alpha U_{i,j} + \beta U_{i+1,j} + \gamma U_{i,j+1} = f_{i,j}.$$

Exemplo 3.1.1. Vamos considerar como exemplo o caso em que temos $m = 4$ e $n = 4$ partições. Nesse caso, a malha do nosso problema será constituída de $5 \times 5 = 25$ pontos, sendo $3 \times 3 = 9$ desses pontos são pontos internos da malha, e o restante, pontos na borda da malha. Assim, para cada ponto interno da malha, teremos uma equação associada.

Assim, o sistema de equações será formado por 9 equações da forma:

$$\begin{aligned}
\gamma U_{1,0} + \beta U_{0,1} - \alpha U_{1,1} + \beta U_{2,2} + \gamma U_{1,2} &= f_{1,1}, \\
\gamma U_{2,0} + \beta U_{1,1} - \alpha U_{2,1} + \beta U_{3,2} + \gamma U_{2,2} &= f_{2,1}, \\
\gamma U_{3,0} + \beta U_{2,1} - \alpha U_{3,1} + \beta U_{4,2} + \gamma U_{3,2} &= f_{3,1}, \\
\gamma U_{1,1} + \beta U_{0,2} - \alpha U_{1,2} + \beta U_{2,3} + \gamma U_{1,3} &= f_{1,2}, \\
\gamma U_{2,1} + \beta U_{1,2} - \alpha U_{2,2} + \beta U_{3,3} + \gamma U_{2,3} &= f_{2,2}, \\
\gamma U_{3,1} + \beta U_{2,2} - \alpha U_{3,2} + \beta U_{4,3} + \gamma U_{3,3} &= f_{3,2}, \\
\gamma U_{1,2} + \beta U_{0,3} - \alpha U_{1,3} + \beta U_{2,4} + \gamma U_{1,3} &= f_{1,3}, \\
\gamma U_{2,2} + \beta U_{1,3} - \alpha U_{2,3} + \beta U_{3,4} + \gamma U_{2,3} &= f_{2,3}, \\
\gamma U_{3,2} + \beta U_{2,3} - \alpha U_{3,3} + \beta U_{4,4} + \gamma U_{3,3} &= f_{3,3}.
\end{aligned}$$

Aplicando as condições de contorno de Dirichlet ao sistema, podemos substituir as variáveis $u_{i,j}$ associadas à borda do domínios pelos valores dados pela condição de contorno do nosso problema, logo sobrarão somente 9 incógnitas, que estão associadas aos pontos internos da malha.

Isolando as constantes no lado direito do do sistema, e colocando o sistema linear na forma matricial, obtemos então o seguinte sistema:

$$\begin{bmatrix}
\alpha & \beta & & & & & & & \\
\beta & \alpha & \beta & & & & & & \\
& \beta & \alpha & & & & & & \\
\gamma & & & \alpha & \beta & & & & \\
& \gamma & & \beta & \alpha & \beta & & & \\
& & \gamma & & \beta & \alpha & & & \\
& & & \gamma & & \alpha & \beta & & \\
& & & & \gamma & \beta & \alpha & \beta & \\
& & & & & \gamma & \beta & \alpha &
\end{bmatrix}
\begin{bmatrix}
U_{11} \\
U_{21} \\
U_{31} \\
U_{12} \\
U_{22} \\
U_{32} \\
U_{13} \\
U_{23} \\
U_{33}
\end{bmatrix}
=
\begin{bmatrix}
f_{11} & -\beta u_{01} & -\gamma u_{10} \\
f_{21} & & -\gamma u_{20} \\
f_{31} & -\beta u_{41} & -\gamma u_{30} \\
f_{12} & -\beta u_{02} & \\
f_{22} & & \\
f_{32} & -\beta u_{42} & \\
f_{13} & -\beta u_{03} & -\gamma u_{14} \\
f_{23} & & -\gamma u_{24} \\
f_{33} & -\beta u_{43} & -\gamma u_{34}
\end{bmatrix},$$

em que a matriz dos coeficientes é uma matriz penta-diagonal. Além disso a matriz possui uma estrutura de blocos. Podendo ser dividida então na seguinte matriz de blocos:

$$A = \begin{bmatrix} D & B & \\ B & D & B \\ & B & D \end{bmatrix}, \text{ onde } D = \begin{bmatrix} \alpha & \beta \\ \beta & \alpha & \beta \\ & \beta & \alpha \end{bmatrix} \text{ e } B = \begin{bmatrix} \gamma & & \\ & \gamma & \\ & & \gamma \end{bmatrix}.$$

A matriz acima obtida, corresponde ao caso em que temos $m = 4$ e $n = 4$ partições. Para o caso com um número diferente de partições a matriz continuará sendo uma matriz penta diagonal com a mesma estrutura de blocos.

O tamanho dos blocos da matriz está associado ao número de pontos por linha da malha, enquanto que a quantidade linha contidos matriz está associados ao número de colunas na

malha. Dessa forma, para uma malha com $(n - 1) \times (m - 1)$ pontos internos teremos uma matriz com $(n - 1) \times (n - 1)$ blocos, tendo cada bloco dimensão $(m - 1) \times (m - 1)$, da forma:

$$A = \begin{bmatrix} D & B & & & \\ B & D & B & & \\ & B & D & B & \\ & & \ddots & \ddots & \ddots \\ & & & B & D & B \\ & & & & B & D \end{bmatrix}, \quad (3.6)$$

onde A é uma matriz com $(n - 1) \times (n - 1)$ blocos com formas:

$$D = \begin{bmatrix} \alpha & \beta & & & \\ \beta & \alpha & \beta & & \\ & \beta & \alpha & \beta & \\ & & \ddots & \ddots & \ddots \\ & & & \beta & \alpha & \beta \\ & & & & \beta & \alpha \end{bmatrix} \text{ e } B = \begin{bmatrix} \gamma & & & & \\ & \gamma & & & \\ & & \gamma & & \\ & & & \ddots & \\ & & & & \gamma \end{bmatrix}.$$

com D e B tendo dimensões $(m - 1) \times (m - 1)$

3.2 ORDENAÇÃO DOS PONTOS

Quando criamos nossa malha, obtemos um conjunto de pontos no qual iremos calcular a solução aproximada da nossa função $u = u(x, y)$. Dessa forma obtemos um sistema linear onde queremos descobrir um vetor U contendo as aproximações nos pontos.

Para a criação desse vetor U , é necessário que escolhamos uma ordenação para os pontos. No exemplo 3.1.1, escolhemos ordenar os pontos escolhendo-os pontos sequenciamento por linhas, portanto, o vetor U terá a forma:

$$\mathbf{U}^t = [U_{11}, U_{21}, U_{31}, U_{12}, U_{22}, U_{32}, U_{13}, U_{23}, U_{33}]^t.$$

A ordenação natural por linhas segue a seguinte forma

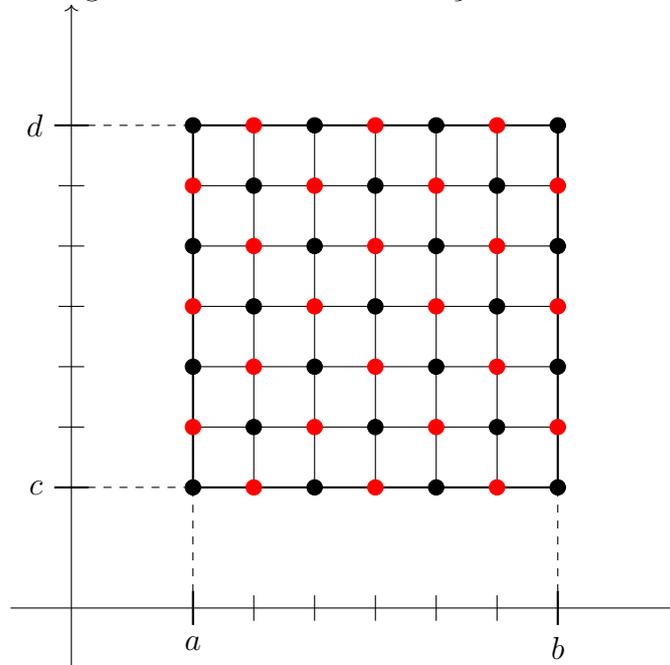
$$\mathbf{U} = \begin{bmatrix} \mathbf{U}^{(1)} \\ \mathbf{U}^{(2)} \\ \mathbf{U}^{(3)} \\ \vdots \\ \mathbf{U}^{(m-1)} \end{bmatrix}, \text{ onde } \mathbf{U}^{(j)} = \begin{bmatrix} U_{1,j} \\ U_{2,j} \\ U_{3,j} \\ \vdots \\ U_{n-1,j} \end{bmatrix}.$$

De forma semelhante, a ordenação natural por colunas possui a forma

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}^{(1)} \\ \mathbf{U}^{(2)} \\ \mathbf{U}^{(3)} \\ \vdots \\ \mathbf{U}^{(n-1)} \end{bmatrix}, \text{ onde } \mathbf{U}^{(j)} = \begin{bmatrix} U_{i,1} \\ U_{i,2} \\ U_{i,3} \\ \vdots \\ U_{i,m-1} \end{bmatrix}.$$

Outro tipo de ordenação possível de ser usado é a ordenação tipo *red-black*, que consiste em classificar os pontos da malha em dois tipos: vermelhos (*red*) e pretos (*black*). Os pontos da malha são classificados de forma alternada entre vermelhos e pretos, conforme a Figura 3.2, de forma que os pontos tipo vermelho são adjacentes apenas aos pontos tipo pretos.

Figura 3.2: Malha na ordenação *red-black*.



Fonte: O autor.

Uma vez feito a classificação dos pontos em vermelho e preto, podemos fazer a ordenação dos pontos. Fazemos a ordenação dos pontos, começando com todos os pontos tipo vermelho, e terminamos com todos os pontos tipo preto, assim o vetor na ordenação *red-black* será

$$\mathbf{U}_{rb} = \begin{bmatrix} \mathbf{U}_{red} \\ \mathbf{U}_{black} \end{bmatrix},$$

onde \mathbf{U}_{red} é o vetor contendo os U_{ij} classificados como vermelho, e \mathbf{U}_{black} é o vetor contendo os U_{ij} classificados como preto.

Como o vetor \mathbf{U}_{rb} se difere do vetor original \mathbf{U} apenas pela reordenação dos termos em relação, podemos então relacionar os dois vetores a partir de uma matriz de permutação P , ou

como P é uma matriz de permutação, então $P^{-1} = P^t$, logo $PP^t = I$, e portanto

$$PAP^tPU = Pf^*.$$

Considerando $U_{rb} = PU$ e $f_{rb}^* = Pf^*$, temos que

$$A_{rb} = PAP^t.$$

Dado que, na ordenação *red-black*, os pontos da malha podem ser bipartidos em dois conjuntos: *red* e *black*. Então, devido aos pontos *red* só se relacionarem com os pontos tipo *black*, e os pontos tipo *black* só se relacionarem aos pontos tipo *red*, a matriz dos coeficientes na forma *red-black* deverá ter a estrutura

$$A_{rb} = \begin{bmatrix} D_1 & T_1 \\ T_2 & D_2 \end{bmatrix},$$

onde D_1 e D_2 são matrizes diagonais.

3.3 ANÁLISE DO MÉTODO

Analogamente ao que foi feito no capítulo 2, podemos definir os conceitos de Erro de Truncamento Local e Erro Global. O que nos permitirá analisar a convergência do método de diferenças finitas.

Definição 3.3.1. Dado L um operador diferencial, e $L_{h,K}$ um operador discreto. Chama-se erro de truncamento local no ponto (x_i, y_j) a diferença:

$$\tau_{h,k}u(x_i, y_j) = Lu(x_i, y_j) - L_{h,k}u(x_i, y_j).$$

Definimos o vetor $\tau_{h,k}$ como sendo vetor contendo os termos $\tau_{h,k}u(x_i, y_j)$, para cada (x_i, y_j) da malha.

Definição 3.3.2. Seja \mathbf{u} o vetor composto por $u(x_i, y_j)$, e \mathbf{U} o vetor contendo as aproximações $U_{i,j}$, então chama-se *erro global*, a diferença:

$$\mathbf{e}_{h,k} = \mathbf{U} - \mathbf{u}.$$

Dado o operador discreto $L_{h,k}$ da forma

$$L_{h,k}u(x_i, y_j) = \delta_h^2u(x_i, y_j) + \delta_k^2u(x_i, y_j),$$

pela expansão em Taylor, obtemos

$$\begin{aligned}\delta_h^2 u(x_i, y_j) - \frac{\partial^2 u}{\partial x^2}(x_i, y_j) &= \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4}(x_i, y_j) + O(h^4), \\ \delta_k^2 u(x_i, y_j) - \frac{\partial^2 u}{\partial y^2}(x_i, y_j) &= \frac{k^2}{12} \frac{\partial^4 u}{\partial y^4}(x_i, y_j) + O(k^4).\end{aligned}$$

Logo, temos que

$$L_{h,k}u(x_i, y_j) - Lu(x_i, y_j) = \delta_h^2 u(x_i, y_j) + \delta_k^2 u(x_i, y_j) - \frac{\partial^2 u}{\partial x^2}(x_i, y_j) - \frac{\partial^2 u}{\partial y^2}(x_i, y_j).$$

Portanto, o erro de truncamento é dado por

$$\tau_{h,k}(x_i, y_j) = \frac{h^2}{12} \frac{\partial^2 u}{\partial x^2}(x_i, y_j) + \frac{k^2}{12} \frac{\partial^2 u}{\partial y^2}(x_i, y_j) + O(h^4) + O(k^4).$$

Dessa forma, o erro de truncamento local $\|\tau_{h,k}\| \rightarrow 0$, quando $h, k \rightarrow 0$, e portanto o método é consistente.

Similarmente ao problema unidimensional, podemos determinar os autovalores de matriz $A_{h,k}$. Segundo [9], os autovalores $\lambda_{i,j}$ da matriz $A_{h,k}$ são dados por

$$\lambda_{i,j} = \frac{2}{h^2} [\cos(\pi i h) - 1] + \frac{2}{k^2} [\cos(\pi j k) - 1],$$

para $i = 1, 2, \dots, n - 1$ e $j = 1, 2, \dots, m - 1$.

E os autovetores $\mathbf{u}_{i,j}$ associados à $\lambda_{i,j}$ são da forma

$$\mathbf{u}_{i,j} = \begin{bmatrix} \text{sen}(\pi i h) \text{sen}(\pi j k) \\ \text{sen}(2\pi i h) \text{sen}(\pi j k) \\ \vdots \\ \text{sen}((n-1)\pi i h) \text{sen}(\pi j k) \\ \text{sen}(\pi i h) \text{sen}(2\pi j k) \\ \vdots \\ \text{sen}((n-1)\pi i h) \text{sen}((m-1)\pi j k) \end{bmatrix}.$$

O menor autovalor em módulo será dado por $\lambda_{1,1}$, expandindo os termos \cos em série de Taylor, obtemos:

$$\lambda_{1,1} = \frac{2}{h^2} \left[1 - \frac{\pi^2}{2} h^2 + O(h^4) - 1 \right] + \frac{2}{k^2} \left[1 - \frac{\pi^2}{2} k^2 + O(k^4) - 1 \right],$$

juntando os termos temos que:

$$\lambda_{1,1} = \frac{2}{h^2} \left[-\frac{\pi^2}{2} h^2 + O(h^4) \right] + \frac{2}{k^2} \left[-\frac{\pi^2}{2} k^2 + O(k^4) \right],$$

portanto,

$$\lambda_{1,1} = [-\pi^2 + O(h^2)] + [-\pi^2 + O(k^2)],$$

e portanto,

$$\lambda \rightarrow -2\pi^2, \text{ para } h, k \rightarrow 0.$$

simplificando a expressão temos que o $\lambda_{1,1}$ é dado por:

$$\lambda_{1,1} = -2\pi^2 + O(h^2) + O(k^2).$$

Assim, temos que:

$$\|A_{h,k}^{-1}\|_2 = \frac{1}{\min |\lambda|} = \frac{1}{|\lambda_{1,1}|} \rightarrow \frac{1}{2\pi^2} \text{ quando } h \rightarrow 0,$$

portanto para existe uma constante $C > 0$, e $h_0, k_0 > 0$, tal que

$$\|A_{h,k}^{-1}\|_2 \leq C, \text{ para } h < h_0 \text{ e } k < k_0.$$

E portanto, o método é estável.

Como o método é consistente e estável, então pelo Teorema de Lax, ele é convergente. De fato, temos que pela relação $A_{h,k} \mathbf{e}_{h,k} = \boldsymbol{\tau}_{h,k}$ que:

$$\|\mathbf{e}_{h,k}\|_2 = \|A_{h,k}^{-1} \boldsymbol{\tau}_{h,k}\|_2 \leq \|A_{h,k}^{-1}\|_2 \|\boldsymbol{\tau}_{h,k}\|_2 \rightarrow 0, \text{ quando } h, k \rightarrow 0.$$

3.4 PRINCIPIO DO MÁXIMO

Consideremos primeiro o caso contínuo do princípio máximo, isto é, o caso em que temos uma função contínua u definida em um domínio limitado $\Omega \subset \mathbb{R}^2$. Definimos inicialmente o operador elíptico, o qual usaremos para aplicar o princípio do máximo.

Para a demonstração do princípio do máximo contínuo se utilizou das referências [7, 13, 15].

Definição 3.4.1 (Operador elíptico). Seja L um operador diferencial linear, $u = u(x, y)$ definida em uma região Ω . Então L é um operador elíptico se

$$Lu = a \frac{\partial^2 u}{\partial x^2} + 2b \frac{\partial^2 u}{\partial x \partial y} + c \frac{\partial^2 u}{\partial y^2}, \text{ onde } b^2 - ac < 0,$$

para $(x, y) \in \Omega$.

Exemplo 3.4.2. O operador laplaciano em coordenadas cartesianas

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2},$$

é um operador elíptico. Basta tomarmos um operador L com $a = 1$, $b = 0$ e $c = 1$, logo $b^2 - ac = -1 < 0$, satisfazendo a definição de operador elíptico.

Teorema 3.4.3 (Princípio do máximo). *Seja u uma função definida em um domínio limitado $\Omega \subset \mathbb{R}^2$. Se*

$$Lu \geq 0,$$

então $u(x, y)$ possui máximo na borda $\partial\Omega$ do domínio. Por outro lado, se

$$Lu \leq 0,$$

então $u(x, y)$ possui mínimo na borda $\partial\Omega$ do domínio.

Demonstração. Suponha que o teorema é falso, ou seja, que existe um ponto interior $(x^*, y^*) \in \Omega$ que é ponto máximo, isto é, $u(x^*, y^*) \geq u(x, y)$, para todo $(x, y) \in \Omega$.

Para que (x^*, y^*) seja ponto crítico, devemos ter $\nabla u(x^*, y^*) = \mathbf{0}$, ou seja,

$$\frac{\partial u}{\partial x}(x^*, y^*) = 0, \quad \frac{\partial u}{\partial y}(x^*, y^*) = 0.$$

Tomemos $x(t) = x^* + t\Delta x$ e $y(t) = y^* + t\Delta y$, consideremos então a função $g(t) = u(x(t), y(t)) = u(x^* + t\Delta x, y^* + t\Delta y)$. Aplicando a regra da cadeia:

$$\frac{dg}{dt} = \frac{\partial u}{\partial x} \frac{dx}{dt} + \frac{\partial u}{\partial y} \frac{dy}{dt} = \Delta x \frac{\partial u}{\partial x} + \Delta y \frac{\partial u}{\partial y},$$

aplicando a regra da cadeia novamente obtemos:

$$\begin{aligned} \frac{d^2g}{dt^2} &= \Delta x \frac{d}{dt} \left(\frac{\partial u}{\partial x} \right) + \Delta y \frac{d}{dt} \left(\frac{\partial u}{\partial y} \right) \\ &= \Delta x \left(\frac{\partial^2 u}{\partial x^2} \frac{dx}{dt} + \frac{\partial^2 u}{\partial y \partial x} \frac{dy}{dt} \right) + \Delta y \left(\frac{\partial^2 u}{\partial x \partial y} \frac{dx}{dt} + \frac{\partial^2 u}{\partial y^2} \frac{dy}{dt} \right) \\ &= \Delta x^2 \frac{\partial^2 u}{\partial x^2} + 2\Delta x \Delta y \frac{\partial^2 u}{\partial x \partial y} + \Delta y^2 \frac{\partial^2 u}{\partial y^2}. \end{aligned}$$

Como (x^*, y^*) é ponto máximo então $\frac{d^2g}{dt^2}(0) \leq 0$, logo

$$\frac{d^2g}{dt^2}(0) = \Delta x^2 \frac{\partial^2 u}{\partial x^2}(x^*, y^*) + 2\Delta x \Delta y \frac{\partial^2 u}{\partial x \partial y}(x^*, y^*) + \Delta y^2 \frac{\partial^2 u}{\partial y^2}(x^*, y^*) \leq 0. \quad (3.8)$$

Por outro lado, da hipótese do teorema temos:

$$Lu(x^*, y^*) = a_0 \frac{\partial^2 u}{\partial x^2}(x^*, y^*) + 2b_0 \frac{\partial^2 u}{\partial x \partial y}(x^*, y^*) + c_0 \frac{\partial^2 u}{\partial y^2}(x^*, y^*) \geq 0 \quad (3.9)$$

onde $a_0 = a(x^*, y^*)$, $b_0 = b(x^*, y^*)$ e $c_0 = c(x^*, y^*)$.

Podemos reescrever a expressão 3.9 como:

$$Lu(x^*, y^*) = (\sqrt{a_0})^2 \frac{\partial^2 u}{\partial x^2}(x^*, y^*) + 2\sqrt{a_0} \frac{b_0}{\sqrt{a_0}} \frac{\partial^2 u}{\partial x \partial y}(x^*, y^*) + \left(\frac{b_0}{\sqrt{a_0}}\right)^2 \frac{\partial^2 u}{\partial y^2}(x^*, y^*) + \left(c_0 - \frac{(b_0)^2}{a_0}\right) \frac{\partial^2 u}{\partial y^2}(x^*, y^*) \geq 0 \quad (3.10)$$

Dessa forma, tomemos Δx e Δy como:

$$\Delta x = \sqrt{a_0}, \quad \Delta y = \frac{b_0}{\sqrt{a_0}}.$$

Da equação 3.8, temos que:

$$a_0 \frac{\partial^2 u}{\partial x^2}(x^*, y^*) + 2b_0 \frac{\partial^2 u}{\partial x \partial y}(x^*, y^*) + \frac{(b_0)^2}{a_0} \frac{\partial^2 u}{\partial y^2}(x^*, y^*) \leq 0$$

Tomando agora Δx e Δy como

$$\Delta x = 0, \quad \Delta y = \sqrt{\left(c_0 - \frac{(b_0)^2}{a_0}\right)}$$

assim, obtemos pela equação 3.8 que

$$\left(c_0 - \frac{(b_0)^2}{a_0}\right) \frac{\partial^2 u}{\partial y^2}(x^*, y^*) \leq 0.$$

Logo o lado esquerdo da equação 3.9 deve ser negativo, o que é um absurdo, pois, $Lu \geq 0$, logo o ponto máximo deve estar na borda de Ω . \square

Exemplo 3.4.4. Consideremos a seguinte problema de valor de contorno definida no domínio $\Omega = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$.

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 4 & , \text{ para } (x, y) \in \Omega, \\ u(x, y) = 1 & , \text{ para } (x, y) \in \partial\Omega. \end{cases}$$

A solução analítica do problema é um parabolóide com concavidade para cima dado por $u(x, y) = x^2 + y^2$. Considerando o operador L , temos que

$$Lu = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 4 > 0,$$

logo pelo princípio do máximo o ponto máximo de u acontece na borda $\partial\Omega$.

Para o caso discreto, consideraremos um operador definido nos pontos da malha. Usaremos o esquema de diferenças finitas aplicado no problema bidimensional como base para definir o operador discreto.

Para a demonstração do princípio do máximo discreto se utilizou das referências [6, 15].

Definição 3.4.5 (Operador discreto). Seja $U_{i,j}$ uma função definida nos pontos da malha,

então definimos o operador discreto $L_{h,k}$ como

$$L_{h,k}U_{ij} = \delta_h^2 U_{i,j} + \delta_k^2 U_{i,j}$$

Teorema 3.4.6 (Princípio do máximo discreto). *Dado uma função U_{ij} definida malha de pontos, se*

$$L_{h,k}U_{ij} \geq 0$$

para todos os ponto interiores da malha, então o máximo de U_{ij} está na borda do conjunto. Do contrário se

$$L_{h,k}U_{ij} \leq 0$$

então o mínimo está na borda.

Demonstração. Suponha que o teorema é falso, isto é, existe um ponto interior (x_{i_0}, y_{j_0}) , tal que $U_{i_0,j_0} \geq U_{ij}$ para todo i e j . Como $L_{h,k}U_{i_0,j_0} \geq 0$ então

$$\frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{h^2} + \frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{k^2} \geq 0,$$

dessa forma temos:

$$\frac{U_{i-1,j} + U_{i+1,j}}{h^2} + \frac{U_{i,j-1} + U_{i,j+1}}{k^2} \geq \left(\frac{2}{h^2} + \frac{2}{k^2} \right) U_{i,j},$$

isolando $U_{i,j}$, temos então

$$U_{i,j} \leq \left(\frac{2}{h^2} + \frac{2}{k^2} \right)^{-1} \left(\frac{U_{i-1,j} + U_{i+1,j}}{h^2} + \frac{U_{i,j-1} + U_{i,j+1}}{k^2} \right).$$

Por outro lado, como $U_{i_0,j_0} \geq U_{ij}$ para todo i e j , então, em particular temos:

$$U_{i_0,j_0} \geq U_{i_0-1,j_0}, \quad U_{i_0,j_0} \geq U_{i_0+1,j_0}, \quad U_{i_0,j_0} \geq U_{i_0,j_0-1}, \quad U_{i_0,j_0} \geq U_{i_0,j_0+1}. \quad (3.11)$$

Devemos então ter que o ponto máximo é igual aos pontos adjacentes da malha, isto é

$$U_{i_0,j_0} = U_{i_0-1,j_0} = U_{i_0+1,j_0} = U_{i_0,j_0-1} = U_{i_0,j_0+1},$$

pois do contrário, se alguma das expressões em 3.11 forem estritamente desiguais teremos

$$\begin{aligned} U_{i_0,j_0} &\leq \left(\frac{2}{h^2} + \frac{2}{k^2} \right)^{-1} \left(\frac{U_{i_0-1,j_0} + U_{i_0+1,j_0}}{h^2} + \frac{U_{i_0,j_0-1} + U_{i_0,j_0+1}}{k^2} \right) \\ &< \left(\frac{2}{h^2} + \frac{2}{k^2} \right)^{-1} \left(\frac{U_{i_0,j_0} + U_{i_0,j_0}}{h^2} + \frac{U_{i_0,j_0} + U_{i_0,j_0}}{k^2} \right) \\ &= \left(\frac{2}{h^2} + \frac{2}{k^2} \right)^{-1} \left(\frac{2}{h^2} + \frac{2}{k^2} \right) U_{i_0,j_0} \\ &= U_{i_0,j_0}, \end{aligned}$$

o que implicaria em

$$U_{i_0, j_0} < \left(\frac{2}{h^2} + \frac{2}{k^2} \right)^{-1} \left(\frac{U_{i_0-1, j_0} + U_{i_0+1, j_0}}{h^2} + \frac{U_{i_0, j_0-1} + U_{i_0, j_0+1}}{k^2} \right) < U_{i_0, j_0}.$$

Absurdo.

Logo se um ponto (x_{i_0}, y_{j_0}) for ponto máximo ele deverá ser igual à seus pontos adjacentes da malha, assim, os pontos adjacentes da malha também serão pontos máximos. Repetindo o mesmo argumento para todos os pontos internos temos que (x_{i_0}, y_{j_0}) só pode ser máximo se U for uma função constante. \square

Lema 3.4.7. *Seja U_{ij} uma função definida em domínio $[a, b] \times [c, d]$, tal que*

$$L_{h,k}U_{ij} = f_{ij},$$

para $i = 0, 1, \dots, m$, e $j = 0, 1, \dots, n$. Sujeito a condição de contorno $U_{ij} = 0$, então temos que:

$$\|\mathbf{U}\|_{\infty} \leq C \|L_{h,k}\mathbf{U}\|_{\infty},$$

$$\text{onde } C = \frac{(b-a)^2 + (d-c)^2}{16}.$$

Demonstração. Definiremos a função $w = w(x, y)$ no domínio $[a, b] \times [c, d]$ por:

$$w(x, y) = \frac{1}{4} \left[\left(x - \frac{a+b}{2} \right)^2 + \left(y - \frac{c+d}{2} \right)^2 \right].$$

A função $w(x, y)$, define um parabolóide no retângulo $[a, b] \times [c, d]$, centrado no ponto médio $(\frac{a+b}{2}, \frac{c+d}{2})$ do retângulo.

Logo as derivadas parciais são:

$$\frac{\partial w}{\partial x} = \frac{1}{2} \left(x - \frac{a+b}{2} \right), \quad \text{e} \quad \frac{\partial w}{\partial y} = \frac{1}{2} \left(y - \frac{c+d}{2} \right).$$

e portanto,

$$\frac{\partial^2 w}{\partial x^2} = \frac{1}{2}, \quad \text{e} \quad \frac{\partial^2 w}{\partial y^2} = \frac{1}{2}; \quad \frac{\partial^4 w}{\partial x^4} = 0, \quad \text{e} \quad \frac{\partial^4 w}{\partial y^4} = 0. \quad (3.12)$$

Definimos então, uma malha de pontos com espaçamento uniforme $h = \frac{b-a}{m}$ e $k = \frac{d-c}{n}$, que discretiza o domínio $[a, b] \times [c, d]$ do problema. Os pontos (x_i, y_j) da malha são dados da forma

$$x_i = a + ih, \quad \text{onde } i = 0, 1, \dots, m, \quad \text{e } h = \frac{b-a}{m},$$

$$y_j = c + jk, \quad \text{onde } j = 0, 1, \dots, n, \quad \text{e } k = \frac{d-c}{n}.$$

Considerando $w_{i,j} = w(x_i, y_j)$, como pela pelo Teorema de Taylor,

$$\frac{w_{i-1,j} - 2w_{i,j} + w_{i+1,j}}{h^2} = \frac{\partial^2 w}{\partial x^2}(x_i, y_j) + \frac{h^2}{12} \frac{\partial^4 w}{\partial x^4}(\xi_i, \eta_j), \quad (3.13)$$

$$\frac{w_{i,j-1} - 2w_{i,j} + w_{i,j+1}}{k^2} = \frac{\partial^2 w}{\partial y^2}(x_i, y_j) + \frac{k^2}{12} \frac{\partial^4 w}{\partial y^4}(\xi_i, \eta_j), \quad (3.14)$$

para $\xi_i \in (x_i - h, x_i + h)$ e $\eta_j \in (y_j - k, y_j + k)$, então, substituindo 3.12 em 3.13, concluimos que

$$L_{h,k}w_{i,j} = 1.$$

Além disso $Lw(x, y) = 1$, logo pelo princípio do máximo contínuo, o máximo de $w = w(x, y)$ se encontra na borda do domínio $[a, b] \times [c, d]$. Em particular o máximo se encontra nos pontos (a, c) , (a, d) , (b, c) e (b, d) , assim

$$\max_{(x,y)} w(x, y) = \frac{(b-a)^2 + (d-c)^2}{16}.$$

Pela linearidade do operador $L_{h,k}$, temos que

$$\begin{aligned} L_{h,k}(U_{ij} + \|\mathbf{f}\|_\infty w_{ij}) &= L_{h,k}U_{ij} + \|\mathbf{f}\|_\infty L_{h,k}w_{ij}, \\ L_{h,k}(U_{ij} - \|\mathbf{f}\|_\infty w_{ij}) &= L_{h,k}U_{ij} - \|\mathbf{f}\|_\infty L_{h,k}w_{ij}. \end{aligned}$$

onde \mathbf{f} é o vetor contendo os termos $f(x_i, y_j)$. Como $L_{h,k}U_{ij} = f_{ij}$ e $L_{h,k}w_{ij} = 1$, então temos que

$$\begin{aligned} L_{h,k}(U_{ij} + \|\mathbf{f}\|_\infty w_{ij}) &= f_{ij} + \|\mathbf{f}\|_\infty \geq 0, \\ L_{h,k}(U_{ij} - \|\mathbf{f}\|_\infty w_{ij}) &= f_{ij} - \|\mathbf{f}\|_\infty \leq 0. \end{aligned}$$

Assim, pelo princípio do máximo discreto, $U_{ij} + \|\mathbf{f}\|_\infty w_{ij}$ tem o máximo na borda de Ω , enquanto que $U_{ij} - \|\mathbf{f}\|_\infty w_{ij}$, possui o mínimo na borda de Ω .

Por hipótese, $U_{ij} = 0$ na borda de Ω , então temos

$$\begin{aligned} \max(U_{ij} + \|\mathbf{f}\|_\infty w_{ij}) &= \max(\|\mathbf{f}\|_\infty w_{ij}), \\ \min(U_{ij} - \|\mathbf{f}\|_\infty w_{ij}) &= \min(-\|\mathbf{f}\|_\infty w_{ij}), \end{aligned}$$

Como $\|\mathbf{f}\|_\infty w_{ij} \geq 0$, então $\min(-\|\mathbf{f}\|_\infty w_{ij}) = -\max(\|\mathbf{f}\|_\infty w_{ij})$, portanto

$$U_{ij} + \|\mathbf{f}\|_\infty w_{ij} \leq \max(\|\mathbf{f}\|_\infty w_{ij}), \quad (3.15)$$

$$U_{ij} - \|\mathbf{f}\|_\infty w_{ij} \geq -\max(\|\mathbf{f}\|_\infty w_{ij}), \quad (3.16)$$

logo,

$$U_{ij} - \|\mathbf{f}\|_\infty w_{ij} \leq U_{ij} \leq U_{ij} + \|\mathbf{f}\|_\infty w_{ij},$$

pelas equações 3.15 e 3.16 temos que

$$-\max(\|\mathbf{f}\|_\infty w_{ij}) \leq U_{ij} \leq \max(\|\mathbf{f}\|_\infty w_{ij}),$$

logo $|U_{ij}| \leq \max(\|\mathbf{f}\|_\infty w_{ij})$, para todo i, j . Portanto

$$\|\mathbf{U}\|_\infty \leq \max(\|\mathbf{f}\|_\infty w_{ij}). \quad (3.17)$$

Como $\|\mathbf{f}\|_\infty \geq 0$, e $\max w_{ij} = \frac{(b-a)^2 + (d-c)^2}{16}$, então

$$\max(\|\mathbf{f}\|_\infty w_{ij}) = \|\mathbf{f}\|_\infty \max w_{ij} = \frac{(b-a)^2 + (d-c)^2}{16} \|\mathbf{f}\|_\infty. \quad (3.18)$$

Como $L_{h,k}U_{ij} = f_{ij}$, então $\|\mathbf{f}\|_\infty = \|L_{h,k}\mathbf{U}\|_\infty$, logo pelas equações 3.17 e 3.18,

$$\|\mathbf{U}\|_\infty \leq C \|L_{h,k}\mathbf{U}\|_\infty,$$

onde $C = \frac{(b-a)^2 + (d-c)^2}{16}$. □

Teorema 3.4.8. *Seja $U_{i,j}$ a solução de $L_{h,k}U_{i,j} = f_{i,j}$. Então o erro global $\|e_{h,k}\|_\infty$ satisfaz:*

$$\|e_{h,k}\|_\infty \leq C_x h^2 + C_y k^2,$$

onde C_x e C_y são constantes da forma

$$C_x = \frac{(b-a)^2 + (d-c)^2}{192} \left\| \frac{\partial^4 u}{\partial x^4} \right\|_\infty, \quad e \quad C_y = \frac{(b-a)^2 + (d-c)^2}{192} \left\| \frac{\partial^4 u}{\partial y^4} \right\|_\infty.$$

Demonstração. Pelo Teorema de Taylor, o esquema de diferenças finitas centradas satisfaz

$$\delta_h^2 u + \delta_k^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4}(\xi, \eta) + \frac{k^2}{12} \frac{\partial^4 u}{\partial y^4}(\xi, \eta). \quad (3.19)$$

Por causa da suavidade admitida da função, as derivadas da função devem ser contínuas e portanto admitir máximo. Assim, temos que

$$\left| \frac{\partial^4 u}{\partial x^4}(\xi, \eta) \right| \leq \left\| \frac{\partial^4 u}{\partial x^4} \right\|_\infty, \quad e \quad \left| \frac{\partial^4 u}{\partial y^4}(\xi, \eta) \right| \leq \left\| \frac{\partial^4 u}{\partial y^4} \right\|_\infty,$$

Portanto, pela equação 3.19, temos que

$$\|\boldsymbol{\tau}_{h,k}\|_\infty = \|L_{h,k}\mathbf{u} - L\mathbf{u}\|_\infty \leq \frac{h^2}{12} \left\| \frac{\partial^4 u}{\partial x^4} \right\|_\infty + \frac{k^2}{12} \left\| \frac{\partial^4 u}{\partial y^4} \right\|_\infty. \quad (3.20)$$

Como $L_{h,k}e_{h,k} = \boldsymbol{\tau}$, aplicando o lema 3.4.7 no vetor \mathbf{e} , concluímos que

$$\|e_{h,k}\|_\infty \leq C \|L_{h,k}e_{h,k}\|_\infty = C \|\boldsymbol{\tau}_{h,k}\|_\infty,$$

onde $C = \frac{(b-a)^2 + (d-c)^2}{16}$. Logo pela equação 3.20,

$$\|\mathbf{e}_{h,k}\|_{\infty} \leq \frac{(b-a)^2 + (d-c)^2}{16} \left(\frac{h^2}{12} \left\| \frac{\partial^4 u}{\partial x^4} \right\|_{\infty} + \frac{k^2}{12} \left\| \frac{\partial^4 u}{\partial y^4} \right\|_{\infty} \right).$$

Dessa forma, o erro global \mathbf{e} satisfaz

$$\|\mathbf{e}_{h,k}\|_{\infty} \leq C_x h^2 + C_y k^2,$$

onde C_x e C_y são constantes da forma

$$C_x = \frac{(b-a)^2 + (d-c)^2}{192} \left\| \frac{\partial^4 u}{\partial x^4} \right\|_{\infty}, \quad \text{e} \quad C_y = \frac{(b-a)^2 + (d-c)^2}{192} \left\| \frac{\partial^4 u}{\partial y^4} \right\|_{\infty}.$$

□

4. MÉTODOS ITERATIVOS

Ao longo dos capítulos 2 e 3, aplicamos o método das diferenças finitas para problemas de valores de contorno sujeito às condições de tipo Dirichlet. Uma vez substituídas as derivadas da equação diferencial por diferenças finitas, e aplicado as condições de contorno, obtivemos um sistema linear com uma estrutura

$$A\mathbf{x} = \mathbf{b},$$

onde A era uma matriz esparsa, que é uma matriz majoritariamente composta por zeros.

Para resolvermos os sistemas da forma $A\mathbf{x} = \mathbf{b}$, usaremos os chamados métodos iterativos. Os métodos iterativos são métodos que consistem em atribuir uma aproximação inicial $\mathbf{x}^{(0)}$, e a partir disso criar uma sequência de aproximações $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \dots$ que espera-se que convirja para a solução exata \mathbf{x} do sistema linear.

Os métodos iterativos que estudaremos aqui neste trabalho são conhecidos como métodos iterativos estacionários, pois cada aproximação $\mathbf{x}^{(k)}$ é obtida a partir da aproximação anterior $\mathbf{x}^{(k-1)}$, sempre através do mesmo processo iterativo.

Para o desenvolvimento desse capítulo usamos como referências: [4], [8], [6], [11], [12] e [13].

4.1 CONCEITOS PRELIMINARES

Antes de definirmos os métodos iterativos estacionários, precisamos definir alguns conceitos sobre convergências que serão usados para fazer uma análise dos métodos iterativos.

Definição 4.1.1. Dado uma sequência de vetores $(\mathbf{x}^{(k)})_{k=0}^{\infty}$ e uma norma $\|\cdot\|$, dizemos que $\mathbf{x}^{(k)}$ converge para $\mathbf{x} \in \mathbb{R}^n$, se

$$\|\mathbf{x}^{(k)} - \mathbf{x}\| \rightarrow 0, \text{ quando } k \rightarrow \infty.$$

Definição 4.1.2. Uma matriz A é chamada de convergente se

$$\lim_{k \rightarrow \infty} A^k = O,$$

onde O é a matriz nula.

Exemplo 4.1.3. Dado a matriz A definida por

$$A = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix},$$

então para todo $k \in \mathbb{N}$, temos que

$$A^k = \begin{pmatrix} \frac{1}{2^k} & 0 & 0 \\ 0 & \frac{1}{2^k} & 0 \\ 0 & 0 & \frac{1}{2^k} \end{pmatrix},$$

logo $\lim_{k \rightarrow \infty} A^k = O$, e portanto A é convergente.

Sabemos que $\|A^k - O\| = \|A^k\| = |\|A^k\| - 0|$, portanto podemos dizer que A é convergente se $\lim_{k \rightarrow \infty} \|A^k\| = 0$.

Lema 4.1.4. *Dado uma matriz A , para todo $\epsilon > 0$, existe uma norma $\|\cdot\|$, tal que $\|A\| \leq \rho(A) + \epsilon$.*

Demonstração. Demonstração pode ser encontrada em [4]. □

Teorema 4.1.5. *Dado uma matriz A , são equivalentes as seguintes*

(a) $\lim_{k \rightarrow \infty} A^k = O$.

(b) $\lim_{k \rightarrow \infty} A^k \mathbf{x} = \mathbf{0}$, para todo $\mathbf{x} \in \mathbb{R}^n$.

(c) $\rho(A) < 1$.

(d) *Existe uma norma $\|\cdot\|$ tal que $\|A\| < 1$.*

Demonstração. (a) \Rightarrow (b): Temos que $\|A^k \mathbf{x}\| \leq \|A^k\| \|\mathbf{x}\|$ para todo $\mathbf{x} \in \mathbb{R}^n$, como A é convergente, $\|A^k\| \rightarrow 0$, e portanto pela desigualdade $\|A^k \mathbf{x}\| \rightarrow 0$, assim, $\lim_{k \rightarrow \infty} A^k \mathbf{x} = \mathbf{0}$.

(b) \Rightarrow (c): Dado um autovalor λ com autovetor associado \mathbf{v} de A , como $\|A^k \mathbf{v}\| = \|\lambda^k \mathbf{v}\| = |\lambda|^k \|\mathbf{v}\|$, então por $\lim_{k \rightarrow \infty} A^k \mathbf{x} = \mathbf{0}$, temos então que

$$\lim_{k \rightarrow \infty} |\lambda|^k \|\mathbf{v}\| = 0,$$

e portanto $|\lambda| < 1$, para todo autovalor λ de A . Logo $\rho(A) < 1$

(c) \Rightarrow (d): Dado $\rho(A) < 1$, tomemos $\epsilon > 0$ tal que $\epsilon < 1 - \rho(A)$, então $\rho(A) + \epsilon < 1$. Pelo lema 4.1.4, existe uma norma $\|\cdot\|$ tal que $\|A\| \leq \rho(A) + \epsilon$, e portanto $\|A\| < 1$.

(d) \Rightarrow (a): Dado uma norma $\|\cdot\|$ tal que $\|A\| < 1$, então $\lim_{k \rightarrow \infty} \|A\|^k = 0$. Pela propriedade sub-multiplicativa da norma de matriz temos que $\|A^k\| \leq \|A\|^k$, logo, $\|A^k\| \rightarrow 0$ quando $k \rightarrow \infty$, portanto,

$$\lim_{k \rightarrow \infty} A^k = O.$$

□

Dado uma matriz A , algumas propriedades dessa matriz podem ser observadas a partir da construção de uma grafo direcionado $G(A)$, associado à matriz A .

Definição 4.1.6. Seja A uma matriz de dimensões $n \times n$. Então o grafo (direcionado) $G(A)$ associado à matriz A é constituído pelos vértices P_1, P_2, \dots, P_n e pelas arestas $P_i \rightarrow P_j$, para todo i, j tais que $a_{ij} \neq 0$.

Exemplo 4.1.7. Consideremos a matriz A da forma

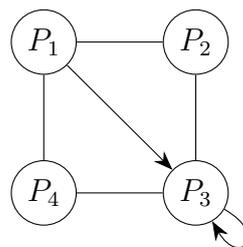
$$A = \begin{bmatrix} 0 & 2 & 3 & 1 \\ 1 & 0 & 5 & 0 \\ 0 & 4 & 2 & 7 \\ 5 & 0 & 8 & 0 \end{bmatrix}.$$

Então o grafo $G(A)$ associado à A é constituído pelos vértices P_1, P_2, P_3 e P_4 . A partir dos termos $a_{ij} \neq 0$ podemos construir um conjunto de arestas da forma $P_i \rightarrow P_j$, ou seja, para cada $a_{ij} \neq 0$, temos

$$\begin{array}{ll} a_{12} : P_1 \rightarrow P_2, & a_{21} : P_2 \rightarrow P_1, \\ a_{14} : P_1 \rightarrow P_4, & a_{41} : P_4 \rightarrow P_1, \\ a_{23} : P_2 \rightarrow P_3, & a_{32} : P_3 \rightarrow P_2, \\ a_{34} : P_3 \rightarrow P_4, & a_{43} : P_4 \rightarrow P_3, \\ a_{13} : P_1 \rightarrow P_3, & a_{33} : P_3 \rightarrow P_3. \end{array}$$

Portanto o grafo $G(A)$ tem a forma de acordo com a figura 4.1.

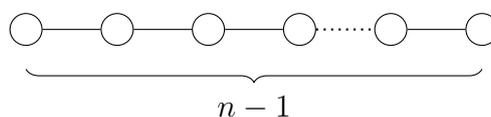
Figura 4.1: Grafo $G(A)$ associado à matriz A .



Fonte: O autor.

Exemplo 4.1.8. A matriz A definida por 2.9 para o problema de valor de contorno unidimensional $u''(x) = f(x)$, sujeito a condições de contorno de tipo Dirichlet, é representado, ignorando os laços, por um grafo $G(A)$ de acordo com a figura 4.2.

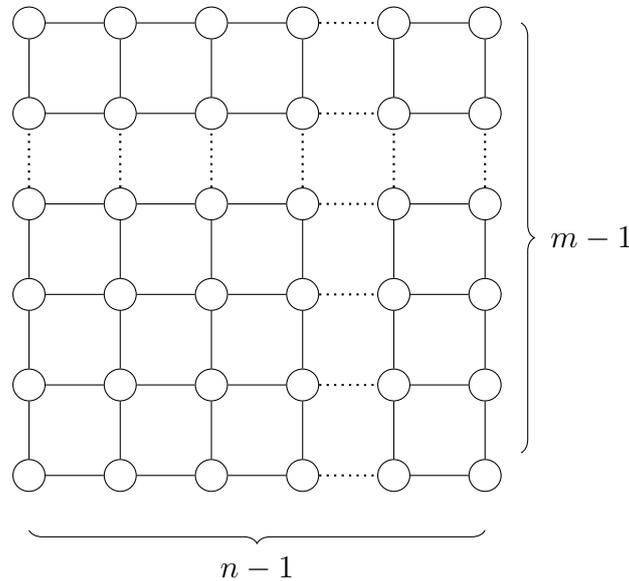
Figura 4.2: Grafo $G(A)$ da matriz do problema de unidimensional.



Fonte: O autor.

Exemplo 4.1.9. A matriz A definida por 3.6 para o problema de Poisson bidimensional sujeito a condições de contorno do tipo Dirichlet, é representado, ignorando os laços, por um grafo $G(A)$ representado de acordo com a figura 4.3.

Figura 4.3: Grafo $G(A)$ da matriz do problema de Poisson.



Fonte: O autor.

Definição 4.1.10. Um grafo (direcionado) é chamado de fortemente conexo, se para cada par de vértices $P_i, P_j \in G(A)$, existe um caminho de P_i à P_j .

Os grafos definidos pela matriz do problema de unidimensional, e pela matriz problema de Poisson possuem a propriedade de serem fortemente conexos.

4.2 MÉTODOS ITERATIVOS ESTACIONÁRIOS

Dado um sistema linear com n equações e n incógnitas da forma $A\mathbf{x} = \mathbf{b}$, com $\det(A) \neq 0$ podemos decompor a matriz A como $A = M - N$, de forma que $\det M \neq 0$. Assim, o sistema linear pode ser reescrito como

$$M\mathbf{x} = N\mathbf{x} + \mathbf{b},$$

se multiplicarmos ambos os lados da equação por M^{-1} obtemos

$$\mathbf{x} = M^{-1}N\mathbf{x} + M^{-1}\mathbf{b}.$$

Considerando a matriz $T := M^{-1}N$, e o vetor $\mathbf{c} := M^{-1}\mathbf{b}$, podemos reescrever o sistema linear $A\mathbf{x} = \mathbf{b}$ como

$$\mathbf{x} = T\mathbf{x} + \mathbf{c}. \quad (4.1)$$

Assim, encontrar a solução do sistema linear $A\mathbf{x} = \mathbf{b}$ é equivalente à encontrar a solução do sistema linear 4.1.

Definimos uma sequência de vetores $(\mathbf{x}^{(k)})_{k=0}^{\infty}$, mediante o processo iterativo:

$$\mathbf{x}^{(k+1)} = T\mathbf{x}^{(k)} + \mathbf{c}, \quad (4.2)$$

onde a matriz T é chamada de matriz de iteração. Para analisarmos a convergência do método iterativo, olharemos para o erro de aproximação entre a k -ésima iteração $\mathbf{x}^{(k)}$ do método, e a solução exata \mathbf{x} do sistema linear.

Definição 4.2.1. Dado uma sequência de vetores $(\mathbf{x}^{(k)})_{k=0}^{\infty}$, definimos o erro $\mathbf{e}^{(k)}$ da k -ésima iteração por

$$\mathbf{e}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}.$$

A partir do erro $\mathbf{e}^{(k)}$ da k -ésima aproximação, analisaremos como o método se comporta quando $k \rightarrow \infty$. Para o erro $\mathbf{e}^{(k+1)} = \mathbf{x} - \mathbf{x}^{(k+1)}$, temos que

$$\mathbf{x} - \mathbf{x}^{(k+1)} = (T\mathbf{x} + \mathbf{c}) - (T\mathbf{x}^{(k)} + \mathbf{c}) = T(\mathbf{x} - \mathbf{x}^{(k)}),$$

para $k = 0, 1, 2, \dots$, portanto,

$$\mathbf{e}^{(k+1)} = T\mathbf{e}^{(k)}, \text{ para } k \in \mathbb{N}.$$

Assim, dado um erro inicial $\mathbf{e}^{(0)}$, o k -ésimo erro da aproximação da solução do sistema será

$$\mathbf{e}^{(k)} = T\mathbf{e}^{(k-1)} = T^2\mathbf{e}^{(k-2)} = \dots = T^{k-1}\mathbf{e}^{(1)} = T^k\mathbf{e}^{(0)},$$

e portanto, podemos relacionar o k -ésimo erro com o erro inicial $\mathbf{e}^{(0)}$ por

$$\mathbf{e}^{(k)} = T^k\mathbf{e}^{(0)}, \text{ para todo } k \in \mathbb{N}.$$

Consideremos uma sequência de aproximação $(\mathbf{x}^{(k)})_{k=0}^{\infty}$ definidas esquema iterativo 4.2, essa sequência convergirá para a solução \mathbf{x} se

$$\lim_{k \rightarrow \infty} \mathbf{e}^{(k)} = \lim_{k \rightarrow \infty} [\mathbf{x} - \mathbf{x}^{(k)}] = \mathbf{0}.$$

Como $\mathbf{e}^{(k)} = T^k\mathbf{e}^{(0)}$, e o erro inicial $\mathbf{e}^{(0)}$ é constante, então temos que o método convergirá se

$$\lim_{k \rightarrow \infty} T^k\mathbf{e}^{(0)} = \mathbf{0}, \text{ para todo } \mathbf{e}^{(0)} \in \mathbb{R}^n.$$

Logo, pelo teorema 4.1.5, o método iterativo convergirá se, e somente se, a matriz de iteração T for convergente. Assim, podemos analisar a convergência de qualquer método iterativo da forma 4.2, a partir de sua matriz de iteração.

4.3 MÉTODO DE JACOBI

O primeiro método iterativo que iremos desenvolver, é conhecido como Método de Jacobi, ele é um dos métodos iterativos mais simples e servirá de base para a definição dos próximos métodos.

Consideremos o sistema linear de equações da forma $A\mathbf{x} = \mathbf{b}$ com n equações e n incógnitas, com $\det A \neq 0$. Então o sistema $A\mathbf{x} = \mathbf{b}$ possui a forma:

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

Podemos escrever o sistema $A\mathbf{x} = \mathbf{b}$ na forma

$$\sum_{j=1}^n a_{ij}x_j = b_i, \text{ para } i = 1, 2, \dots, n.$$

Assim, reescrevemos o sistema linear como

$$\sum_{j=1}^{i-1} a_{ij}x_j + a_{ii}x_i + \sum_{j=i+1}^n a_{ij}x_j = b_i, \text{ para } i = 1, 2, \dots, n,$$

isolando o termo $a_{ii}x_i$,

$$a_{ij}x_j = b_i - \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^n a_{ij}x_j, \text{ para } i = 1, 2, \dots, n,$$

ao dividirmos ambos os lados da equação por a_{ii} , obtemos

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

Assim, o método de Jacobi é dado por

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

Podemos obter o método de Jacobi também na forma matricial, consideremos $A = D - L - U$, onde D é uma matriz diagonal, L é uma matriz triangular inferior e U é uma matriz triangular

superior. Então podemos reescrever o sistema $A\mathbf{x} = \mathbf{b}$ como

$$(D - L - U)\mathbf{x} = \mathbf{b}.$$

Isolando o termo $D\mathbf{x}$, temos que

$$D\mathbf{x} = (L + U)\mathbf{x} + \mathbf{b},$$

multiplicando ambos os lados por D^{-1} , obtemos

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

Assim, o método de Jacobi na forma matricial é dado por

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

A matriz de iteração do método de Jacobi é da forma $D^{-1}(L + U)$, assim

$$J := D^{-1}(L + U).$$

Assim, a matriz de iteração J terá a forma

$$J = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & -\frac{a_{23}}{a_{22}} & \dots & -\frac{a_{2n}}{a_{22}} \\ -\frac{a_{31}}{a_{33}} & -\frac{a_{32}}{a_{33}} & 0 & \dots & -\frac{a_{3n}}{a_{33}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & -\frac{a_{n3}}{a_{nn}} & \dots & 0 \end{bmatrix}.$$

Para analisarmos a convergência do método, analisaremos a matriz de iteração do método. Pelo teorema 4.1.5, se existir uma norma $\|\cdot\|$, tal que $\|J\| < 1$, então a matriz de iteração J do método de Jacobi é convergente.

Se tomarmos a norma normas matriciais $\|\cdot\|_1$ e $\|\cdot\|_\infty$, então

$$\|J\|_1 = \max \left\{ \sum_{\substack{i=1 \\ i \neq j}}^n \left| \frac{a_{ij}}{a_{ii}} \right| : j = 1, 2, \dots, n \right\},$$

$$\|J\|_\infty = \max \left\{ \sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right| : i = 1, 2, \dots, n \right\}.$$

Para garantirmos a convergência do método iterativo de Jacobi, pelo teorema 4.1.5, basta

que $\|J\|_1 < 1$, ou $\|J\|_\infty < 1$. Podemos, além disso definir o seguinte critério para a convergência do método de Jacobi:

Teorema 4.3.1. *Seja $A = (a_{ij})$ uma matriz de dimensão $n \times n$, e $J = D^{-1}(L + U)$ a matriz de iteração do método de Jacobi, se*

$$\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| < |a_{ii}|, \text{ para todo } i = 1, 2, \dots, n. \quad (4.3)$$

então $\|J\|_\infty < 1$, e portanto o método de Jacobi converge.

Demonstração. Se a matriz A satisfaz

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

para todo $i = 1, 2, \dots, n$. Então dividindo ambos os lados da equação por $|a_{ii}|$, temos que

$$\sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right| = \frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| < 1$$

para todo $i = 1, 2, \dots, n$. Logo $\|J\|_\infty < 1$, e portanto o método iterativo de Jacobi converge. \square

Uma matriz que satisfaz a equação 4.3 é chamada de estritamente diagonal dominante.

4.4 MÉTODO DE GAUSS-SEIDEL

Conforme visto na seção anterior, o método iterativo de Jacobi possuía a forma

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right), \quad (4.4)$$

para $i = 1, 2, \dots, n$. Podemos ver na equação 4.4, que para calcularmos o termo $x_i^{(k+1)}$, utilizamos os termos $x_1^{(k)}, x_2^{(k)}, \dots, x_{i-1}^{(k)}$ obtidos na iteração anterior. Porém, quando vamos calcular o valor $x_i^{(k+1)}$, já conhecemos quais são os valores anteriores $x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)}$, obtidos na mesma iteração.

Assim, buscando acelerar a convergência do método, podemos utilizar em 4.4, os valores de $x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)}$ no lugar de $x_1^{(k)}, x_2^{(k)}, \dots, x_{i-1}^{(k)}$. Dessa forma, obtemos o método iterativo de Gauss-Seidel:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right), \quad (4.5)$$

para $i = 1, 2, \dots, n$.

Para obtermos o método de Gauss-Seidel na forma matricial, devemos fazer a separação $A = D - L - U$, de maneira análoga a como foi feito na seção anterior para o método de Jacobi. Assim, dado o sistema $A\mathbf{x} = \mathbf{b}$,

$$(D - L - U)\mathbf{x} = \mathbf{b}.$$

Somando o termo $U\mathbf{x}$ à ambos os lados da equação, obtemos

$$(D - L)\mathbf{x} = U\mathbf{x} + \mathbf{b},$$

assim, a sequência $(\mathbf{x}^{(k)})_{k=0}^{\infty}$ deverá satisfazer

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

Dessa forma, o método de iterativo de Gauss-Seidel na forma matricial é

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

Sendo $(D - L)^{-1}U$ a matriz de iteração, definimos então a matriz de iteração do método iterativo de Gauss-Seidel por

$$G := (D - L)^{-1}U.$$

Teorema 4.4.1. *Seja $A = (a_{ij})$ uma matriz de dimensão $n \times n$, e $G = (D - L)^{-1}U$ a matriz de iteração do método de Gauss-Seidel, se*

$$\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| < |a_{ii}|, \text{ para todo } i = 1, 2, \dots, n.$$

então $\|G\|_{\infty} < 1$, e portanto o método de Gauss-Seidel converge.

A demonstração do teorema pode ser encontrada em [12].

Portanto, para matrizes A que sejam estritamente diagonais dominantes, ambos os métodos de Jacobi e Gauss-Seidel convergem.

Este resultado, porém não é suficiente para garantir a convergência das matrizes obtidas a partir do problema unidimensional $u''(x) = f(x)$ e do problema de Poisson bidimensional, pois ambas as matrizes não são estritamente diagonal dominantes.

Portanto precisamos de um outro critério para garantir a convergência dos métodos iterativos de Jacobi e Gauss-Seidel. Com base nisso introduzimos o conceito de matriz francamente diagonal dominante:

Definição 4.4.2. Uma matriz $A = (a_{ij})$ é francamente diagonal dominante se

$$\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \leq |a_{ii}|, \text{ para todo } i = 1, 2, \dots, n.$$

e existe um $i = 1, 2, \dots, n$, tal que vale a desigualdade estrita $\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| < |a_{ii}|$.

As matrizes do problema unidimensional $u''(x) = f(x)$ e do problema Poisson bidimensional são ambas fracamente diagonal dominante.

Definição 4.4.3. Uma matriz A de dimensão $n \times n$ é chamada redutível se existir uma matriz de permutação P tal que

$$PAP^t = \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix},$$

onde O é uma matriz nula, A_{11} é uma matriz de dimensão $r \times r$ e A_{22} é uma matriz de dimensão $(n - r) \times (n - r)$. Se não existir nenhuma permutação que satisfaça a definição, então A é dita ser irredutível.

Assim, podemos dizer que uma matriz irredutível é uma matriz que não pode ser transformada em uma matriz triangular superior por blocos por meio de uma matriz de permutação P . É possível analisar se uma matriz A é irredutível a partir de seu grafo $G(A)$ com a ajuda do seguinte teorema:

Teorema 4.4.4. Dada uma matriz A e seu grafo $G(A)$, então A é irredutível se, e somente se, $G(A)$ é fortemente conexo.

A demonstração do teorema pode ser encontrada em [4].

Pelos exemplos 4.1.8 e 4.1.9, os grafos associados as matrizes do problema unidimensional e do problema de Poisson são ambos fortemente conexos, portanto ambas as matrizes são irredutíveis.

Teorema 4.4.5. Se A é uma matriz irredutível e fracamente diagonal dominante, então os métodos de Jacobi e Gauss-Seidel convergem.

A demonstração do teorema pode ser encontrada em [1].

Como as matrizes do problema unidimensional $u''(x) = f(x)$ e do problema Poisson bidimensional são ambas irredutíveis e fracamente diagonal dominante, então pelo teorema 4.4.5, os métodos iterativos de Jacobi e Gauss-Seidel convergem para o problema unidimensional $u''(x) = f(x)$ e para o problema de Poisson bidimensional.

4.5 MÉTODO DE SOBRE-RELAXAÇÃO SUCESSIVA

Conforme vimos na seção anterior, no método iterativo de Gauss-Seidel, as aproximações $\mathbf{x}^{(k)}$ da solução do sistema linear, são dadas pelo esquema iterativo:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right), \quad (4.6)$$

Aqui apresentaremos o método de Sobre-relaxação Sucessiva, também conhecido como método de SOR (*Successive Over-relaxation*), que é um método que extrapola a solução obtida a partir do método de Gauss-Seidel, proporcionando uma taxa de convergência mais rápida para a solução do sistema.

O método de Sobre-relaxação sucessiva é um método iterativo que consiste em fazer uma soma, usando um peso ω , entre a aproximação atual $\mathbf{x}^{(k)}$ e a próxima aproximação $\mathbf{x}_{gs}^{(k+1)}$ obtida pelo método de Gauss-Seidel, obtendo assim uma nova aproximação $\mathbf{x}^{(k+1)}$ da forma

$$\mathbf{x}^{(k+1)} = (1 - \omega)\mathbf{x}^{(k)} + \omega\mathbf{x}_{gs}^{(k+1)}, \quad (4.7)$$

onde, $\mathbf{x}_{gs}^{(k+1)}$ é calculado conforme 4.6, usando os valores $\mathbf{x}^{(k)}$, e ω é um coeficiente real. Podemos observar que se $\omega = 1$, então $\mathbf{x}^{(k+1)} = \mathbf{x}_{gs}^{(k+1)}$, portanto o método de Gauss-Seidel é o caso particular do método de SOR quando $\omega = 1$.

Portanto, por 4.6 e 4.7, concluímos que a sequência de aproximações $(\mathbf{x}^{(k)})_{k=0}^{\infty}$ da solução \mathbf{x} determinadas pelo método de Sobre-relaxação Sucessiva (SOR) é definida pelo seguinte esquema iterativa:

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \frac{\omega}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right), \quad (4.8)$$

para $\omega \in \mathbb{R}$.

Assim como nas secções anteriores, é possível obter uma forma matricial para método iterativo de SOR, para isso, multiplicamos ambos os lados por a_{ii} ,

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

podemos então reescrever a equação como

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

colocando a equação na forma matricial, obtemos

$$(D - \omega L)x^{(k+1)} = [(1 - \omega)D + \omega U]x^{(k)} + \omega b,$$

multiplicando ambos os lados da equação por D^{-1} , obtemos a forma

$$(I - \omega D^{-1}L)x^{(k+1)} = [(1 - \omega)I + \omega D^{-1}U]x^{(k)} + \omega D^{-1}b. \quad (4.9)$$

Afim de simplificar a notação das equações que seguirem no texto, introduziremos as seguinte matrizes:

$$L^* := D^{-1}L, \quad U^* := D^{-1}U.$$

Assim, podemos reescrever a equação 4.9 como

$$(I - \omega L^*)x^{(k+1)} = [(1 - \omega)I + \omega U^*]x^{(k)} + b.$$

e portanto o método iterativo de *SOR* é dado por

$$x^{(k+1)} = (I - \omega L^*)^{-1}[(1 - \omega)I + \omega U^*]x^{(k)} + (I - \omega L^*)^{-1}b. \quad (4.10)$$

Podemos então definir a matriz de iteração $H(\omega)$ do método iterativo de *SOR* por

$$H(\omega) := (I - \omega L^*)^{-1}[(1 - \omega)I + \omega U^*],$$

para $\omega \in \mathbb{R}$, $L^* := D^{-1}L$, e $U^* := D^{-1}U$.

O próximo teorema permitirá determinar os valores de ω para que o qual o método de *SOR* convergirá.

Teorema 4.5.1. *Seja $H(\omega)$ a matriz de iteração do método de *SOR*, então*

$$\rho(H) \geq |\omega - 1|.$$

Demonstração. A matriz $I - \omega L^*$ é uma matriz triangular inferior com diagonais iguais à 1, desta forma o seu determinante é o produto dos elementos da diagonal principal, ou seja, $\det(I - \omega L^*) = 1$, para todo ω .

Seja $p(\lambda)$ o polinômio característico da matriz $H(\omega)$, então

$$p(\lambda) = \det(\lambda I - H(\omega)) = 1 \cdot \det(\lambda I - H(\omega)),$$

como $\det(I - \omega L^*) = 1$, então

$$\det(I - \omega L^*) \cdot \det(\lambda I - H(\omega)) = \det((I - \omega L^*)(\lambda I - H(\omega))),$$

expandindo o termo $(I - \omega L^*)(\lambda I - H(\omega))$, obtemos

$$p(\lambda) = \det((\lambda + \omega - 1)I - \omega \lambda L^* - \omega U^*).$$

Seja n a dimensão da matriz A , e seja $\lambda_1, \lambda_2, \dots, \lambda_n$, as raízes do polinômio característico $p(\lambda)$, então

$$p(\lambda) = a(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n),$$

além disso, como o polinômio característico de $H(\omega)$ é $p(\lambda) = \det(\lambda I - H(\omega))$, então o coeficiente do termo dominante λ^n de $p(\lambda)$ será 1 para n par e -1 para n ímpar, e portanto $a = (-1)^n$, logo

$$p(\lambda) = (-1)^n \prod_{i=1}^n (\lambda - \lambda_i), \quad (4.11)$$

onde n é a dimensão da matriz A , avaliando o polinômio característico no zero, temos que

$$(-1)^n \prod_{i=1}^n \lambda_i = p(0) = \det((\omega - 1)I - \omega U^*),$$

Como $(\omega - 1)I - \omega U^*$ é uma matriz triangular superior, com diagonais $\omega - 1$, então

$$p(0) = \det((\omega - 1)I - \omega U^*) = (\omega - 1)^n. \quad (4.12)$$

Seja $\mu = \rho(H(\omega))$, o raio espectral de $H(\omega)$, então $|\lambda_i| \leq \mu$ para todos os autovalores λ_i de $H(\omega)$, portanto pelas equações 4.11 e 4.12 temos que

$$|\omega - 1|^n = |p(0)| = \left| (-1)^n \prod_{i=1}^n \lambda_i \right| = \prod_{i=1}^n |\lambda_i| \leq \prod_{i=1}^n \mu = \mu^n.$$

Dessa forma $|\omega - 1|^n \leq \mu^n$, concluímos então que

$$|\omega - 1| \leq \rho(H).$$

□

Para que o método de SOR seja convergente, é necessária que o raio espectral da matriz de iteração $H(\omega)$ seja menor do que 1, aplicando o teorema 4.5.1 concluímos que

$$|\omega - 1| \leq \rho(H(\omega)) < 1,$$

assim, uma condição necessária para a convergência do método é que $\omega \in (0, 2)$.

Teorema 4.5.2. *Seja A uma matriz simétrica definida positivamente, então*

$$\rho(H(\omega)) < 1,$$

para todo $0 < \omega < 2$.

A demonstração do teorema pode ser encontrado em [4] ou [12].

Em particular, como o método converge para $\omega = 1$, então o método de Gauss-Seidel converge para toda matriz simétrica definida positivamente.

Exemplo 4.5.3. As matrizes obtidas pelo problema unidimensional $u''(x) = f(x)$ e pelo problema de Poisson bidimensional, ambos com condições de contorno do tipo Dirichlet, são simétricas definidas positivamente, logo o método é SOR converge para $0 < \omega < 2$ para ambas as matrizes.

4.6 MATRIZES CONSISTENTEMENTE ORDENADAS

Nesta seção iremos introduzir um tipo especial de matrizes, as matrizes consistentemente ordenadas. Estas matrizes possuem propriedades que permitem deduzir informações a respeito

de seus autovalores.

Iremos ver que para essas matrizes, o método de Gauss-Seidel sempre converge 2 vezes mais rápido que o método de Jacobi, além disso, também será possível encontrar qual o melhor parâmetro ω para ser usado no método de Sobre-relação Sucessiva.

Para essa seção usaremos como base as referências [4, 11, 12].

Definição 4.6.1. Uma matriz A é consistentemente ordenada se os vértices do grafo $G(A)$ associado à matriz A podem ser particionados em p conjuntos S_1, S_2, \dots, S_p de tal forma que quaisquer dois vértices adjacentes P_i e P_j pertençam a duas partições consecutivas de forma que:

$$\text{Se } i < j \text{ e } P_i \in S_k, \text{ então } P_j \in S_{k+1},$$

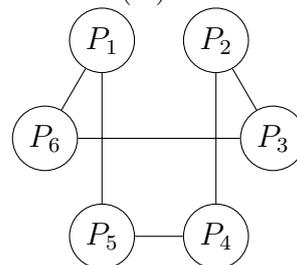
$$\text{Se } i > j \text{ e } P_i \in S_k, \text{ então } P_j \in S_{k-1}.$$

Exemplo 4.6.2. Tomemos como exemplo a seguinte matriz

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 1 & 2 & 0 \\ 1 & 0 & 1 & 0 & 0 & 2 \end{bmatrix}.$$

O grafo $G(A)$ associado a matriz A é representado, ignorando as laços, de acordo com a figura 4.4.

Figura 4.4: Grafo $G(A)$ associado à matriz A .



Fonte: O autor.

Para mostrar que a matriz A é consistentemente ordenada, podemos tomar partições do grafo da forma

$$S_1 = \{P_2\}, \quad S_2 = \{P_1, P_3, P_4\}, \quad S_3 = \{P_5, P_6\}.$$

Considerando as partições S_1 , S_2 e S_3 , podemos analisar cada par de vértices adjacentes do grafo $G(A)$ para verificar se a matriz do grafo é consistentemente ordenada. Analisando cada

par de vértices concluímos que:

- para $1 < 5$, temos $P_1 \in S_2$ e $P_5 \in S_3$,
- para $1 < 6$, temos $P_1 \in S_2$ e $P_6 \in S_3$,
- para $2 < 3$, temos $P_2 \in S_1$ e $P_3 \in S_2$,
- para $2 < 4$, temos $P_2 \in S_1$ e $P_4 \in S_2$,
- para $3 < 6$, temos $P_3 \in S_2$ e $P_6 \in S_3$,
- para $4 < 5$, temos $P_4 \in S_2$ e $P_5 \in S_3$.

Portanto, pela definição, a matriz A é consistentemente ordenada.

Exemplo 4.6.3. A matriz obtida a partir do problema de Poisson bidimensional com condições de contorno de tipo Dirichlet é consistentemente ordenada, tanto na ordenação natural quanto na ordenação *red-black*.

Para a ordenação *red-black* basta tomar duas partições do grafo S_1 e S_2 de tal forma que S_1 contenha os pontos tipo red e S_2 contenha os pontos tipo black.

Para a ordenação natural, tanto por linhas, quanto por colunas, peguemos o conjunto de partições S_k da forma

$$S_k = \{(x_i, y_j) : k = i + j\}.$$

O próximo teorema nos dá uma caracterização equivalente para matrizes consistentemente ordenadas.

Teorema 4.6.4. *Uma matriz A é consistentemente ordenada se, e somente se, os autovalores da matriz*

$$B(\alpha) = \alpha L^* + \frac{1}{\alpha} U^*,$$

são independentes de α , para todo $\alpha \neq 0$.

A demonstração do teorema pode ser encontrado em [12].

Os próximos teoremas nos darão resultados importantes sobre métodos iterativos para matrizes consistentemente ordenadas. E permitirá determinar o melhor valor ω para a convergência das matrizes consistentemente ordenadas.

Teorema 4.6.5. *Seja A uma matriz consistentemente ordenada, e seja $\omega \neq 0$, então*

1. *Os autovalores de J aparecem em pares positivos e negativos.*
2. *Se μ é um autovalor de J e*

$$(\lambda + \omega - 1)^2 = \lambda \omega^2 \mu^2, \tag{4.13}$$

então λ é um autovalor de $H(\omega)$.

3. *Se $\lambda \neq 0$ é um autovalor de $H(\omega)$ e a equação 4.13 é satisfeita, então μ é um autovalor de J .*

Demonstração. (a) Seja μ um autovalor da matriz

$$J = D^{-1}(L + U) = B(1).$$

Por A ser uma matriz consistentemente ordenada, μ também é um autovalor de

$$-J = B(-1) = -D^{-1}(L + U).$$

Como μ é um autovalor de $-J$, então $-\mu$ é autovalor de J .

(b): Como $I - \omega L^*$ é uma matriz triangular inferior com diagonal 1, então $\det(I - \omega L^*) = 1$ para todo ω , logo

$$\begin{aligned} \det(\lambda I - H(\omega)) &= \det(I - \omega L^*) \det(\lambda I - H(\omega)) \\ &= \det((I - \omega L^*)(\lambda I - H(\omega))) \\ &= \det((I - \omega L^*)\lambda I - (I - \omega L^*)H(\omega)) \\ &= \det(\lambda I - \lambda \omega L^* - (1 - \omega)I - \omega U^*) \\ &= \det((\lambda + \omega - 1)I - \lambda \omega L^* - \omega U^*). \end{aligned}$$

Seja μ um autovalor de J , e λ uma solução de 4.13, então

$$\lambda + \omega - 1 = \sqrt{\lambda} \omega \mu, \text{ ou } \lambda + \omega - 1 = -\sqrt{\lambda} \omega \mu.$$

Por (a) assumimos sem perda de generalidade que

$$\lambda + \omega - 1 = \sqrt{\lambda} \omega \mu.$$

Se $\lambda = 0$ e $\omega = 1$, então

$$p(0) = \det(0 \cdot I - H(1)) = \det(-U^*) = 0,$$

logo 0 é um raiz do polinômio característico e portanto um autovalor $H(1)$. Se $\lambda \neq 0$, então

$$\det(\lambda I - H(\omega)) = \det \left[(\lambda + \omega - 1)I - \sqrt{\lambda} \omega \left(\sqrt{\lambda} L^* + \frac{1}{\sqrt{\lambda}} U^* \right) \right],$$

como $(\lambda + \omega - 1) = \sqrt{\lambda} \omega \mu$, então

$$\begin{aligned} \det(\lambda I - H(\omega)) &= \det \left(\sqrt{\lambda} \omega \mu I - \sqrt{\lambda} \omega B(\sqrt{\lambda}) \right) \\ &= \det \left(\sqrt{\lambda} \omega I \right) \cdot \det \left(\mu I - B(\sqrt{\lambda}) \right). \end{aligned}$$

Como $\sqrt{\lambda} \omega I$ é uma matriz diagonal, então $\det(\sqrt{\lambda} \omega I) = (\sqrt{\lambda} \omega)^n$, logo

$$\det(\lambda I - H(\omega)) = \left(\sqrt{\lambda} \omega \right)^n \det \left(\mu I - B(\sqrt{\lambda}) \right),$$

onde $B(\lambda) = \sqrt{\lambda}L^* + \frac{1}{\sqrt{\lambda}}U^*$. Como os autovalores de $B(\sqrt{\lambda})$ são os mesmo de $B(1) = L^* + U^* = J$, então

$$p(\lambda) = \left(\sqrt{\lambda}\omega\right)^n \det(\mu I - J) = 0,$$

pois por hipótese μ é um autovalor de J . Logo, λ definido por 4.13 é um autovalor de $H(\omega)$. \square

Resolvendo

$$(\lambda + \omega - 1)^2 = \lambda\omega^2\mu^2,$$

para λ , encontramos que :

$$\lambda = 1 - \omega + \frac{1}{2}\omega^2\mu^2 \pm \omega\mu\sqrt{1 - \omega + \frac{1}{4}\omega^2\mu^2}.$$

Corolário 4.6.6. *Seja A uma matriz consistentemente ordenada, então*

$$\rho(G) = [\rho(J)]^2.$$

Demonstração. Tomando $\omega = 1$, então $H(1) = G$, temos assim, pelo teorema 4.6.5 que $\lambda^2 = \lambda\mu^2$, dividindo ambos os lados por $\lambda \neq 0$, obtemos $\lambda = \mu^2$, portanto $\rho(G) = [\rho(J)]^2$. \square

Conforme [8], a taxa de convergência do método iterativo estacionário, com matriz de iteração T , é definida por

$$R = -\log(\rho(T)).$$

Consideremos as taxas de convergência do método de Jacobi, R_J , e a taxa de convergência do método de Gauss-Seidel, R_{GS} , então

$$R_{GS} = -\log \rho(G) = -\log \rho(J)^2 = -2 \log \rho(J) = 2R_J.$$

Portanto o método de Gauss-Seidel converge duas vezes mais rápido que o método de Jacobi.

Dado $0 < \omega < 2$ e μ um autovalor de J , então pelo teorema 4.6.5, os autovalores λ de $H(\omega)$ satisfazem

$$(\lambda + \omega - 1)^2 = \lambda\omega^2\mu^2. \quad (4.14)$$

podemos dividir a equação 4.14 por ω^2 , dessa forma obtemos

$$\frac{(\lambda + \omega - 1)^2}{\omega^2} = \lambda\mu^2,$$

retirando a raiz de ambos os lados da equação obtemos

$$\frac{\lambda + \omega - 1}{\omega} = \pm\sqrt{\lambda}\mu. \quad (4.15)$$

Tomando

$$g_\omega(\lambda) = \frac{\lambda + \omega - 1}{\omega}, \text{ e } m_i(\lambda) = \pm\sqrt{\lambda}\mu_i,$$

Dessa forma obtemos uma reta g_ω com inclinação $\frac{1}{\omega}$ que passa pelo ponto $(1, 1)$, e um conjunto de parábolas $\mu_i(\lambda) = \pm\sqrt{\lambda}\mu_i$ para cada autovalor λ_i . Observe que se μ_i é um autovalor de J , então $-\mu_i$ também é autovalor e a parábola associada à $-\mu_i$ será a mesma parábola m_i associado à μ_i .

Assim, dado um autovalor μ_i da matriz J , pela equação 4.15 deverá existir dois autovalores $\lambda_i^{(1)}$ e $\lambda_i^{(2)}$ associados à μ_i , além disso $\mu_j = -\mu_i$ possui os mesmos autovalores $\lambda_i^{(1)}$ e $\lambda_i^{(2)}$ associados.

Os autovalores λ_i associados à μ_i ocorrem quando $g_\omega(\lambda_i) = m_i(\lambda_i)$, ou seja, quando a reta g_ω intercepta a parábola μ_i .

Como $H(\omega)$ é uma função de ω , também são os seus autovalores, portanto $\rho(H(\omega))$ também depende de ω . Para encontrarmos o valor ótimo de ω que faça com que o método de SOR convirja o mais rápido possível, devemos encontrar um ω_{opt} que minimize $\rho(T_{SOR}(\omega))$.

Teorema 4.6.7. *Seja A uma matriz consistentemente ordenada, se J possui autovalores reais e $\rho(J) < 1$, então*

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - [\rho(J)]^2}}, \quad \rho(H(\omega_{opt})) = \omega_{opt} - 1.$$

A demonstração do teorema pode ser encontrado em [12].

Podemos ver que se $\omega > 1$ e $\rho(J) < 1$, então, pelo teorema 4.6.7, $\rho(H(\omega_{opt})) < 1$. Portanto, para matrizes consistentemente ordenadas, a convergência do método de Jacobi, implica a convergência do método de SOR para ω ótimo.

Exemplo 4.6.8. Dado a matriz A da forma

$$A = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix}. \quad (4.16)$$

A matriz A é consistentemente ordenada, e a matriz de iteração $J = D^{-1}(L + U)$ associada à matriz A é da forma

$$J = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix}.$$

Portanto,

$$\lambda I - J = \begin{bmatrix} \lambda & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \lambda & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \lambda \end{bmatrix},$$

logo o polinômio característico de J é

$$p(\lambda) = \det(\lambda I - J) = \lambda^3 - \frac{\lambda}{2},$$

cujas raízes são $\lambda_1 = 0$, $\lambda_2 = \frac{\sqrt{2}}{2}$ e $\lambda_3 = -\frac{\sqrt{2}}{2}$. Assim,

$$\rho(J) = \max_i |\lambda_i| = \frac{\sqrt{2}}{2}.$$

e portanto,

$$\rho(G) = [\rho(J)]^2 = \frac{1}{2}.$$

Aplicando o teorema 4.6.7, concluímos que o w_{opt} é

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \left(\frac{\sqrt{2}}{2}\right)^2}},$$

desenvolvendo o denominador da fração

$$1 + \sqrt{1 - \left(\frac{\sqrt{2}}{2}\right)^2} = 1 + \sqrt{1 - \frac{1}{2}} = 1 + \frac{1}{\sqrt{2}} = \frac{\sqrt{2} + 1}{\sqrt{2}},$$

e portanto

$$w_{opt} = \frac{2}{1 + \sqrt{1 - \left(\frac{\sqrt{2}}{2}\right)^2}} = \frac{2}{\frac{\sqrt{2} + 1}{\sqrt{2}}} = \frac{2\sqrt{2}}{\sqrt{2} + 1},$$

multiplicando a fração pelo conjugado obtemos,

$$\frac{2\sqrt{2}}{\sqrt{2} + 1} = \frac{2\sqrt{2}}{\sqrt{2} + 1} \cdot \frac{\sqrt{2} - 1}{\sqrt{2} - 1} = \frac{4 - 2\sqrt{2}}{2 - 1} = 4 - 2\sqrt{2}$$

e portanto o valor do w_{opt} é

$$w_{opt} = 4 - 2\sqrt{2} \approx 1,7157.$$

Portanto o raio espectral de $H(\omega)$ para w_{opt} é

$$\rho(H(\omega_{opt})) = \omega_{opt} - 1 = 3 - 2\sqrt{2}.$$

Portanto, para a matriz A , definida por 4.16, temos que $w_{opt} = 4 - 2\sqrt{2}$, e

$$\rho(J) = \frac{\sqrt{2}}{2}, \quad \rho(G) = \frac{1}{2}, \quad \rho(H(\omega_{opt})) = 3 - 2\sqrt{2}.$$

5. RESULTADOS NUMÉRICOS

A seguir, são apresentados problemas de valor de contorno unidimensionais e bidimensionais com soluções analíticas conhecidas, visando demonstrar que os métodos iterativos conseguem resolvê-los, como era esperado. Embora os sistemas lineares associados aos problemas de valor de contorno unidimensionais sejam tri-diagonais, permitindo o uso de métodos diretos para sua solução, aqui o foco foi dado ao uso de métodos iterativos, com o intuito de analisar qual deles apresenta o melhor desempenho.

5.1 CONCEITOS PRELIMINARES

Devido à estrutura esparsa das matrizes obtidas pelo método das diferenças finitas, se tentarmos armazenar todos os valores das matrizes, estaremos armazenado $(n - 1)^2$ valores para o problema unidimensional e $(n - 1)^2(m - 1)^2$ valores para o problema bidimensional. Porém, devido à estrutura do nosso problema, não será necessário armazenar todos os valores da matriz obtida pelo método de diferenças finitas.

Para calcular as aproximações da solução U_{ij} através dos métodos iterativos apresentados nesse trabalho, basta conhecermos os valores de U avaliado no ponto (x_i, y_j) e nos pontos vizinhos do ponto (x_i, y_j) . Dessa forma, se armazenarmos todos os pontos da malha, isso será o suficiente para aplicarmos os métodos iterativos.

Devido à nossa malha de pontos ser uma malha estruturada, onde as conexões entre os pontos possuem uma estrutura simples, podemos armazenar todos os valores U_i em um único vetor, para o caso unidimensional, e os valores de U_{ij} em uma matriz, para o caso bidimensional.

Para o problema unidimensional $u''(x) = f(x)$, iremos armazenar os valores de U_i , para $i = 0, 1, 2, \dots, n - 1, n$, na forma de um vetor:

$$U = [U_0 \ U_1 \ U_2 \ U_3 \ \cdots \ U_{n-2} \ U_{n-1} \ U_n] \quad (5.1)$$

Observe que no vetor U , estamos armazenando, além dos valores que queremos calcular, os valores da borda, U_0 e U_n . Isso é necessário pois para calcular os valores de U próximos à borda, U_1 e U_{n-1} precisaremos usar os valores de U avaliado na borda.

Para o problema bidimensional, basta armazenar os valores U_{ij} em uma matriz, de forma que os pontos vizinhos da malha, estejam em posições adjacentes na matriz. Assim, a matriz

ficará da forma

$$U = \begin{bmatrix} U_{0,0} & U_{1,0} & U_{2,0} & U_{3,0} & \cdots & U_{n-1,0} & U_{n,0} \\ U_{0,1} & U_{1,1} & U_{2,1} & U_{3,1} & \cdots & U_{n-1,1} & U_{n,1} \\ U_{0,2} & U_{1,2} & U_{2,2} & U_{3,2} & \cdots & U_{n-1,2} & U_{n,2} \\ U_{0,3} & U_{1,3} & U_{2,3} & U_{3,3} & \cdots & U_{n-1,3} & U_{n,3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ U_{0,m-1} & U_{1,m-1} & U_{2,m-1} & U_{3,m-1} & \cdots & U_{n-1,m-1} & U_{n,m-1} \\ U_{0,m} & U_{1,m} & U_{2,m} & U_{3,m} & \cdots & U_{n-1,m} & U_{n,m} \end{bmatrix}. \quad (5.2)$$

Novamente, devemos incluir os valores de U avaliados na borda da malha, pois eles serão usados para o cálculo dos valores U_{ij} avaliados nos pontos internos.

A matriz U definida por 5.1 é constituída por apenas $(n+1)$ elementos, em comparação com os $(n-1)^2$ da matriz A obtida pelo método das diferenças finitas. Da mesma forma, a matriz U definida por 5.2, armazena $(n+1)(m+1)$ elementos, enquanto que a matriz do método das diferenças finitas possui $(n-1)^2(m-1)^2$ elementos.

Para conseguirmos medir o erro do nosso problema, será necessário o uso de normas. Usaremos as seguintes normas de vetor, conforme [9]:

$$\|\mathbf{e}_h\|_1 = \sum_{i=0}^n |E_i| h, \quad \|\mathbf{e}_h\|_2 = \left(\sum_{i=0}^n |E_i|^2 h \right)^{1/2}, \quad \|\mathbf{e}_h\|_\infty = \max_{0 \leq i \leq n} |E_i|.$$

Para analisarmos o erro \mathbf{e} da solução do problema, observemos que se o erro é de ordem p , ou seja $\|\mathbf{e}_h\| \approx Ch^p$, então

$$\log(\|\mathbf{e}_h\|) \approx \log(C) + p \log(h),$$

portanto, o gráfico em escala log-log do erro $\|\mathbf{e}_h\|$ em função do espaçamento h , deverá ser uma função afim com inclinação p . Assim, introduziremos os quocientes das normas dos erros associados à h e $h/2$:

$$q_1 = \frac{\|\mathbf{e}_h\|_1}{\|\mathbf{e}_{h/2}\|_1}, \quad q_2 = \frac{\|\mathbf{e}_h\|_2}{\|\mathbf{e}_{h/2}\|_2}, \quad q_\infty = \frac{\|\mathbf{e}_h\|_\infty}{\|\mathbf{e}_{h/2}\|_\infty}.$$

Devido à $\|\mathbf{e}_h\| \approx Ch^p$, os quocientes apresentados devem se aproximar de $4 = 2^2$.

5.2 PROBLEMA UNIDIMENSIONAL

Consideremos o seguinte problema de valor de contorno com condições de tipo Dirichlet

$$\begin{cases} u''(x) = e^x, & \text{para } x \in (0, 1), \\ u(0) = 1 \quad \text{e} \quad u(1) = e, \end{cases}$$

A solução exata do problema é $u(x) = e^x$.

Faremos uma comparação entre os 3 métodos iterativos apresentados: método de Jacobi, método de Gauss-Seidel, e método de Sobre-relação Sucessiva. O vetor nulo será utilizado como aproximação inicial para o método iterativo. Utilizaremos como critério de parada para o método iterativo, que o erro entre as sucessivas iterações satisfaça $\|\mathbf{U}^{(k+1)} - \mathbf{U}^{(k)}\|_\infty < 10^{-6}$.

De acordo com [15], os autovalores da matriz de iteração do método de Jacobi para o problema unidimensional são da forma

$$\lambda_i = -\cos\left(\frac{\pi i}{n}\right), \text{ para } i = 1, 2, \dots, n - 1.$$

Assim, $\rho(J) = |\lambda_1| = \cos\left(\frac{\pi}{n}\right)$, e portanto o ω ótimo para o método de SOR é

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \cos^2\left(\frac{\pi}{n}\right)}} = \frac{2}{1 + \text{sen}\left(\frac{\pi}{n}\right)}.$$

Assim, comparando os métodos de Jacobi, Gauss-Seidel, e SOR conforme a tabela 5.1, podemos ver que o método de Gauss-Seidel converge quase 2 vezes mais rápido que o método de Jacobi, e que o método de SOR com valor ótimo ω_{opt} , supera ambos os métodos de Jacobi, e Gauss-Seidel.

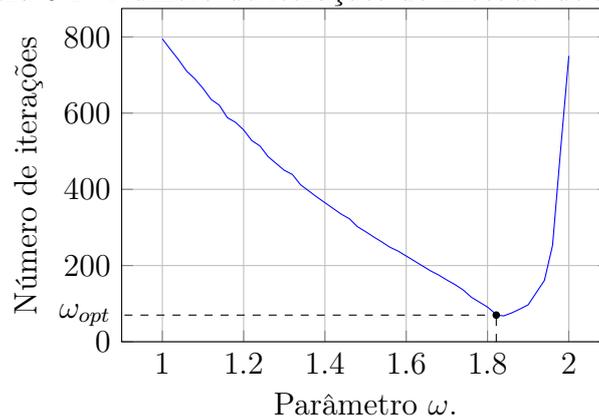
Tabela 5.1: Número de iterações dos métodos.

n	Jacobi	Gauss-Seidel	SOR com ω_{opt}
4	33	19	10
8	125	67	20
16	431	235	38
32	1443	795	70
64	4623	2601	133

Fonte: O autor.

Resolvendo o problema para diferentes valores de ω usando $n = 32$, podemos ver pela Figura 5.1, que o número de iterações necessárias para a convergência do método muda de acordo com o parâmetro de ω . Para $n = 32$, temos que $\omega_{opt} \approx 1.821465$. Conforme podemos ver na Figura 5.1, o valor ótimo ω_{opt} é o onde ocorre o menor número de iterações.

Figura 5.1: Número de iterações do método de SOR.



Fonte: O autor.

Para analisarmos a ordem do método das diferenças finitas, consideremos o seguinte problema de valor de contorno com condições de tipo Dirichlet,

$$\begin{cases} u''(x) = -\pi^2 \operatorname{sen}(\pi x), & \text{para } x \in (-1, 1), \\ u(-1) = 0 \quad \text{e} \quad u(1) = 0. \end{cases}$$

A solução exata do problema é $u(x) = \operatorname{sen}(\pi x)$.

De acordo com vimos na teoria apresentada no segundo capítulo,

$$\frac{u(x-h) - 2u(x) + u(x+h)}{h^2} = u''(x) + O(h^2),$$

dessa forma, o erro de aproximação da solução numérica deverá ser de segunda ordem. E portanto, espera-se que quando reduzimos o espaçamento h pela metade, o erro caia aproximadamente $4 = 2^2$ vezes.

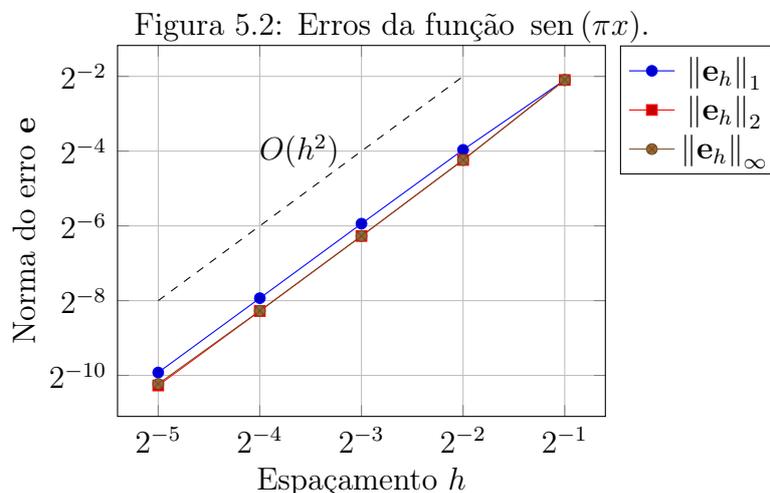
Para a resolução do problema, foi utilizado o método de SOR com o valor ótimo ω_{opt} , utilizando como critério de parada para o método iterativo, o erro $\|\mathbf{U}^{(k+1)} - \mathbf{U}^{(k)}\|_\infty < 10^{-6}$ entre sucessivas aproximações. Usamos como a aproximação inicial, $\mathbf{U}^{(0)}$, o vetor nulo $\mathbf{U}^{(0)} = [0, 0, \dots, 0]^t$.

Podemos ver pela tabela 5.2 que os quocientes do erro q_1 , q_2 e q_∞ se aproximam de $2^2 = 4$, indicando assim que o método é de segunda ordem. Também podemos ver pela inclinação da reta do gráfico 5.2, que o erro é de segunda ordem.

Tabela 5.2: Erro da função $\operatorname{sen}(\pi x)$.

n	h	$\ \mathbf{e}_h\ _1$	$\ \mathbf{e}_h\ _2$	$\ \mathbf{e}_h\ _\infty$	q_1	q_2	q_∞
4	0.5	0.233701	0.233701	0.233701			
8	0.25	0.064013	0.053030	0.053032	3.650813	4.406974	4.406790
16	0.125	0.016280	0.012952	0.012958	3.932130	4.942349	4.940123
32	0.0625	0.004090	0.003222	0.003234	3.980420	5.053057	5.033307
64	0.031250	0.001030	0.000810	0.000835	3.969073	5.050287	4.900005

Fonte: O autor.



Consideremos o problema de condição de contorno de Dirichlet definido por

$$\begin{cases} u''(x) = -2, & \text{para } x \in (-1, 1), \\ u(-1) = 0 \quad \text{e} \quad u(1) = 0. \end{cases}$$

A solução analítica do problema é dada por $u(x) = 1 - x^2$.

Conforme vimos na teoria, pela expansão em polinômio Taylor, o erro de truncamento local do método é

$$\frac{u(x-h) - 2u(x) + u(x+h)}{h^2} - u''(x) = \frac{h^2}{12}u^{(4)}(\xi).$$

Como $u^{(4)} = 0$, então $\frac{h^2}{12}u^{(4)}(\xi) = 0$, dessa forma, o erro de truncamento local será $\tau_h u = 0$. Como o erro de truncamento local é zero, espera-se que o erro da solução numérica obtida também se aproxime de zero, e que os erros encontrados sejam consequência dos erros de arredondamento e dos erros de aproximação do método iterativo.

Para a resolução numérica do problema usamos novamente o método de SOR com valor ótimo ω_{opt} , e com critério de parada que $\|\mathbf{U}^{(k+1)} - \mathbf{U}^{(k)}\|_\infty < 10^{-6}$, conforme a tabela 5.3 e com o critério $\|\mathbf{U}^{(k+1)} - \mathbf{U}^{(k)}\|_\infty < 10^{-10}$, conforme a tabela 5.4.

Podemos ver pelas tabelas 5.3 e 5.4, que o erro da aproximação se aproxima de 0, conforme esperado da teoria. O erro apresentado pela tabela, pode ser interpretado como sendo os erros de arredondamentos acumulados. Conforme o número de pontos da malha aumentam, o número de operações necessárias para calcular os valores U_i também aumentam, dessa forma, espera-se que os erros numéricos de arredondamentos ao longo do processo iterativo cresçam.

Tabela 5.3: Erro da função $1 - x^2$, quando $< 10^{-6}$.

n	h	$\ \mathbf{e}_h\ _1$	$\ \mathbf{e}_h\ _2$	$\ \mathbf{e}_h\ _\infty$
4	0.5	7.450581×10^{-7}	6.322027×10^{-7}	5.960464×10^{-7}
8	0.25	4.740472×10^{-6}	3.805325×10^{-6}	3.841594×10^{-6}
16	0.125	1.261314×10^{-5}	1.000687×10^{-5}	1.011076×10^{-5}
32	0.0625	2.542740×10^{-5}	2.011592×10^{-5}	2.036017×10^{-5}
64	0.03125	4.913274×10^{-5}	3.884201×10^{-5}	3.932360×10^{-5}

Fonte: O autor.

Tabela 5.4: Erro da função $1 - x^2$, quando $< 10^{-10}$.

n	h	$\ \mathbf{e}_h\ _1$	$\ \mathbf{e}_h\ _2$	$\ \mathbf{e}_h\ _\infty$
4	0.5	9.094947×10^{-11}	7.717318×10^{-11}	7.275958×10^{-11}
8	0.25	5.122295×10^{-10}	4.111827×10^{-10}	4.151016×10^{-10}
16	0.125	1.168212×10^{-9}	9.268228×10^{-10}	9.364447×10^{-10}
32	0.0625	2.534509×10^{-9}	2.005080×10^{-9}	2.029426×10^{-9}
64	0.03125	4.857011×10^{-9}	3.839721×10^{-9}	3.887329×10^{-9}

Fonte: O autor.

5.3 PROBLEMA DE BIDIMENSIONAL

Nesta seção apresentaremos alguns resultados numéricos em relação ao método de diferenças finitas para a equação de Poisson bidimensional. Para podermos aplicar o método de SOR com ω_{opt} precisamos saber o raio espectral do método de Jacobi. De acordo com [12], o raio espectral do método de Jacobi, para $n = m$ é $\rho(J) = \cos\left(\frac{\pi}{n}\right)$, logo

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \cos^2\left(\frac{\pi}{n}\right)}} = \frac{2}{1 + \operatorname{sen}\left(\frac{\pi}{n}\right)}.$$

Consideremos o problema de condição de contorno de Dirichlet, em um domínio $\Omega = [-1, 1] \times [-1, 1]$, definido por

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = e^{xy}(x^2 + y^2), & \text{para } (x, y) \in \Omega, \\ u(x, -1) = e^{-x} & \text{e } u(y, 1) = e^x, & \text{para } x \in [-1, 1], \\ u(-1, y) = e^{-y} & \text{e } u(1, y) = e^y, & \text{para } y \in [-1, 1]. \end{cases} \quad (5.3)$$

A solução analítica do problema é dada por $u(x, y) = e^{xy}$.

Resolvendo o problema 5.3 usando como critério de parada, o erro $\|\mathbf{U}^{(k+1)} - \mathbf{U}^{(k)}\| < 10^{-6}$, podemos comparar o número de iterações necessário para cada método, conforme apresentado na tabela 5.5.

Vemos novamente que o método de Gauss-Seidel converge aproximadamente 2 vezes mais rápido que o método de Jacobi, e que o método de SOR com ω_{opt} foi o método que convergiu mais rápido.

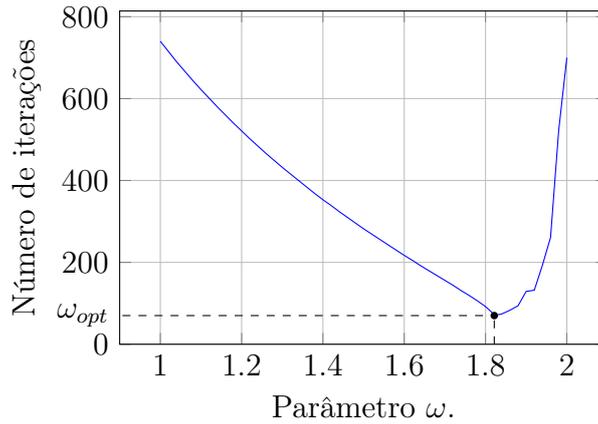
Aplicando o método de SOR para o problema 5.3 mantendo fixo o número de partições $n = 32$ e $m = 32$, e o coeficiente ω variando, obtemos o número de iterações necessárias para cumprir o critério de parada $\|\mathbf{U}^{(k+1)} - \mathbf{U}^{(k)}\| < 10^{-6}$, conforme o gráfico 5.3. Novamente, podemos ver que o $\omega_{opt} \approx 1.821465$ nos proporciona o menor número de iterações possíveis para a convergência do método de Sobre-relação Sucessiva.

Tabela 5.5: Número de iterações dos métodos.

$n \times m$	Jacobi	Gauss-Seidel	SOR com ω_{opt}
4 x 4	30	17	10
8 x 8	117	64	20
16 x 16	403	221	38
32 x 32	1334	740	70
64 x 64	4190	2384	139

Fonte: O autor.

Figura 5.3: Número de iterações do método de SOR.



Fonte: O autor.

Vimos na seção anterior que o erro do método das diferenças finitas centradas para o problema unidimensional é de segunda ordem. Veremos que para o problema de Poisson bidimensional que o erro também é de segunda ordem. Para isso, consideremos como exemplo, o problema de valor de contorno de tipo Dirichlet, em um domínio $\Omega = [-1, 1] \times [-1, 1]$, definido

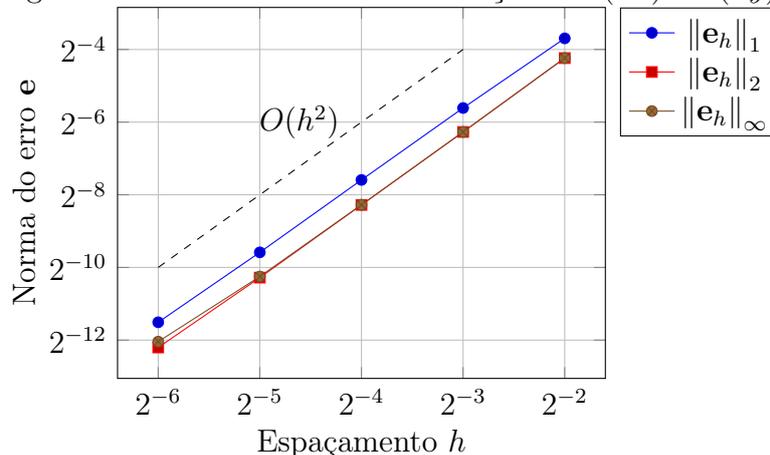
$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -2\pi^2 \text{sen}(\pi x) \text{sen}(\pi y), & \text{para } (x, y) \in \Omega, \\ u(x, y) = 0, & \text{para } (x, y) \in \partial\Omega. \end{cases} \quad (5.4)$$

A solução analítica do problema é dada por $u(x, y) = \text{sen}(\pi x) \text{sen}(\pi y)$.

Resolvendo o problema 5.4 pelo método de SOR, usando ω_{opt} para o caso em que $n = m$. Consideremos como critério de parada o $\|\mathbf{U}^{(k+1)} - \mathbf{U}^{(k)}\|_\infty < 10^{-6}$ e utilizando como aproximação inicial $U_{ij}^{(0)} = 0$ para todo os pontos da malha.

Podemos ver pela tabela 5.6, e pelo gráfico 5.4 que o erro decai, aproximadamente, de forma quadrática, confirmando o resultado obtido na teoria.

Figura 5.4: Gráfico dos erros da função $\text{sen}(\pi x) \text{sen}(\pi y)$.



Fonte: O autor.

Tabela 5.6: Erro da função $\text{sen}(\pi x) \text{sen}(\pi y)$.

n	h	$\ \mathbf{e}_h\ _1$	$\ \mathbf{e}_h\ _2$	$\ \mathbf{e}_h\ _\infty$	q_1	q_2	q_∞
8	0.25	0.077271	0.053029	0.053031			
16	0.125	0.020460	0.012951	0.012955	3.776772	4.094743	4.093421
32	0.0625	0.005187	0.003219	0.003227	3.944513	4.023634	4.012966
64	0.03125	0.001302	0.000804	0.000821	3.982447	4.005349	3.921153
128	0.015625	0.000343	0.000212	0.000237	3.792553	3.787424	3.386068

Fonte: O autor.

Nos exemplos anteriores, vimos que o uso do método de Sobre-relaxação Sucessiva com ω_{opt} apresentou o menor número de iterações, considerando, em todos os casos, o vetor nulo como aproximação inicial. No entanto, se a aproximação inicial estiver mais próximo da solução esperada, é possível que o número de iterações seja ainda menor. Para demonstrar isso, analisaremos o seguinte problema

Consideremos o problema de de valor de contorno com condição de contorno de tipo Dirichlet, um domínio $\Omega = [-1, 1] \times [-1, 1]$, definido por

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -\pi^2 e^{-t} (\cos(\pi x) + \cos(\pi y)), & \text{para } (x, y) \in \Omega, \\ u(x, -1) = u(x, 1) = e^{-t} (\cos(\pi x) - 1), & \text{para } x \in [-1, 1], \\ u(-1, y) = u(1, y) = e^{-t} (\cos(\pi y) - 1), & \text{para } y \in [-1, 1]. \end{cases} \quad (5.5)$$

para tempo fixo $t \geq 0$. A solução analítica do problema é $u(x, y) = e^{-t} (\cos(\pi x) + \cos(\pi y))$.

Vamos comparar os métodos usando uma aproximação inicial próximo da solução exata, $u(x, y) = e^{-t_0} (\cos(\pi x) + \cos(\pi y))$, onde $t_0 = t - 0.01$, com uma aproximação inicial $U_{ij}^{(0)} = 0$ para todos os pontos internos da malha. Para comparar, resolveremos o problema 5.5 para $t = 1$ usando o critério de parada $\|\mathbf{U}^{(k+1)} - \mathbf{U}^{(k)}\| < 10^{-6}$, utilizaremos o método de Gauss-Seidel e o método de SOR com parâmetro ótimo,

Tabela 5.7: Número de iterações dos métodos.

$n \times m$	GS, início 0	GS, início em t_0	SOR, início 0	SOR, início t_0
8 x 8	101	85	29	24
16 x 16	396	262	54	44
32 x 23	1334	750	109	81
63 x 64	4760	2779	223	149
128 x 128	57753	27141	456	286

Fonte: O autor.

Podemos comparar o número de iterações necessário para cada método conforme a tabela 5.5. Vemos que em ambos os casos, a solução próxima da solução exata está convergindo mais rápido. Podemos ver também que tanto para o método de Gauss-Seidel, quanto para o SOR, conforme aumentamos o número de pontos da malha, a diferença da velocidade de convergência entre as duas aproximações iniciais também aumenta.

6. CONCLUSÕES

Os problemas de valores de contorno estão presentes em diversas áreas da matemática, ciências exatas e engenharias. Neste trabalho apresentamos uma introdução às equações diferenciais ordinárias e às equações diferenciais parciais, vimos problemas de valores de contorno, dando um foco para os problemas de tipo Dirichlet.

Vimos o método das diferenças finitas, que consiste em discretizar o domínio do problema por uma malha de pontos. Uma vez discretizado o domínio, podemos substituir, em cada ponto da malha, as derivadas por esquemas de diferenças finitas, resultando em um sistema de equações lineares, os quais podem ser resolvidos por métodos numéricos.

Além da introdução do método das diferenças finitas, utilizamos o Teorema de Taylor para fazer um estudo sobre a ordem do erro de aproximação dos esquemas de diferenças finitas. Concluimos que as diferenças finitas centradas, apresentadas neste trabalho, para a primeira e segunda derivada são de segunda ordem.

Foram introduzidos os conceitos de erro global e erro de truncamento local dos métodos de diferenças finitas, o que possibilitou fazer um estudo sobre a convergência, consistência e estabilidade dos métodos para os problemas de valor de contorno unidimensional e bidimensional. No caso bidimensional, vimos ainda o princípio do máximo contínuo e discreto, que nos possibilitou encontrar um limitante para o erro global do método.

Para resolver os sistemas lineares obtidos através do método de diferenças finitas, introduzimos os métodos iterativos estacionários, em especial, os métodos de Jacobi, Gauss-Seidel e Sobre-relação Sucessiva. Vimos os critérios de convergência para os métodos iterativos apresentados e vimos que os três métodos convergem para o problema unidimensional apresentado e para o problema de Poisson bidimensional.

Vimos que podemos analisar algumas propriedades das matrizes por meio de um grafo associado a ela, em especial, o grafo associado possibilitou a identificação das matrizes consistentemente ordenadas. Concluimos que as matrizes obtidas pelo método das diferenças finitas eram consistentemente ordenadas e concluimos várias propriedades a respeito dos métodos iterativos, em especial, encontramos o parâmetro ótimo para o método de Sobre-relação Sucessiva.

A partir dos resultados numéricos obtidos, confirmamos a coerência e a consistência da teoria apresentada, validando a aplicação das abordagens discutidas. Confirmamos que os métodos de diferenças centradas são de segunda ordem, analisamos a convergência dos três métodos iterativos e concluimos que o melhor cenário para a convergência do método de Sobre-relação Sucessiva é, de fato, o parâmetro ótimo encontrado na teoria.

Para trabalhos futuros, podemos propor uma análise da convergência do método das diferenças finitas para os problemas de valor de contorno quando sujeitos a condições de tipo de Neumann. Também propõe-se um estudo mais aprofundado de como o método se comporta usando malhas não uniformes, ou usando a ordenação *red-black*.

De maneira geral, com este estudo foi possível fazer uma conexão da aplicação de alguns métodos numéricos utilizados para resolver problemas de valor de contorno, com a teoria que se encontra por trás dos métodos numéricos utilizados.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Ackleh, A. S., e. a.: *Classical and Modern Numerical Analysis: Theory, Methods and Practice*. Chapman and Hall/CRC, 1ª ed., 2009.
- [2] Boyce, W. E. e Diprima, R. C.: *Equações Diferenciais Elementares e Problemas de Valores de Contorno*. LTC, 8ª ed., 2005.
- [3] Cuminato, J. A. e Junior, M. M.: *Discretização de Equações Diferenciais Parciais*. Coleção Matemática Aplicada. Sociedade Brasileira de Matemática, 2013.
- [4] Demmel, J. W.: *Applied Numerical Linear Algebra*. Society for Industrial and Applied Mathematics, 1ª ed., 1997.
- [5] Figueiredo, D. G.: *Análise de Fourier e Equações Diferenciais Parciais*. SBM, Rio de Janeiro, 2008.
- [6] Franco, N. M. B.: *Cálculo Numérico*. Pearson Universidades, 2006.
- [7] Guidorizzi, H. L.: *Um Curso de Cálculo, Vol 2*. Livros Técnicos e Científicos, 5ª ed., 1989.
- [8] Isaacson, E., H. B. K.: *Analysis of Numerical Methods*. Dover Publications, 1994.
- [9] LeVeque, R. J.: *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady State and Time Dependent Problems*. Texts in applied mathematics. Society for Industrial and Applied Mathematics, Philadelphia, 2007.
- [10] Morton, K. W.: *Numerical Solution of Partial Differential Equations: An Introduction*. Cambridge University Press, 2005.
- [11] Smith, G. D.: *Numerical Solution of Partial Differential Equations: Finite Difference Methods*. Oxford Applied Mathematics and Computing Science Series. Clarendon Press, 3ª ed., 1986.
- [12] Stoer, J. e Bulirsch, R.: *Introduction to numerical analysis*. Texts in applied mathematics. Springer, 2002.
- [13] Strikwerda, J. C.: *Finite Difference Schemes and Partial Differential Equations*. Society for Industrial and Applied Mathematics, 2ª ed., 2004.

- [14] Suli E., M.D.: *An introduction to numerical analysis*. Cambridge, 2003.
- [15] Zhilin, J. e Bulirsch, R.: *Numerical Solution of Differential Equations: Introduction to Finite Difference and Finite Element Methods*. Texts in applied mathematics. Cambridge University Press, 2018.