

UNIVERSIDADE FEDERAL  
DE ALFENAS

NATALLY RODRIGUES SILVA

UMA INTRODUÇÃO À PROGRAMAÇÃO NÃO  
LINEAR

Alfenas  
2018

Natally Rodrigues Silva

*Uma introdução à programação não linear*

Trabalho de Conclusão de Curso apresentado à Universidade Federal de Alfenas, como parte dos requisitos para obtenção do título de Licenciado em Matemática.

Orientador  
Prof. Dr. José Claudinei Ferreira

UNIVERSIDADE FEDERAL DE ALFENAS  
Alfenas  
2018

Natally Rodrigues Silva

# Uma Introdução à Programação Não Linear

A banca examinadora abaixo-assinada, aprova o Trabalho de Conclusão de Curso apresentado como parte dos requisitos para obtenção do Certificado de Conclusão do Curso de Licenciatura em Matemática pela Universidade Federal de Alfenas.

Aprovado em: \_\_\_\_ de \_\_\_\_\_ de 2018.

---

Prof. Dr. José Claudinei Ferreira  
Orientador

---

Prof. Dr. Anderson José Oliveira  
Universidade Federal de Alfenas

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Andréa Cardoso  
Universidade Federal de Alfenas

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Angela Leite Moreno (Suplente)  
Universidade Federal de Alfenas

*Agradeço à Deus por guiar meus caminhos.  
À minha família pelo incentivo, amor e carinho.  
Ao meu orientador pelo apoio.  
Aos meus amigos pela convivência, apoio e atenção.*

*“Nunca deixe que lhe digam que não vale a pena  
Acreditar no sonho que se tem  
Ou que seus planos nunca vão dar certo  
Ou que você nunca vai ser alguém.”*

Renato Russo

## Resumo

Os problemas reais tratados por diferentes profissionais podem estar relacionados com o problema de determinação de pontos de máximo ou de mínimo de funções de várias variáveis. Para lidar com esse tipo de problema surgiu a área de Pesquisa Operacional (PO). No Brasil, essa área surgiu no final da década de 50, do século passado, com alguns trabalhos pioneiros. Provavelmente os problemas de PO e particularmente de Problemas de Programação Não Linear começaram a ser trabalhados apenas no século passado e são temas recentes devido à dificuldade computacional. O propósito desta monografia é estudar algumas formas de obtermos aproximações de soluções de problemas de otimização não lineares. Estudamos e compreendemos modelos de resolução de problemas de programação não linear sem utilizar softwares prontos para essa finalidade. Para isso, utilizaremos métodos Numéricos implementados na linguagem de programação R. Damos destaque para métodos relacionados com a localização de raízes do gradiente da função, com o método dos gradientes e com métodos Numéricos para equações diferenciais ordinárias. Utilizamos alguns desses métodos para a determinação de autovalores e autovetores de transformações lineares, que são úteis na análise de problemas de classificação de pontos críticos funções.

**Palavras-chave:** Otimização Não Linear. Autovalores e Autovetores. Métodos Numéricos de Otimização. Linguagem de Programação.

## Abstract

There are many real problems related to optimization problems of functions of several variables. This kind of problems motivate Operational Research Theory to emerge in the late 50's of the last century, here in Brazil, with some pioneering works. Probably Operational Research problems and particularly Nonlinear Programming Problems started to be worked only in the last century due to its computational complexity and the beginning of computation theory. The aim of this work is to introduce some basic concepts of Nonlinear Programming Problems. To this end, we use numerical methods and the language R to construct and implement algorithms. We focus on Newton's methods and numerical methods to ordinary differential equations, including some gradient based methods. Particularly, we use some of these methods to determine eigenvalues and eigenvectors of linear transformations.

**Keywords:** Nonlinear Optimization. Eigenvalues and Eigenvectors. Numerical Methods of Optimization. Programming language.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>7</b>
<b>2</b>	<b>Máximos e mínimos de funções</b>	<b>10</b>
2.1	Máximos e Mínimos . . . . .	10
2.2	Autovalores e Autovetores . . . . .	17
<b>3</b>	<b>Métodos numéricos para equações não lineares</b>	<b>21</b>
3.1	Método de Newton . . . . .	21
3.1.1	Análise de Convergência do Método de Newton . . . . .	23
3.2	Implementação do Método do Tipo Newton . . . . .	27
3.2.1	Método da Eliminação de Gauss . . . . .	28
3.2.2	Decomposição LU . . . . .	31
<b>4</b>	<b>Problemas de otimização</b>	<b>46</b>
4.1	Multiplicadores de Lagrange . . . . .	47
4.1.1	Método dos Multiplicadores de Lagrange . . . . .	48
4.2	Condições de Karush-Kuhn-Tucker (KKT) . . . . .	50
4.3	Métodos Tipo Gradiente . . . . .	52
4.3.1	Introdução a Equações Diferenciais Ordinárias . . . . .	53
4.4	Métodos Numéricos para PVI . . . . .	54
4.4.1	Método de Taylor . . . . .	55
4.4.2	Método de Euler . . . . .	56
4.4.3	Método de Runge-Kutta . . . . .	57
4.5	Método de Declive . . . . .	59
4.6	Método do M-Grande . . . . .	61
<b>5</b>	<b>Considerações finais</b>	<b>66</b>
	<b>Referências Bibliográficas</b>	<b>67</b>
<b>A</b>	<b>APÊNDICE - Algoritmos em linguagem R</b>	<b>69</b>
A.1	Método de Newton do Exemplo (3.1.7) . . . . .	69



A.2	Cálculo de autovalores por Decomposição LU do Exemplo (3.2.4)	69
A.3	Algoritmo de Eliminação de Gauss	70
A.4	Algoritmo de Eliminação de Gauss com Pivotamento Parcial	72
A.5	Método de Newton do Exemplo (3.2.6)	73
A.6	Método de Newton do Exemplo (3.2.7)	75
A.7	Runge-Kutta do Exemplo (4.4.2)	76
A.8	Método de Euler do Exemplo (4.6.2)	77
A.9	Método de Euler do Exemplo (4.6.3)	78

# Capítulo 1

## Introdução

Em um curso inicial de Cálculo Diferencial e Integral, é muito comum termos como motivação para os estudos a resolução de problemas de otimização de funções, que surgem de problemas reais adaptados ao contexto didático. A resolução desses problemas resume-se a encontrar as raízes das derivadas das funções envolvidas. Um exemplo bem simples de aplicações em negócios e economia pode ser encontrado em [1, p. 315]. Outro exemplo bastante útil é encontrado em [2, p. 359] e trata da dedução de fórmulas de regressão ou mínimos quadrados, que é usada para ajuste de funções a conjuntos de dados em várias áreas do conhecimento.

Por outro lado, os problemas reais tratados por diferentes profissionais podem envolver funções bem mais complicadas e, por exemplo, o problema de encontrar as raízes do gradiente destas funções, pode ser extremamente complicado e exigir métodos Numéricos sofisticados, instáveis e lentos. Mesmo que consigamos encontrar essas raízes, ainda precisamos decidir se estas determinam pontos de máximos ou mínimos da função, ou se são apenas pontos de inflexão ou de sela. Uma forma de tomar essa decisão envolve conhecimento dos autovalores da matriz associada ao problema, chamada de Matriz Hessiana. Entretanto, o cálculo de autovalores de uma matriz pode ser visto como um problema de maximização (veja [3, p. 207]) que pode ser relacionado com o método dos multiplicadores de Lagrange ([2, p. 364]) que também envolve a determinação de raízes de equações em várias variáveis (veja [4] e referências lá citadas para as discussões sobre as dificuldades na resolução de sistemas de equações). Assim, sob certos aspectos, podemos ver a Teoria Espectral, que trata de assuntos ligados a autovalores e a valores singulares, como um problema de otimização e assim incluí-la isso na Pesquisa Operacional (PO). O artigo [5] mostra avanços nessa linha, ligados ao Google e ao algoritmo PageRank.

O termo Pesquisa Operacional segundo [6], apareceu no Brasil no final da década de 50 do século passado com alguns trabalhos pioneiros, dentre eles o trabalho do engenheiro Lóss. [7] relata que segundo Lóss, desde o século III a.C. utilizava-se PO, e que o início formal da PO foi através dos militares durante a Segunda Guerra Mundial, para fins de decisão de pontos estratégicos de deslocamento de pessoal e suprimentos, chegando ao

Brasil uma década após sua implementação na Grã-Bretanha.

Segundo [6] e [7] a PO foi implantada no curso de Engenharia em 1957 na USP e após dois anos no ITA e na década de 60 teve-se conhecimento das primeiras aplicações de PO a problemas reais no Brasil e em 1968, houve o I Simpósio Brasileiro de Pesquisa Operacional, que ocorre anualmente desde então.

O primeiro trabalho formal que se tem registro em otimização não linear são as *Condições de Karush-Kuhn-Tucker (KKT)* que determina condições necessárias para que uma solução de um Problema de Programação Não Linear (PPNL) seja ótima. Essas condições generalizam o método dos Multiplicadores de Lagrange, permitindo restrições com desigualdades e foram nomeadas originalmente após as publicações de Harold W. Kuhn e Albert W. Tucker em 1951. Mais tarde, estudiosos descobriram que essas condições já tinham sido propostas, de forma parecida, por William Karush em sua dissertação em 1939 [8]. Provavelmente os problemas de PO, e particularmente de PPNL, começaram a ser trabalhados apenas no século passado e são temas de pesquisas recentes em virtude da dificuldade computacional e do avanço da tecnologia.

Como a maioria dos problemas reais são não lineares e bastante complexos, podemos utilizar a PO para analisá-los através das técnicas existentes para PPNL, que são problemas de otimização em que a função objetivo e/ou pelo menos uma das restrições envolvidas não são funções lineares das variáveis de decisão [9]. Primeiramente procuramos determinar a região viável e depois identificar os extremos da região, ou seja, os pontos onde a função-objetivo apresenta maior ou menor valor. Porém, a fronteira da região de um PPNL pode não ser uma união de poliedros  $n$ -dimensionais, o que faz com que a solução ótima possa estar em qualquer lugar da região viável, ou seja, pode estar em algum extremo, na fronteira, ou em algum ponto interior da região viável e, assim, para descobrir qual o maior ou menor valor possível para a função objetivo, é preciso pesquisar dentro do conjunto de soluções viáveis seus possíveis valores. Como há uma grande dificuldade em encontrar a solução ótima destes problemas, é possível utilizar, para alguns casos, *softwares* como ferramentas, tais como o Solver do Excel e o Lindo Systems, além de outros, ou o uso de métodos Numéricos com o auxílio de uma linguagem de programação de computadores. Outro fator que dificulta ainda mais sabermos qual é a solução ótima dos PPNLs são os pontos de máximos e de mínimos locais da função objetivo que podem atrapalhar alguns métodos, como os métodos relacionados com o método dos gradientes, que estão relacionados com a resolução numérica de equações diferenciais, que podem ser afetados pela instabilidade dos pontos críticos da função objetivo, que por sua vez depende do sinal dos autovalores da Matriz Hessiana associada. Veja, por exemplo, [10] para novos métodos relacionados com o método dos gradientes e envolvendo derivadas fracionárias em problemas de otimização.

O propósito deste estudo bibliográfico foi analisar algumas formas de obtermos boas aproximações de solução de problemas de otimização não lineares. A partir das equações

do modelo matemático de um problema de otimização não linear, escolher um método para fazer a aproximação da solução por meio de algoritmos e métodos Numéricos ou analíticos. Para isso, o objetivo geral deste trabalho é compreender algoritmos e implementação destes na linguagem de programação R e compreender de certa forma, como os softwares prontos para esse fim funcionam. De forma mais específica, neste trabalho fizemos um aprofundamento sobre problemas de otimização não lineares, começando com uma revisão de problemas de otimização simples e sem restrições, que podem ser resolvidos diretamente pelos métodos vistos nas disciplinas de Cálculo Diferencial e Integral. Mesmo nesses casos mais simples pode haver a necessidade do uso de algum método numérico para resolução de equações não lineares, como por exemplo, o Método de Newton ([11, p.79]).

Por fim, estudamos outros métodos de determinação de máximos e mínimos de funções com e sem restrições e possíveis aplicações, como por exemplo o problema de determinação de autovalores e autovetores. Após o estudo teórico, implementamos na linguagem de programação R os métodos estudados.

Este trabalho foi estruturado da forma como a seguir.

No segundo capítulo introduzimos conceitos básicos de Cálculo Diferencial e Integral e Álgebra Linear apresentando uma demonstração para o Teorema Fundamental da Álgebra.

No terceiro capítulo estudamos alguns métodos para resolver equações do tipo  $f(x) = 0$ , com ênfase no Método de Newton.

O quarto capítulo é dedicado aos métodos de resolver problemas de otimização com restrições de igualdade e desigualdade e problemas de otimização por meio de Equações Diferenciais Ordinárias.

Por fim, apresentamos alguns algoritmos em Linguagem R no Apêndice A.

# Capítulo 2

## Máximos e mínimos de funções

Neste capítulo vamos relembrar alguns conceitos relacionados a máximos e mínimos de funções e autovalores e autovetores de transformações lineares, para que possamos relacionar de maneira adequada esses conceitos com os métodos Numéricos para aproximar a solução de problemas de otimização. Para isso, é preciso lembrar alguns conceitos de Cálculo Diferencial e Integral, Análise Real e Álgebra Linear.

### 2.1 Máximos e Mínimos

Para contextualizar o assunto proposto para este trabalho precisamos começar mencionando alguns fatos elementares de Cálculo Diferencial e Integral e de Análise Real e faremos isso baseado em [2] e [12].

Um ponto  $a \in A$  é chamado ponto interior do conjunto  $A$  se esse conjunto  $A$  contém um disco aberto de centro  $a$  (o disco aberto deve estar todo contido no conjunto  $A$ ). Um conjunto qualquer  $A$  é dito aberto se todos os seus pontos são pontos interiores.

**Definição 2.1.1** *Seja  $A$  um subconjunto não vazio do espaço euclidiano  $\mathbb{R}^m$ ,  $f : A \rightarrow \mathbb{R}$  uma função e considere o ponto  $a \in A$ . Então:*

1.  $f(a)$  é um valor **máximo local** de  $f$  se  $f(a) \geq f(x)$  para todos os pontos  $x \in A \cap B$ , em que  $B$  é um disco aberto centrado em  $a$ .
2.  $f(a)$  é um valor **mínimo local** de  $f$  se  $f(a) \leq f(x)$  para todos os pontos  $x \in A \cap B$ , em que  $B$  é um disco aberto centrado em  $a$ .

**Observação 2.1.2** *Nem sempre existem pontos de máximo ou de mínimo de uma função, como na definição anterior, mas sempre existem o ínfimo de  $f(x)$  e o supremo de  $f(x)$ , que podem ser finitos ou não.*

Dizemos que um subconjunto  $K \subset \mathbb{R}^m$  é limitado se existir um número real  $c > 0$  tal que  $|k| \leq c$  para todo  $k \in K$ . Isto equivale a dizer que  $K$  está contido na bola fechada de

centro na origem e raio  $c$ . E, um conjunto  $K$  é fechado se, e somente se, seu complementar  $K^c$  é aberto. Além disso, temos que um conjunto  $K \subset \mathbb{R}^m$  é compacto quando for limitado e fechado.

O próximo resultado pode ser visto como uma aplicação direta do Axioma da Completude do conjunto dos números reais, que diz que todo conjunto limitado de números reais possui supremo.

**Teorema 2.1.3** (*Bolzano-Weierstrass*): *Todo conjunto infinito e limitado  $K \subset \mathbb{R}$  possui algum ponto de acumulação.*

Esse resultado pode ser usado para demonstrar o próximo Teorema, e por isto, recebe a mesma denominação de alguns autores.

**Teorema 2.1.4** *Seja  $K \subset \mathbb{R}^m$  um conjunto compacto e seja  $f : K \rightarrow \mathbb{R}$  uma função contínua. Existem  $a, b \in K$  tais que*

$$f(a) \leq f(x) \leq f(b), \quad \text{para todo } x \in K,$$

ou seja,  $f(a)$  é o valor mínimo de  $f(x)$  e  $f(b)$  é o valor máximo.

A demonstração desse teorema passa pela aplicação do resultado anterior em cada coordenada do vetor  $x \in K$ .

Lembremos alguns conceitos importantes.

Seja  $f : D \rightarrow \mathbb{R}^n$  uma função definida no conjunto aberto  $D \subset \mathbb{R}^m$ . Dizemos que  $f$  é diferenciável no ponto  $a \in D$  quando existir uma transformação linear  $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$  tal que

$$f(a+h) = f(a) + f'(a)h + r(h), \quad \lim_{h \rightarrow 0} \frac{r(h)}{|h|} = 0, \quad (2.1)$$

onde  $|h|$  denota a norma do vetor  $h$ .

No caso em que  $D$  é um intervalo aberto da reta  $\mathbb{R}$ , podemos definir a derivada  $f'(a)$  de uma função  $f : D \rightarrow \mathbb{R}^n$  no ponto  $a \in U$ , através do limite

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}, \quad h \in \mathbb{R},$$

desde que o limite exista.

Diremos que a função  $f : D \rightarrow \mathbb{R}^n$  é derivável no conjunto  $D \subset \mathbb{R}$  quando existir a derivada de  $f$  em todos os pontos de acumulação de  $\overset{\circ}{D} \in D$ , em que  $\overset{\circ}{D} = \text{int}(D)$  denota o interior de  $D$ .

Seja  $a$  um ponto de acumulação. Assim, neste caso temos que

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

se esse limite existir. Assim,

$$\frac{f(a+h) - f(a)}{h} \simeq f'(a), \quad \text{para } h \text{ bem próximo de } 0,$$

e

$$f(a+h) - f(a) \simeq f'(a) \cdot h, \quad \text{para } h \text{ bem próximo de } 0. \quad (2.2)$$

Podemos então escrever a equação (2.2), com auxílio do polinômio de Taylor de ordem  $p$  da seguinte maneira:

$$f(a+h) - f(a) = f'(a) \cdot h + \frac{r(h)}{h},$$

com  $r(h)$  tendendo a zero mais rapidamente do que  $h$  tende a zero, veja [13, cap.5].

Um ponto  $a$  do conjunto aberto  $D \subset \mathbb{R}^m$  pode ser escrito como  $a = (a_1, a_2, \dots, a_m)$ , com  $a_1, a_2, \dots, a_m \in \mathbb{R}$ . Dada a função  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}$ , podemos considerá-la como função de uma variável, fixando  $n - 1$  coordenadas do ponto  $a$  exceto a  $i$ -ésima, isto é, considerando a função

$$x \rightarrow f(a_1, a_2, \dots, a_{i-1}, x, a_{i+1}, \dots, a_m), \quad \text{para } x \simeq a_i,$$

definida num intervalo  $(a_i - \varepsilon, a_i + \varepsilon)$ . E podemos considerar a derivada desta função, caso exista, através da definição a seguir.

**Definição 2.1.5** *Seja  $D \subset \mathbb{R}^m$  um conjunto aberto,  $f : D \rightarrow \mathbb{R}^n$  uma função e  $a \in D$ . A  $i$ -ésima derivada parcial de  $f$ , no ponto  $a$ , que se representa por  $\frac{\partial f}{\partial x_i}(a)$  é, caso exista, dada por*

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h \rightarrow 0} \frac{f(a_1, a_2, \dots, a_i + h, \dots, a_m) - f(a_1, a_2, \dots, a_m)}{h}$$

ou melhor,

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h \rightarrow 0} \frac{f(a + he_i) - f(a)}{h},$$

com  $e_i$  sendo o vetor da base canônica de  $\mathbb{R}^m$ , cuja  $i$ -ésima componente é 1 e as demais são nulas.

Pode-se mostrar (veja [12, cap.3]), que se  $\frac{\partial f}{\partial x_i}(a)$  existir para  $i = 1, \dots, m$  e  $h = (h_1, h_2, \dots, h_m) \in \mathbb{R}^m$ , então temos

$$f'(a)h = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(a)h_i, \quad (2.3)$$

em (2.1). Ainda, se  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}$  for uma função diferenciável de  $m$  variáveis, o gradiente de  $f$  é a função vetorial  $\nabla f$  definido por:

$$\nabla f(x) = \left( \frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_m} \right).$$

Em particular,  $f'(a)h = \nabla f(a) \cdot h$ , onde  $\cdot$  denota o produto escalar em  $\mathbb{R}^m$ .

De modo geral, a derivada de  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ , no ponto  $a$ , pode ser vista como uma matriz chamada de *Matriz Jacobiana* de  $f$  no ponto  $a$ .

A função  $f(x)$  pode ser vista como um vetor, se escrita da seguinte maneira

$$f(x) = (f_1(x), f_2(x), \dots, f_n(x)).$$

Assim, temos que cada coordenada de  $f(x)$  pode ser vista como  $f_i : D \rightarrow \mathbb{R}$ , com  $f'_i(a) : \mathbb{R}^m \rightarrow \mathbb{R}$  e  $f'_i(a) \cdot h = \nabla f_i(a) \cdot h$ .

Pela definição de derivada temos que  $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ . (veja [12, cap. 3]) e assim,

$$f'(a) \cdot h = (f'_1(a) \cdot h, f'_2(a) \cdot h, \dots, f'_n(a) \cdot h) = Jf(a) \cdot h,$$

onde

$$Jf(a) = \begin{bmatrix} \nabla f_1(a) \\ \nabla f_2(a) \\ \vdots \\ \nabla f_n(a) \end{bmatrix}$$

é uma matriz cujas linhas são os vetores gradiente de  $f_i$  é chamada de *Matriz Jacobiana* de  $f$  no ponto  $a$ .

O Lema que segue relaciona o conceito de derivada ao conceito de máximo e mínimo de funções.

**Lema 2.1.6** *Sejam  $I$  um intervalo aberto, uma função  $f : I \rightarrow \mathbb{R}$  e  $a \in I$ . Se  $f(a) \geq f(x)$ , para todo  $x \in I$ , então  $f'(a) = 0$ .*

**Demonstração:** Suponha que  $f(a) \geq f(x)$ .

Se  $x < a$  temos que

$$\frac{f(x) - f(a)}{x - a} \geq 0,$$

aplicando a definição de derivada, temos:

$$\lim_{x \rightarrow a^+} \frac{f(x) - f(a)}{x - a} = f'(a) \geq 0 \tag{2.4}$$



Se  $x > a$  temos que

$$\frac{f(x) - f(a)}{x - a} \leq 0,$$

e, da mesma forma, segue que:

$$f'(a) \leq 0 \tag{2.5}$$

Como  $f$  é derivável, de (2.4) e (2.5) temos que  $f'(a) = 0$ . ■

O próximo resultado nos dá uma condição necessária para que um ponto seja ponto de máximo ou mínimo local.

**Teorema 2.1.7** *Se  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}$  tiver um valor de máximo ou mínimo local em um ponto interior  $a$  do seu domínio  $D$  e se as derivadas parciais de primeira ordem de  $f$  existem nesse intervalo, então  $\nabla f(a) = 0$ .*

**Demonstração:** Suponha, por simplicidade que  $a$  é um ponto de máximo local de  $f$  em  $\mathring{D}$ . Considere  $v \in \mathbb{R}^m$  um vetor arbitrário e  $\varepsilon > 0$  tal que  $a + hv \in D$ , quando  $h \in \mathbb{R}$  e  $|h| < \varepsilon$ . Então  $h = 0$  é um ponto de máximo local da função de uma variável real  $g(h) = f(a + hv)$ . Pelo Lema 2.1.6, temos que

$$g'(0) = 0.$$

Pela Regra da Cadeia (veja [12, p. 129]) temos ainda que:

$$g'(0) = \nabla f(a) \cdot v = f'(a)v = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(a)v_i,$$

ou seja,  $\nabla f(a)$  é ortogonal a  $h$ .

Escolhendo,  $h = \nabla f(a)$  temos que  $|\nabla f(a)| = 0$  e que  $\nabla f(a)$  é nulo. ■

Esse teorema fornece uma condição para pontos de máximo e de mínimo da função  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ , para pontos no interior de  $D$ . Esses pontos são as raízes da equação  $\nabla f(x) = 0$ .

**Observação 2.1.8** *Dizemos que o ponto  $a$  é um ponto crítico de  $f$  quando  $\nabla f(a)$  for zero.*

Dado um conjunto  $D \subset \mathbb{R}^n$  e um ponto  $a \in \mathbb{R}^n$  dizemos que:

- $a$  é um *ponto interior* de  $D$  se existe uma bola centrada em  $a$ , que esteja contida em  $D$ , ou seja,  $a \in \text{int}(D)$ .

- $a$  é um *ponto exterior* de  $D$  se existe uma bola centrada em  $a$ , que esteja contida em  $D^c$ , ou seja,  $a \in \text{ext}(D)$ .
- $a$  é um *ponto da fronteira* se toda bola centrada em  $a$  intercepta  $D$  e  $D^c$ .

Sabemos que os únicos pontos nos quais uma função  $f(x)$  pode assumir valores de máximo ou de mínimo são os pontos críticos e os pontos de fronteira. Assim, como as funções diferenciáveis de uma variável real podem ter um ponto de inflexão, as funções diferenciáveis com várias variáveis podem ter pontos de sela.

**Definição 2.1.9** *Uma função diferenciável  $f(x)$  tem um **ponto de sela** em um ponto crítico  $a$  se em todo disco aberto centrado em  $a$ , existem pontos  $x$  e  $y$  do domínio tais que  $f(x) > f(a)$  e  $f(y) < f(a)$ . O ponto correspondente  $(a, f(a))$  na superfície  $z = f(x)$  é chamado de ponto de sela da superfície. Veja a Figura 2.1 ilustrativa a seguir.*

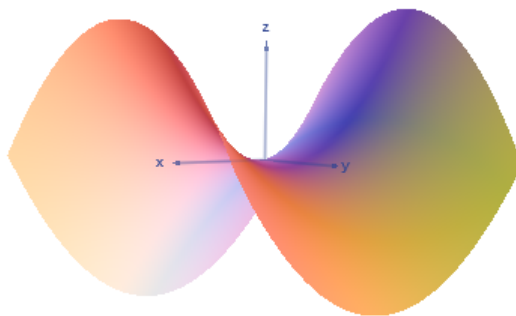


Figura 2.1: Ponto de Sela.

Como vimos, o gradiente de  $f$  no ponto  $a$  pode ser escrito em forma de vetor e consequentemente em forma de matriz. Desse modo, como o vetor gradiente de  $f(x)$  é  $n \times 1$ , podemos escrever as derivadas de segunda ordem de  $f(x)$  como a matriz:

$$Hf(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n x_1} & \frac{\partial^2 f(x)}{\partial x_n x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix},$$

chamada de Matriz Hessiana de  $f(x)$ .

**Observação 2.1.10** *Como vimos anteriormente se  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}$  é diferenciável no aberto  $D$ , para  $i = 1, 2, \dots, m$ , temos que as funções  $\frac{\partial f}{\partial x_i} : D \rightarrow \mathbb{R}$  podem ser diferenciáveis num ponto  $a \in D$ . Se todas as funções  $\frac{\partial f}{\partial x_i}$  forem diferenciáveis, dizemos que*

$f$  é duas vezes diferenciável no ponto  $a$ . Assim, para todos os inteiros  $i, j = 1, 2, \dots, m$ , existem as derivadas parciais de segunda ordem dadas por:

$$\frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_j} \right) (a) = \frac{\partial^2 f}{\partial x_i \partial x_j} (a).$$

Formalmente podemos demonstrar que uma função de duas vezes diferenciável  $f : D \rightarrow \mathbb{R}$ , com ponto crítico  $a$ , ou seja, com  $f'(a) = 0$ , no conjunto aberto  $D \subset \mathbb{R}^m$ , pode ser escrita como

$$f(a + h) = f(a) + \langle Ah, h \rangle + r(h), \quad \lim_{h \rightarrow 0} \frac{r(h)}{|h|^2} = 0,$$

onde  $A$  denota a matriz Hessiana de  $f$ , no ponto  $a$ , que é autoadjunta, ou seja, a matriz é igual à sua transposta conjugada, [14, p. 68], e  $\langle \cdot, \cdot \rangle$  denota um produto interno de  $\mathbb{R}^m$ .

É possível encontrar uma base ortonormal de  $\mathbb{R}^m$  de tal forma que uma mudança de base possibilite reescrever  $a$  como  $b$  na nova base e valha a igualdade

$$f(b + u) = f(b) + \langle Cu, u \rangle + r(u), \quad \lim_{u \rightarrow 0} \frac{r(u)}{|u|^2} = 0, \quad (2.6)$$

com

$$C = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

Isso será demonstrado na seção seguinte e garante que o ponto crítico  $a$  determina um ponto de máximo local, quando todos os autovalores forem negativos, ou um ponto de mínimo local, quando todos os autovalores forem positivos. Quando houver um autovalor com sinal positivo e um autovalor com sinal negativo, o ponto crítico  $a$  determina um ponto de sela. Caso o sinal dos autovalores seja maior ou igual a 0, por exemplo, o ponto  $a$  é de mínimo mas pode não ser único naquela região.

A relação entre ponto crítico e a Matriz Hessiana é estabelecida pelo teorema seguinte, que confirma o que acabamos de comentar sobre o sinal dos autovalores da matriz Hessiana  $Hf(a)$ .

**Teorema 2.1.11** *Sejam  $f : D \rightarrow \mathbb{R}$  uma função de classe  $C^2$ , ou seja, função onde as derivadas parciais de 2ª ordem são contínuas, com  $D$  aberto. Se  $a \in D$  for um ponto crítico de  $f$  e  $A = Hf(a)$  a matriz Hessiana de  $f$  no ponto  $a$ . Então:*

- i) Se  $A$  é positiva, ou seja, possui apenas autovalores positivos,  $a$  é um ponto de mínimo local;*

- ii) Se  $A$  é negativa, ou seja, possui apenas autovalores negativos,  $a$  é um ponto de máximo local;
- iii) Se  $A$  é indefinida, ou seja, possui um autovalor positivo e um autovalor negativo,  $a$  é um ponto sela.

A demonstração do Teorema 2.1.11 está relacionada com (2.6) e pode ser encontrada em [12, p.158].

**Observação 2.1.12** Considere  $D \subset \mathbb{R}^2$ . Um resultado que nos auxilia em saber se um ponto  $a$  é ponto de máximo ou mínimo local ou um ponto de sela, no caso em que  $f(x)$  e suas derivadas parciais de primeira e segunda ordens sejam contínuas em um disco centrado em  $a$  e que  $\nabla f(a) = 0$  é o seguinte:

- i)  $f$  tem um **máximo local** em  $a$  se

$$\frac{\partial^2 f(a)}{\partial x^2} < 0 \quad e \quad \frac{\partial^2 f(a)}{\partial x^2} - \left( \frac{\partial^2 f(a)}{\partial x \partial y} \right)^2 > 0.$$

- ii)  $f$  tem um **mínimo local** em  $a$  se

$$\frac{\partial^2 f(a)}{\partial x^2} > 0 \quad e \quad \frac{\partial^2 f(a)}{\partial x^2} - \left( \frac{\partial^2 f(a)}{\partial x \partial y} \right)^2 > 0.$$

- iii)  $f$  tem um **ponto de sela** em  $a$  se

$$\frac{\partial^2 f(a)}{\partial x^2} - \left( \frac{\partial^2 f(a)}{\partial x \partial y} \right)^2 < 0.$$

- iv) **Nada se pode afirmar** em  $a$  se

$$\frac{\partial^2 f(a)}{\partial x^2} - \left( \frac{\partial^2 f(a)}{\partial x \partial y} \right)^2 = 0.$$

Essa Observação é um caso particular do Teorema (2.1.11).

Se o conjunto  $D$  não for limitado e fechado temos que  $f(x)$  pode não ter pontos de máximo e de mínimo local ou global em  $D$ , mas sempre terá um supremo e um ínfimo. Assim, partindo de  $a \in D$  podemos procurar um valor melhor para  $f(x)$  do que  $f(a)$ .

## 2.2 Autovalores e Autovetores

A Equação (2.7) motiva o estudo da existência de autovalores e de autovetores de uma matriz, ou de uma transformação linear. Como o processo de determinação de autovalores

muitas vezes utiliza raízes de polinômios associados ao operador, começamos apresentando uma demonstração do clássico Teorema Fundamental da Álgebra. A demonstração apresentada pode ser vista num curso de Cálculo Diferencial e Integral, e é devida a Gauss e adaptada por Cauchy (veja [15] para uma revisão histórica, embora use conceitos básicos de números complexos, sobre equações polinomiais).

**Lema 2.2.1 (Teorema Fundamental da Álgebra)** *Seja*

$$p(z) = \sum_{j=0}^n a_j z^j, \quad z \in \mathbb{C},$$

um polinômio, com cada  $a_j \in \mathbb{C}$ ,  $n \geq 1$  é um número natural e  $a_n = 1$ . Existe  $b \in \mathbb{C}$  tal que  $p(b) = 0$ .

**Demonstração:** Primeiramente note que  $p(z) = p(x + iy)$ , com  $x, y \in \mathbb{R}$  e  $i^2 = -1 = e^{i\pi}$ , é uma função contínua e que podemos escolher  $t > 0$  de tal forma que  $|p(z)| > |p(0)| + 1$ , sempre que  $|z| > t$ . O Teorema 2.1.4 garante que existe  $b \in \mathbb{C}$  tal que  $\min_{|z| \leq t} |p(z)| = |p(b)| \leq |p(0)|$  e podemos reescrever (após manipulação de potências de  $z = \pm b$ )

$$p(z + b) = p(b) + c_k z^k + r(z) z^{k+1}, \quad z \in \mathbb{C},$$

onde  $r(z)$  é polinômio,  $c_k > 0$  e  $k \geq 1$  é um número natural,

Se  $p(b) \neq 0$ , então

$$\left| \frac{p(z + b)}{p(b)} \right| \leq \left| 1 + \frac{c_k}{p(b)} z^k \right| + \left| \frac{r(z)}{p(b)} z^{k+1} \right|.$$

Escolhendo  $z = |z| e^{i(\theta + \pi)/k}$  suficientemente próximo de 0, com  $c_k/p(b) = |c_k/p(b)| e^{-i\theta}$ , obtemos a desigualdade

$$\left| \frac{p(z + b)}{p(b)} \right| \leq \left| 1 - \frac{c_k}{p(b)} |z|^k \right| + \left| \frac{r(z)}{p(b)} |z|^{k+1} \right| < 1,$$

uma vez que um múltiplo (limitado) de  $|z|^{k+1}$  decresce para 0 mais rápido que qualquer múltiplo de  $|z|^k$ , quando  $|z|$  se aproxima de 0. Isso mostra que existe  $\delta > 0$  tal que

$$|p(z + b)| < |p(b)|, \quad |z + b| < t, \quad 0 < |z| < \delta,$$

o que contradiz o fato de  $|p(b)|$  ser mínimo. Sendo assim,  $p(b)$  tem que ser nulo. ■

Podemos usar esse teorema para demonstrar o que segue, de forma mais geral, mas preferimos uma demonstração mais direta.

Consideramos o espaço vetorial  $E$  real ou complexo e de dimensão finita (euclidiano), com produto interno denotado por  $\langle \cdot, \cdot \rangle$ . Lembramos que uma transformação linear  $H : E \rightarrow E$  é autoadjunta quando  $H = H^*$ , onde  $H^*$  denota o adjunto de  $H$ , e  $\bar{K}$  denota o conjugado de  $K$ . O resultado mencionado é o seguinte:

**Teorema 2.2.2** *Sejam  $E$  um espaço euclidiano e  $H : E \rightarrow E$  uma transformação linear autoadjunta. Então existe uma base ortonormal de  $E$  formada por autovetores de  $H$ , e os autovalores são todos reais.*

**Demonstração:** Seja  $E$  um espaço vetorial com produto interno de dimensão  $n \geq 1$  sobre  $\mathbb{R}$  ou  $\mathbb{C}$  (veja [16] para o caso  $\mathbb{C}$ ). Dobrando os elementos da base, se necessário, podemos sempre manipular os elementos de  $E$  como se estivessem em  $\mathbb{R}^n$ , que é um espaço vetorial sobre  $\mathbb{R}$ , e é isso que faremos.

Seja

$$K = \{x \in E : |x| = 1\},$$

que é compacto, e seja

$$\psi(x) = \langle Hx, x \rangle = \overline{\langle Hx, x \rangle}, \quad x \in K,$$

que é contínua.

Segue do Teorema 2.1.4 que existe  $u \in K$  e  $\lambda_1 \in \mathbb{R}$  tais que

$$\psi(u) = \max_{x \in K} \psi(x) = \lambda_1.$$

Note que, se  $x \in E - \{0\}$ ,

$$\begin{aligned} \langle (H - \lambda_1 I)x, x \rangle &= |x|^2 \left( \psi \left( \frac{x}{|x|} \right) - \lambda_1 \right) \\ &\leq |x|^2 (\lambda_1 - \lambda_1) = 0. \end{aligned} \tag{2.7}$$

Escolhendo  $x = u + ty$ , com  $y \in E$  e  $t \in \mathbb{R}$  temos de (2.7) que

$$\begin{aligned} \langle (H - \lambda_1 I)(u + ty), u + ty \rangle &= t \langle (H - \lambda_1 I)u, y \rangle + t \langle (H - \lambda_1 I)y, u \rangle + t^2 \langle (H - \lambda_1 I)y, y \rangle \\ &= 2t \langle (H - \lambda_1 I)u, y \rangle + t^2 \langle (H - \lambda_1 I)y, y \rangle \\ &\leq 0 \end{aligned} \tag{2.8}$$

Como a Equação (2.8) é um polinômio de grau 2 em  $t \in \mathbb{R}$ , temos

$$\langle (H - \lambda_1 I)u, y \rangle = 0, \quad y \in E,$$

o que produz a igualdade  $(H - \lambda_1 I)u = 0$ .

Assim, mostramos a existência de um autovalor e de um autovetor de  $H$ .

Para prosseguir, considere  $E_1$  como o subespaço de  $E$  ortogonal ao vetor  $u_1$ , observe que  $H(v) \in E_1$ , para cada  $v \in E_1$ , e repita o raciocínio para  $H$  restrito a  $E_1$ .

Note que o procedimento garante que  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ . Ainda, se  $Hu_j = \lambda_j u_j$  e  $E_p$  é o subespaço ortogonal a cada  $u_j$ , para  $j < p$ , então

$$\lambda_{p+1} = \max_{\substack{x \in E_p \\ \|x\|=1}} \langle Hx, x \rangle.$$

■

Esse resultado garante que, se  $A = (a_{ij})$  é uma matriz de ordem  $n \times n$  tal que  $\overline{A}^t = A$ , então existe uma matriz  $P$ ,  $n \times n$  cujas colunas são autovetores de  $A$ , com norma, tal que  $\overline{P}^t = P^{-1}$  e  $\overline{P}^t C P = A$ . Em particular, se  $f(x)$  é de classe  $C^2$ , então  $\overline{Hf(x)}^t = Hf(x)$  e (2.7) é verdadeiro.

# Capítulo 3

## Métodos numéricos para equações não lineares

A análise numérica é o estudo de algoritmos que buscam resultados numéricos de problemas envolvendo modelos matemáticos que podem estar relacionados com diferentes áreas do conhecimento. Este capítulo trata basicamente de algoritmos iterativos que apresentam uma sucessão de passos que converge ou não para o valor aproximado da solução exata de uma equação ou modelo matemático. O campo da análise numérica antecede a invenção do computador em séculos. Grandes matemáticos no passado trabalharam com análise numérica, o que é obviamente percebido pelo nome de importantes algoritmos como o Método de Newton, Eliminação de Gauss e Método de Euler.

No Capítulo 2 estudamos pontos críticos de funções, com a definição de máximo e mínimo e alguns resultados que garantem a existência desses pontos e analisamos o caso em que a função possui derivada no interior do domínio, pois se  $a$  for um ponto crítico do domínio então  $\nabla f(a)$  é um vetor nulo. Assim, concluímos que uma maneira de encontrar pontos críticos de funções é encontrar onde o gradiente é zero. Para decidir se este é ponto de máximo ou de mínimo analisamos os autovalores da Matriz Hessiana.

Neste capítulo procuramos entender como alguns métodos numéricos funcionam para resolver equações algébricas e então utilizá-los na resolução de problemas de otimização não linear, tendo como base o livro [11]. O objetivo principal é resolver equações do tipo  $f(x) = 0$ , onde  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}^m$ , que estão relacionados com problemas de otimização, justificado pelo Teorema 2.1.7.

### 3.1 Método de Newton

Do Teorema 2.1.7 temos que se a função diferenciável  $f$  tiver um ponto crítico no ponto  $a$ , então  $Jf(a) = 0$ . Com isso podemos procurar os pontos críticos de  $f$ , a fim de determinar pontos de máximo ou de mínimo da função  $f$ . Assim, nesta seção vamos



trabalhar com métodos numéricos para a resolução de equações não lineares da forma  $f(x) = 0$ . O Método de Newton é uma técnica para aproximar raízes desse tipo de equações. Uma das maneiras de deduzir o Método de Newton é baseado no Método de Iteração Linear, onde procuramos reescrever a equação  $f(x) = 0$  como a equação

$$\varphi(x) = x + A(x)f(x), \quad (3.1)$$

para alguma função  $\varphi(x)$  e alguma função  $A(x) \neq 0$  para  $x$  pertencente ao domínio da  $f$ .

A ideia do Método da iteração linear é escrever a equação  $f(x) = 0$  como a equação  $\varphi(x) = x$ . Para isso, devemos escolher  $A(x)$  de tal forma que a sequência

$$\begin{aligned} x_0 &= \text{aproximação inicial} \\ x_1 &= \varphi(x_0) \\ x_2 &= \varphi(x_1) \\ &\vdots \\ x_{n+1} &= \varphi(x_n) \\ &\vdots \end{aligned} \quad (3.2)$$

seja convergente para uma raiz de  $f(x) = 0$ . O valor inicial  $x_0$  deve estar no domínio de  $f$  e preferencialmente estar próximo da raiz de  $f(x) = 0$ .

A forma mais usada para encontrar os pontos críticos de  $f(x)$  caso a função seja derivável e  $\bar{x}$  seja uma raiz de  $f(x) = 0$ , no interior do domínio da  $f$ , é escolher  $A(x) = -[Jf(x)]^{-1}$ , sendo  $Jf(x)$  a Matriz do Jacobiano de  $f(x)$  desde que  $A(x)$  exista.

Para entender essa escolha, note que, o Polinômio de Taylor temos que

$$\begin{aligned} f(\bar{x}) &= f(x) + Jf(x)(\bar{x} - x) + r(\bar{x} - x), \\ &\simeq f(x) + Jf(x)(\bar{x} - x), \quad \text{com } x \simeq \bar{x} \text{ e } \lim_{h \rightarrow 0} \frac{r(h)}{|h|} = 0. \end{aligned}$$

Como  $f(\bar{x}) = 0$  segue que

$$-f(x) \simeq Jf(x)(\bar{x} - x), \quad x \simeq \bar{x},$$

ou

$$[Jf(x)]^{-1}(\bar{x} - x) \simeq -f(x).$$

Logo,

$$x \simeq \bar{x} + [Jf(x)]^{-1}f(x).$$

Quando  $A(x) = -[Jf(x)]^{-1}$  a sequência (3.2) recebe o nome de Método de Newton e quando  $A(x) = -[Jf(x_0)]^{-1}$  a sequência (3.2) recebe o nome de Método de Newton Modificado.

### 3.1.1 Análise de Convergência do Método de Newton

Pode ser que a sequência obtida pelo Método de iteração linear não seja convergente para qualquer valor inicial. E mesmo que escolhamos um  $x_0$  razoavelmente perto da raiz, a sequência pode divergir.

Assim, para garantir a convergência do Método, precisamos dos seguintes resultados, baseados em [2], [12] e [14].

**Teorema 3.1.1** (*Desigualdade do Valor Médio*) *Seja  $f : [a, b] \rightarrow \mathbb{R}^n$  uma função contínua e diferenciável no intervalo aberto  $(a, b)$ . Se  $|f'(t)| \leq M$ , para todo  $t \in (a, b)$ , então*

$$|f(b) - f(a)| \leq M \cdot (b - a). \quad (3.3)$$

**Demonstração:** Note que  $\varphi(t) = \langle f(t), f(b) - f(a) \rangle$ , para  $t \in [a, b]$ , é contínua e possui derivada  $\varphi'(t) = \langle f'(t), f(b) - f(a) \rangle$ , em  $[a, b]$ .

Do teorema do valor médio para funções  $g : [a, b] \rightarrow \mathbb{R}$ , (veja [12, p. 89]), segue que existe  $c \in (a, b)$  tal que  $\varphi(b) - \varphi(a) = \varphi'(c)(b - a)$ , ou seja,

$$\begin{aligned} \varphi(b) - \varphi(a) &= \langle f(b) - f(a), f(b) - f(a) \rangle \\ |f(b) - f(a)|^2 &= \langle f'(c), f(b) - f(a) \rangle \cdot (b - a) \end{aligned}$$

Pela Desigualdade de Cauchy-Schwarz, (veja [14, p. 3]) temos:

$$\begin{aligned} |f(b) - f(a)|^2 &\leq |f'(c)| |f(b) - f(a)| (b - a) \\ &\leq M |f(b) - f(a)| (b - a) \end{aligned}$$

e segue que

$$|f(b) - f(a)| \leq |f'(c)|(b - a).$$

■

**Corolário 3.1.2** *Dado  $D \subset \mathbb{R}^m$  aberto, seja  $f : D \rightarrow \mathbb{R}^n$  diferenciável em cada ponto do segmento de reta aberto  $(a, a + h)$  e contínua no segmento fechado  $[a, a + h] \subset D$ . Se  $|f'(x)| \leq M$  para todo  $x \in (a, a + h)$ , então*

$$|f(a + h) - f(a)| \leq M|h|. \quad (3.4)$$

**Demonstração:** A aplicação  $\varphi : [0, 1] \rightarrow \mathbb{R}^n$ , definido por  $\varphi(t) = f(a + th)$  é contínua e diferenciável no intervalo aberto  $(0, 1)$ , com  $\varphi(0) = f(a)$  e  $\varphi(1) = f(a + h)$ . Pela Regra da Cadeia, (veja [12, p. 87]), temos que

$$\varphi'(t) = f'(a + th) \cdot h,$$

onde,

$$|\varphi'(t)| \leq |f'(a + th)| \cdot |h| \leq M \cdot |h|,$$

para todo  $t \in [0, 1]$ . Segue do Teorema 3.1.1 que  $|\varphi(1) - \varphi(0)| \leq M \cdot |h|$ , e o resultado segue. ■

**Corolário 3.1.3** *Seja  $D \subset \mathbb{R}^m$  aberto e convexo. Se  $f : D \rightarrow \mathbb{R}^n$  é diferenciável, com  $|f'(x)| \leq M$  para todo  $x \in D$  então  $|f(y) - f(x)| \leq M \cdot |y - x|$ , para quaisquer  $x, y \in D$ .*

**Corolário 3.1.4** *Se  $f'$  for contínua em  $a$  e  $f'(a) = 0$ . Existe  $r > 0$  tal que  $|f'(x)| \leq M < 1$ , com  $|x - a| \leq r$ .*

Para concluir que o Método de Newton converge vamos usar um teorema muito importante em várias áreas, como na teoria e resolução de equações diferenciais ordinárias, como veremos na Seção 4.3.

**Teorema 3.1.5** *(Teorema do Ponto Fixo de Banach) Seja  $(M, d)$  um espaço métrico completo. Se  $\varphi : M \rightarrow M$  tal que*

$$d(\varphi(u), \varphi(v)) \leq c d(u, v), \quad u, v \in M, \quad (3.5)$$

*e alguma constante  $0 \leq c < 1$ . Então, existe um único ponto  $a \in M$  tal que  $\varphi(a) = a$ . Além disso, para qualquer  $x_0 \in M$ , a sequência*

$$\begin{aligned} x_1 &= \varphi(x_0) \\ x_2 &= \varphi(x_1) \\ &\vdots \\ x_{n-1} &= \varphi(x_n) \\ &\vdots \end{aligned}$$

*converge para  $a$ .*

**Demonstração:** Se  $\varphi(a) = a$  e  $\varphi(b) = b$  para algum  $a, b \in M$ . Então

$$d(a, b) = d(\varphi(a), \varphi(b))$$

Por hipótese temos que

$$d(\varphi(a), \varphi(b)) \leq c d(a, b), \quad \text{com } (1 - c) > 0.$$

Então  $(1 - c) d(a, b) \leq 0$  e segue que  $d(a, b) = 0$ , ou seja,  $a = b$ .

Agora vamos mostrar que para qualquer  $x_0 \in M$ , a sequência  $x_1 = \varphi(x_0)$ ,  $x_2 = \varphi(x_1), \dots, x_{n+1} = \varphi(x_n)$  converge para  $a$ . Para isso, mostraremos que  $(x_n)$  é sequência de Cauchy. Tomemos  $m = n + p \in \mathbb{N}$ . Daí, usando a Desigualdade Triangular e (3.5),

$$\begin{aligned} d(x_{n+1}, x_{m+1}) &= d(\varphi(x_n), \varphi(x_m)) \\ &= d(\varphi(x_n), \varphi(x_{n+p})) \\ &\leq c d(x_n, x_{n+p}) \\ &\leq c [d(x_n, x_{n+1}) + \dots + d(x_{n+p-1}, x_{n+p})] \\ &\leq c [c d(x_{n-1}, x_n) + c^2 d(x_{n-1}, x_n) + \dots + c^p d(x_{n-1}, x_n)] \\ &= c^2 (1 + c + \dots + c^{p-1}) d(x_{n-1}, x_n) \\ &\leq c^2 \left( \frac{1}{1 - c} \right) c^{n-1} d(x_0, x_1) \\ &= \left( \frac{c^{n+1}}{1 - c} \right) d(x_0, x_1) \end{aligned}$$

Assim  $d(x_n, x_m)$  converge para 0 e  $(x_n)$  é uma sequência de Cauchy (veja [12, p. 5]). Logo, existe  $a \in M$  tal que

$$a = \lim(x_n) = \lim(x_{n+1}) = \lim f(x_n) = f(\lim(x_n)) = f(a),$$

pois  $M$  é completo e  $f$  é contínua. (veja [17, p. 154] para uma outra demonstração deste teorema). ■

**Corolário 3.1.6** *Seja  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$  duas vezes derivável com  $f''$  contínua. Se  $f(a) = 0$  e existir  $[Jf(a)]^{-1}$ , então existe  $r > 0$  tal que, se  $|\varphi'(x)| \leq M < 1$ , se  $|x - a| \leq r$ . Em particular, o Método de Newton converge quando  $|x_0 - a| \leq r$ .*

**Demonstração:** A equação  $\varphi(x) = x - [Jf(x)]^{-1} f(x)$  tem um ponto fixo. De fato, pela regra do produto, temos

$$\varphi'(x) = x' - ([Jf(x)]^{-1})' f(x) - [Jf(x)]^{-1} f'(x)$$

Logo,

$$\begin{aligned}
 \varphi'(a) &= Id - ([Jf(a)]^{-1})' f(a) - [Jf(a)]^{-1} (Jf(a)) \\
 &= Id - ([Jf(a)]^{-1})' f(a) - Id \\
 &= -([Jf(a)]^{-1})' f(a) \\
 &= 0.
 \end{aligned}$$

Note que a Matriz Jacobiana de  $f(x)$  tem linhas que são os vetores gradientes das funções coordenadas de  $f(x)$ . Então cada coordenada do Jacobiano da função será derivável, logo a Matriz Jacobiana será derivável e como a inversa de  $Jf(x)$  está relacionada com a função determinante ela terá derivada, ou seja, a derivada da inversa é a inversa da derivada. Como  $Jf(x)$  é contínua temos existe  $r > 0$ , onde  $\det(Jf(x)) \neq 0$ , para  $|x - a| < r$ . Pelo Corolário 3.1.4 temos que

$$|\varphi'(x)| \leq M < 1, \quad |x - a| < r.$$

O resultado segue do Teorema 3.1.1 e do Corolário 3.1.3. ■

Veja o Exemplo a seguir para determinar uma raiz pelo Método de Newton.

**Exemplo 3.1.7** Usando o Método de Newton determine uma raiz da equação  $f(x) = x^3 + 3x^2 + 2x - 1 = 0$ .

Note que  $f'(x) = 3x^2 + 6x + 2$ . Pelo Método de Newton temos que:

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

é a função de iteração  $x_{n+1} = \phi(x_n)$ . Observando a Figura (3.1), tomamos como valor inicial  $x_0 = 1$ . Assim,

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)},$$

ou seja,

$$x_1 = 1 - \frac{f(1)}{f'(1)}.$$

Daí,

$$x_1 = 1 - \frac{5}{11} = \frac{6}{11} \simeq 0,5454545.$$

Da mesma forma,

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = \frac{2951}{8206} \sim 0.3596149$$

Quanto mais iterações foram feitas mais próximo do valor da raiz  $x_n$  chegará.

Utilizando o Algoritmo A.1 pudemos observar que a partir de  $n = 6$ ,  $f(x_6)$  permanece próximo de  $10^{-16}$ , ou seja, a sequência  $x_n$  estaciona.

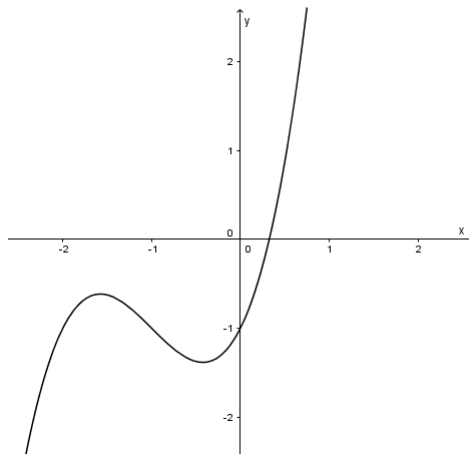


Figura 3.1: Gráfico do Exemplo.

As raízes dos polinômios são autovalores da matriz companheira do polinômio  $p(x)$ , que é uma matriz construída a partir de um polinômio, colocando os negativos dos coeficientes do polinômio em ordem crescente de grau na última linha de uma unidade,

$$\text{Matriz Companheira} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -2 & -3 \end{bmatrix}.$$

## 3.2 Implementação do Método do Tipo Newton

Como vimos na seção anterior, um Método para resolver equações não lineares da forma  $f(x) = 0$ , com  $f$  duas vezes diferenciáveis, é o Método de Newton através da seguinte função de iteração linear:

$$\varphi(x) = x - [Jf(x)]^{-1}f(x), \quad x \in D(f).$$

Ou melhor, escolhendo  $x_0 \in D(f)$  e tomando

$$x_{n+1} = x_n - [Jf(x_n)]^{-1}f(x_n).$$

Assim, precisamos então para  $n = 0, 1, 2, \dots$  resolver um sistema linear a cada passo

do método escrevendo

$$\begin{cases} Jf(x_n) s = -f(x_n) \\ x_{n+1} = s + x_n \end{cases},$$

para  $n = 0, 1, 2, \dots$ .

Para resolver o sistema anterior, podemos utilizar alguns métodos exatos ou iterativos que podem ser vistos em ([11, Cap. 4 e 5]). Detalhemos no que segue o Método de Eliminação de Gauss.

### 3.2.1 Método da Eliminação de Gauss

O Método de Eliminação de Gauss é um dos métodos para resolução de sistemas lineares e consiste em transformar um sistema linear dado num sistema triangular equivalente através de operações elementares sobre as linhas da matriz original, sendo essas:

- trocar a ordem de duas linhas da matriz;
- multiplicar uma linha da matriz por uma constante não-nula;
- somar as duas linhas da matriz.

O objetivo é organizar essa sequencia de operações de tal forma que o sistema linear resultante seja triangular superior. Considere o sistema linear genérico:

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

onde  $a_{ij}$  e  $b_j$  são dados, para  $i, j = 1, 2, \dots, n$ .

Primeiramente devemos escrever o sistema como uma matriz aumentada como

$$\left[ \begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_{11} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & b_{22} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} & b_{33} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} & b_{nn} \end{array} \right].$$

O primeiro passo é eliminar os elementos da primeira coluna abaixo da diagonal, ou seja, devemos eliminar o  $x_1$  da 2ª à  $n$ -ésima linha do sistema, supondo  $a_{11} \neq 0$ , da seguinte maneira:

- para eliminar  $a_{21}$  temos que somar a 2ª equação e a 1ª equação multiplicada por  $-\frac{a_{21}}{a_{11}}$ .

- para eliminar  $a_{31}$  temos que somar a 3ª equação e a 1ª equação multiplicada por  $-\frac{a_{31}}{a_{11}}$ .
- ⋮
- para eliminar  $a_{n1}$  temos que somar a  $n$ -ésima equação e a 1ª equação multiplicada por  $-\frac{a_{n1}}{a_{11}}$ .

lembrando que após essa iteração, apenas a 1ª linha ficará com os mesmos números da matriz original, ou seja, temos a nova matriz aumentada

$$\left[ \begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & | & b_1 \\ 0 & c_{22} & c_{23} & \dots & c_{2n} & | & d_2 \\ 0 & c_{32} & c_{33} & \dots & c_{3n} & | & d_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & c_{n2} & c_{n3} & \dots & c_{nn} & | & d_n \end{array} \right].$$

O segundo passo é eliminar os elementos da segunda coluna abaixo da diagonal, ou seja, devemos eliminar o  $x_2$  da 3ª à  $n$ -ésima linha do sistema obtido, supondo  $c_{22} \neq 0$ , da seguinte maneira:

- para eliminar  $c_{32}$  temos que somar a 3ª equação e a 2ª equação multiplicada por  $-\frac{c_{32}}{c_{22}}$ .
- para eliminar  $c_{42}$  temos que somar a 4ª equação e a 2ª equação multiplicada por  $-\frac{c_{42}}{c_{22}}$ .
- ⋮
- para eliminar  $c_{n2}$  temos que somar a  $n$ -ésima equação e a 2ª equação multiplicada por  $-\frac{c_{n2}}{c_{22}}$ .

Após essa iteração temos a matriz aumentada

$$\left[ \begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & | & b_1 \\ 0 & c_{22} & c_{23} & \dots & c_{2n} & | & d_2 \\ 0 & 0 & e_{33} & \dots & e_{3n} & | & f_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & e_{n3} & \dots & e_{nn} & | & f_n \end{array} \right].$$

Esse processo se repetirá até que a matriz aumentada fique da forma



$$\left[ \begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\ 0 & c_{22} & c_{23} & \dots & c_{2n} & d_2 \\ 0 & 0 & e_{33} & \dots & e_{3n} & f_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & g_{nn} & h_n \end{array} \right].$$

Assim, teremos no final o seguinte sistema triangular superior:

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1 \\ 0 \quad c_{22}x_2 + c_{23}x_3 + \dots + c_{2n}x_n = d_2 \\ 0 \quad 0 \quad e_{33}x_3 + \dots + e_{3n}x_n = f_3 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots = \vdots \\ 0 \quad 0 \quad 0 \quad \dots \quad g_{nn}x_n = h_n \end{array} \right.$$

que é equivalente ao original. Assim, para resolver esse sistema basta isolar  $x_n$  na última linha e do resultado obtido, substitua na linha  $n - 1$  para encontrar o valor da incógnita  $x_{n-1}$  e assim por diante até que encontremos os valores de todas as incógnitas.

Veja o Exemplo a seguir.

**Exemplo 3.2.1** Resolver o sistema linear

$$\left\{ \begin{array}{l} 5x_1 + 2x_2 + x_3 = 0 \\ 3x_1 + x_2 + 4x_3 = -7 \\ x_1 + x_2 + 3x_3 = -5 \end{array} \right.$$

por *Eliminação de Gauss*.

Primeiramente vamos escrever o sistema como matriz aumentada.

$$\left[ \begin{array}{ccc|c} 5 & 2 & 1 & 0 \\ 3 & 1 & 4 & -7 \\ 1 & 1 & 3 & -5 \end{array} \right]$$

Precisamos eliminar o primeiro elemento das linhas 2 e 3. Assim, para eliminar o primeiro elemento da linha 2 vamos multiplicar a primeira linha por  $-\frac{3}{5}$  e somar seu resultado com a linha 2. Do mesmo modo vamos multiplicar a primeira linha por  $-\frac{1}{5}$  e somar seu resultado com a linha 3 para eliminar o primeiro elemento da linha 3.

$$\left[ \begin{array}{ccc|c} 5 & 2 & 1 & 0 \\ 0 & -\frac{1}{5} & \frac{17}{5} & -7 \\ 0 & \frac{3}{5} & \frac{14}{5} & -5 \end{array} \right]$$

Agora precisamos eliminar o segundo elemento da linha 3. Para isso vamos multiplicar a

segunda linha por  $-3$  e somar seu resultado com a linha 3.

$$\left[ \begin{array}{ccc|c} 5 & 2 & 1 & 0 \\ 0 & -\frac{1}{5} & \frac{17}{5} & -7 \\ 0 & 0 & 13 & -26 \end{array} \right]$$

Desse modo teremos o seguinte sistema:

$$\begin{cases} 5x_1 + 2x_2 + x_3 = 0 \\ -\frac{1}{5}x_2 + \frac{17}{5}x_3 = -7 \\ 13x_3 = -26 \end{cases}$$

Logo,

$$x = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}.$$

(veja no anexo o algoritmo em linguagem R desse método no Apêndice (A.3)).

O Método de Eliminação de Gauss por Pivotamento é uma variação do Método de Eliminação de Gauss, trocando-se linhas e colunas da matriz aumentada, de forma a minimizar os erros de arredondamento nas operações de eliminação e evitar divisões por zero, caso  $a_{ii} = 0$ , por exemplo.

Uma forma de pensar o método (de Pivotamento parcial) é escolher o pivô como o elemento de maior módulo da primeira coluna.

$$B = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pn} & b_p \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{bmatrix}$$

Suponhamos que  $a_{p1}$  seja o pivô da matriz, basta trocar a linha  $p$  da matriz com a linha 1.

A partir daí, aplica-se o Método de Eliminação de Gauss (veja no anexo o algoritmo em linguagem R desse método no Apêndice (A.4)).

### 3.2.2 Decomposição LU

Muitas vezes precisamos resolver vários sistemas da forma  $Ax = b_i$ , onde  $A = (a_{ij})_{n \times n}$  e  $b_i$  é um vetor que surge, por exemplo, em um processo iterativo.

Este método consiste em decompor a matriz quadrada no produto de uma matriz triangular inferior  $L$  por uma matriz triangular superior  $U$ .

$$A = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ l_{21} & 1 & 0 & \dots & 0 \\ l_{31} & l_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \dots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & u_{22} & u_{23} & \dots & u_{2n} \\ 0 & 0 & u_{33} & \dots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & u_{nn} \end{bmatrix} \quad (3.6)$$

Nesse caso, é vantajoso decompor  $A = LU$ , como na Expressão (3.6).

Se for possível fazer a Eliminação de Gauss da matriz  $A$ , sem fazer troca de linhas ou colunas, então podemos obter a fatoração  $A = LU$  de forma bem simples. Basta tomar  $U$  como a matriz triangular resultante e

$$L = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ l_{21} & 1 & 0 & \dots & 0 \\ l_{31} & l_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \dots & 1 \end{bmatrix},$$

onde

$$\begin{aligned} l_{21} &= \frac{a_{21}}{a_{11}}, l_{31} = \frac{a_{31}}{a_{11}}, \dots, l_{n1} = \frac{a_{n1}}{a_{11}}, \\ l_{32} &= \frac{c_{32}}{c_{22}}, l_{42} = \frac{c_{42}}{c_{22}}, \dots, l_{n2} = \frac{c_{n2}}{c_{22}}, \\ &\vdots \end{aligned}$$

são obtidos por meio dos multiplicadores do processo de Eliminação de Gauss.

No Exemplo (3.2.1) temos que a matriz  $U$  é

$$U = \begin{bmatrix} 5 & 2 & 1 \\ 0 & -\frac{1}{5} & \frac{17}{5} \\ 0 & 0 & 13 \end{bmatrix}$$

e a matriz  $L$  são os multiplicadores utilizados para escalonar a matriz  $U$ .

$$L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{3}{5} & 1 & 0 \\ \frac{1}{5} & -3 & 1 \end{bmatrix}$$

Para que esse método seja válido é suficiente que o determinante dos menores principais (veja [11, p. 123]) de  $A$  sejam diferentes de 0.

**Teorema 3.2.2** *Sejam  $A = (a_{ij})$  uma matriz quadrada de ordem  $n$ , e  $A_k$  o menor principal, constituído dos  $k$  primeiros elementos das  $k$  primeiras colunas de  $A$ . Assumindo que  $\det(A_k) \neq 0$  para  $k = 1, 2, \dots, n - 1$ . Então, existe uma única matriz triangular inferior  $L = (l_{ij})$ , com  $l_{11} = l_{22} = \dots = l_{nn} = 1$ , e uma matriz triangular superior  $U = (u_{ij})$ , tais que  $A = LU$ .*

**Demonstração:** Vamos provar por indução finita.

Se  $n = 1$ , temos que

$$a_{11} = 1 \cdot a_{11}.$$

Tomando  $L = [l_{11}] = [1]$  e  $U = [a_{11}]$  temos que

$$A = LU.$$

Além disso,  $\det(A) = a_{11}$ .

Suponhamos agora que o teorema seja válido para  $n = k - 1$ , ou seja, que toda matriz  $B$  de ordem  $n = k - 1$  se decompõe unicamente no produto  $B = LU$  nas condições do Teorema.

Seja  $A$  uma matriz de ordem  $k$ . Vamos escrever

$$A = \begin{bmatrix} A_{k-1} & r \\ s & a_{kk} \end{bmatrix},$$

onde  $r$  e  $s$  são vetores, ambos com  $k - 1$  coordenadas.

A matriz  $A_{k-1}$  é de ordem  $k - 1$  satisfaz as condições do teorema e, por hipótese de indução, temos que  $A_{k-1} = L_{k-1} \cdot U_{k-1}$ . Afirmamos que  $A = LU$ , onde

$$L = \begin{bmatrix} L_{k-1} & 0 \\ p & 1 \end{bmatrix}, \quad U = \begin{bmatrix} U_{k-1} & q \\ 0 & u_{kk} \end{bmatrix},$$

sendo  $p$  e  $q$  vetores com  $k - 1$  coordenadas. Note que  $p, q$  e  $u_{kk}$  são desconhecidos e precisamos determiná-los. Efetuando o produto temos que

$$LU = \begin{bmatrix} L_{k-1} U_{k-1} & L_{k-1} q \\ p U_{k-1} & pq + u_{kk} \end{bmatrix} = \begin{bmatrix} A_{k-1} & r \\ s & a_{kk} \end{bmatrix}.$$

Com isso,

$$L_{k-1} U_{k-1} = A_{k-1} \tag{3.7}$$

e

$$\begin{aligned} L_{k-1} q &= r \\ p U_{k-1} &= s \\ pq + u_{kk} &= a_{kk} \end{aligned}$$

Por hipótese de indução, a Expressão (3.7) é válida. Como  $A_{k-1}$  é invertível, pois  $\det \neq 0$  temos que  $L_{k-1}$  e  $U_{k-1}$  são invertíveis e segue que:

$$\begin{aligned} q &= L_{k-1}^{-1} r \\ p &= s U_{k-1}^{-1} \\ u_{kk} &= a_{kk} - pq \end{aligned}$$

Logo, pelo Princípio da Indução Finita segue que  $A = LU$  e ainda,

$$\det(A) = \det(L) \cdot \det(U) = \det(U) = u_{11} \cdot u_{22} \cdot \dots \cdot u_{kk}$$

■

**Observação 3.2.3** *Pode-se mostrar que, se  $A = A^t = LU$ , então a sequência de matrizes  $A_1 = UL = L_1U_1$ ,  $A_2 = U_1L_1 = L_2U_2$ ,  $\dots$ ,  $A_n = U_{n-1}L_{n-1} = L_nU_n, \dots$ , com*

$$L_i = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ l_{21}^i & 1 & 0 & \dots & 0 \\ l_{31}^i & l_{32}^i & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1}^i & l_{n2}^i & l_{n3}^i & \dots & 1 \end{bmatrix}, \quad U_i = \begin{bmatrix} u_{11}^i & u_{12}^i & u_{13}^i & \dots & u_{1n}^i \\ 0 & u_{22}^i & u_{23}^i & \dots & u_{2n}^i \\ 0 & 0 & u_{33}^i & \dots & u_{3n}^i \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & u_{nn}^i \end{bmatrix},$$

converge para

$$B = \begin{bmatrix} \lambda_1 & q_{12} & q_{13} & \dots & q_{1n} \\ 0 & \lambda_2 & q_{23} & \dots & q_{2n} \\ 0 & 0 & \lambda_3 & \dots & q_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix},$$

onde  $\lambda_i$  é autovalor de  $A$  (veja [11, p. 223]).

Cabe notar que, se  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}$  é de classe  $C^2$ , então a matriz Hessiana de  $f$  é simétrica e esse processo pode ser usado em conjunto com o Teorema 2.1.11, para avaliar se um ponto crítico é de máximo, mínimo ou de sela.

Quando calculamos a Eliminação de Gauss por linhas da matriz, temos um processo

de eliminação de matriz. Assim, pelo algoritmo do Apêndice (A.2) calculamos a Decomposição  $LU$  e a partir desta, encontramos aproximações para os autovalores de  $A$ .

**Exemplo 3.2.4** *Consideremos a seguinte matriz:*

$$A = \begin{bmatrix} 54 & 2 & 5 & 4 & 1 \\ 2 & 34 & 6 & 0 & 9 \\ 5 & 6 & 11 & 1 & 1 \\ 4 & 0 & 1 & 81 & 22 \\ 1 & 9 & 1 & 22 & 78 \end{bmatrix}$$

*Pela Observação (3.2.3) temos que se  $A = LU$ , então a sequência de matrizes  $A_1 = UL = L_1U_1$ ,  $A_2 = U_1L_1 = L_2U_2$ ,  $\dots$ ,  $A_n = U_{n-1}L_{n-1} = L_nU_n, \dots$ , converge para os autovalores de  $A$ .*

*Assim, pelo algoritmo do Apêndice (A.2) calculamos a Decomposição  $LU$  e a partir desta, encontramos os autovalores de  $A$  sendo:  $\lambda_1 = 102.464580$ ,  $\lambda_2 = 59.398305$ ,  $\lambda_3 = 54.359077$ ,  $\lambda_4 = 32.714838$  e  $\lambda_5 = 9.063199$  que são os pontos de mínimo, pois os valores obtidos são todos positivos.*

Para encontrarmos os elementos da matriz  $L$  e da matriz  $U$  da forma direta, devemos calcular os elementos das linhas de  $U$  e os elementos da coluna de  $L$  na seguinte ordem e maneira:

- **1ª linha de  $U$**  : Para determinar a primeira linha de  $U$  deve-se fazer o produto da primeira linha de  $L$  por todas as colunas de  $U$  e igualar os resultados com os elementos da primeira linha de  $A$ , ou seja,

$$\begin{aligned} 1 \cdot u_{11} &= a_{11} \Rightarrow u_{11} = a_{11} \\ 1 \cdot u_{12} &= a_{12} \Rightarrow u_{12} = a_{12} \\ \vdots &= \vdots \Rightarrow \vdots \\ 1 \cdot u_{1n} &= a_{1n} \Rightarrow u_{1n} = a_{1n} \end{aligned}$$

Portanto,  $u_{1j} = a_{1j}$ ;  $j = 1, 2, \dots, n$

- **1ª coluna de  $L$**  : Para determinar a primeira coluna de  $L$  deve-se fazer o produto da 2ª até a  $n$ -ésima linha de  $L$  pela primeira coluna de  $U$  e igualar os resultados com os elementos da primeira coluna de  $A$ , ou seja,

$$\begin{aligned} l_{21}u_{11} &= a_{21} \Rightarrow l_{21} = \frac{a_{21}}{u_{11}} \\ l_{31}u_{11} &= a_{31} \Rightarrow l_{31} = \frac{a_{31}}{u_{11}} \\ \vdots &= \vdots \Rightarrow \vdots \\ l_{n1}u_{11} &= a_{n1} \Rightarrow l_{n1} = \frac{a_{n1}}{u_{11}} \end{aligned}$$

Portanto,  $l_{i1} = \frac{a_{i1}}{u_{11}}$ ;  $i = 1, 2, \dots, n$

- **2ª linha de  $U$**  : Para determinar a segunda linha de  $U$  deve-se fazer o produto da segunda linha de  $L$  pelas colunas de  $U$  (da 2ª à  $n$ -ésima) e igualar os resultados com os elementos da segunda linha de  $A$ , ou seja,

$$\begin{aligned} l_{21}u_{12} + u_{22} &= a_{22} \Rightarrow u_{22} = a_{22} - l_{21}u_{12} \\ l_{21}u_{13} + u_{23} &= a_{23} \Rightarrow u_{23} = a_{23} - l_{21}u_{13} \\ \vdots &= \vdots \Rightarrow \vdots \\ l_{21}u_{1n} + u_{2n} &= a_{2n} \Rightarrow u_{2n} = a_{2n} - l_{21}u_{1n} \end{aligned}$$

Portanto,  $u_{2j} = a_{2j} - l_{21}u_{1j}$ ;  $j = 1, 2, \dots, n$

- **2ª coluna de  $L$**  : Para determinar a segunda coluna de  $L$  deve-se fazer o produto da 3ª até a  $n$ -ésima linha de  $L$  pela segunda coluna de  $U$  e igualar os resultados com os elementos da segunda coluna de  $A$ , ou seja,

$$\begin{aligned} l_{31}u_{12} + l_{32}u_{22} &= a_{32} \Rightarrow l_{32} = \frac{a_{32} - l_{31}u_{12}}{u_{22}} \\ l_{41}u_{12} + l_{42}u_{22} &= a_{42} \Rightarrow l_{42} = \frac{a_{42} - l_{41}u_{12}}{u_{22}} \\ \vdots &= \vdots \Rightarrow \vdots \\ l_{n1}u_{12} + l_{n2}u_{22} &= a_{n2} \Rightarrow l_{n2} = \frac{a_{n2} - l_{n1}u_{12}}{u_{22}} \end{aligned}$$

Portanto,  $l_{i2} = \frac{a_{i2} - l_{i1}u_{12}}{u_{22}}$ ;  $i = 1, 2, \dots, n$

Assim, continuando os cálculos da 3ª linha de  $U$ , a 3ª coluna de  $L$  assim por diante teremos as seguintes formulas gerais:

$$\begin{aligned} u_{ij} &= a_{ij} - \sum_{k=1}^{i-1} l_{ik}u_{kj}, & i \leq j, \\ l_{ij} &= \frac{a_{ij} - \sum_{k=1}^{j-1} l_{ik}u_{kj}}{u_{jj}}, & i > j. \end{aligned}$$

No próximo Exemplo, vamos resolver o mesmo sistema linear do Exemplo (3.2.1), utilizando agora a Decomposição LU.

**Exemplo 3.2.5** Resolver o sistema linear

$$\begin{cases} 5x_1 + 2x_2 + x_3 = 0 \\ 3x_1 + x_2 + 4x_3 = -7 \\ x_1 + x_2 + 3x_3 = -5 \end{cases}$$

Para aplicar o Teorema 3.2.2 da decomposição  $A = LU$  precisamos primeiramente fazer o determinante dos menores principais de  $A$  e verificar se são todos diferentes de zero. Consideremos

$$A = \begin{bmatrix} 5 & 2 & 1 \\ 3 & 1 & 4 \\ 1 & 1 & 3 \end{bmatrix}$$

Como

$$\det(A_1) = [5] = 5 \neq 0$$

$$\det(A_2) = \begin{vmatrix} 5 & 2 \\ 3 & 1 \end{vmatrix} = 5 - 6 = -1 \neq 0$$

$$\det(A_3) = \begin{vmatrix} 5 & 2 & 1 \\ 3 & 1 & 4 \\ 1 & 1 & 3 \end{vmatrix} = 19 - 39 = -20 \neq 0$$

Pelo Teorema 3.2.2,  $A$  se decompõe unicamente no produto  $LU$

$$\begin{bmatrix} 5 & 2 & 1 \\ 3 & 1 & 4 \\ 1 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

Para encontrar os elementos das matrizes  $L$  e  $U$ , vamos utilizar as fórmulas do Capítulo 3 da Subseção 3.2.2. Fazendo os cálculos obtemos as seguintes matrizes:

$$\begin{bmatrix} 5 & 2 & 1 \\ 3 & 1 & 4 \\ 1 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{3}{5} & 1 & 0 \\ \frac{1}{5} & -3 & 1 \end{bmatrix} \begin{bmatrix} 5 & 2 & 1 \\ 0 & -\frac{1}{5} & \frac{17}{5} \\ 0 & 0 & 13 \end{bmatrix}$$

Agora vamos determinar a solução do sistema dado

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

Como  $A = LU$ , segue que

$$\left( \begin{bmatrix} 1 & 0 & 0 \\ \frac{3}{5} & 1 & 0 \\ \frac{1}{5} & -3 & 1 \end{bmatrix} \begin{bmatrix} 5 & 2 & 1 \\ 0 & -\frac{1}{5} & \frac{17}{5} \\ 0 & 0 & 13 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -7 \\ -5 \end{bmatrix}.$$

Note que se

$$Ax = b$$

então

$$(LU)x = b$$



e

$$L(Ux) = b$$

Tomando  $Ux = y$  temos  $Ly = b$ . Assim,

$$\begin{bmatrix} 1 & 0 & 0 \\ \frac{3}{5} & 1 & 0 \\ \frac{1}{5} & -3 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -7 \\ -5 \end{bmatrix},$$

ou seja,

$$\begin{cases} y_1 & = & 0 \\ \frac{3}{5}y_1 + y_2 & = & -7 \\ \frac{1}{5}y_1 - 3y_2 + y_3 & = & -5 \end{cases}$$

e

$$y = \begin{bmatrix} 0 \\ -7 \\ -26 \end{bmatrix}.$$

Agora vamos resolver o sistema  $Ux = y$

$$\begin{bmatrix} 5 & 2 & 1 \\ 0 & -\frac{1}{5} & \frac{17}{5} \\ 0 & 0 & 13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -7 \\ -26 \end{bmatrix},$$

ou seja,

$$\begin{cases} 5x_1 + 2x_2 + x_3 & = & 0 \\ -\frac{1}{5}x_2 + \frac{17}{5}x_3 & = & -7 \\ 13x_3 & = & -26 \end{cases}.$$

Segue que

$$x = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}.$$

Note que tanto no Exemplo (3.2.1) quanto no Exemplo (3.2.5) o resultado é o mesmo.

Se  $A = LU$  e  $Ax = b$ , então  $Ly = b \Rightarrow y = L^{-1}b$  e  $Ux = y$  é sistema triangular superior, com

$$L^{-1} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -l_{21} & 1 & 0 & \dots & 0 \\ -l_{31} & -l_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -l_{n1} & -l_{n2} & -l_{n3} & \dots & 1 \end{bmatrix}.$$

Cabe notar que, o Método de Newton Modificado é da forma

$$\begin{cases} Jf(x_0) \cdot s = -f(x_n) \\ x_{n+1} = s + x_n \end{cases},$$

para algum  $x_0 \in D(f)$ .

Quando  $Jf(x_0) = A = LU$ , esse método tem forma muito mais simples de implementar, por que o sistema a ser resolvido, em cada iteração toma a forma

$$\begin{cases} Us = -L^{-1}f(x_n) \\ x_{n+1} = s + x_n \end{cases},$$

que é triangular superior.

Veja o Exemplo a seguir deste método.

**Exemplo 3.2.6** *Considere o polinômio  $p(x) = x^4 + x^3 + x^2 + x + 1$ . Uma forma de aproximar as quatro raízes complexas de  $p(x)$  é forçando a fatoração dele como*

$$p(x) = x^4 + x^3 + x^2 + x + 1 = (x^2 - ax - b)(x^2 - cx - d),$$

com  $a$ ,  $b$ ,  $c$  e  $d$  constantes a determinar. Uma multiplicação distributiva e ajustes na ordem das equações produzem as igualdades

$$\begin{cases} a + c + 1 = 0 \\ bd - 1 = 0 \\ -ac + b + d + 1 = 0 \\ ad + bc - 1 = 0 \end{cases} \quad (3.8)$$

Utilizando o Método de Newton

$$\begin{cases} JF(u_n) \cdot s = -F(u_n) \\ x_{n+1} = s + x_n \end{cases},$$

para aproximação da solução do sistema (3.8) da forma  $F(u) = 0$ , podemos determinar os valores aproximados de  $a$ ,  $b$ ,  $c$  e  $d$ . Escolhendo como aproximação inicial  $u_0 = (-2, -1, 0, -1)$  e  $JF(u_n)$  como a matriz do Jacobiano de  $F$  em  $u_n$ , em que

$$F(u) = \begin{bmatrix} a + c + 1 \\ bd - 1 \\ -ac + b + d + 1 \\ ad + bc - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad u = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

Logo,

$$JF(u) = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & d & 0 & b \\ -c & 1 & -a & 1 \\ d & c & b & a \end{bmatrix}.$$

Assim, vamos decompor  $JF(u_0)$  em  $LU$  e resolver

$$\begin{cases} U \cdot s = -L^{-1}F(u_n) \\ u_{n+1} = s + u_n \end{cases}. \quad (3.9)$$

Então,

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

ou melhor,

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix} \cdot \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} = - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

e

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} = \begin{bmatrix} 0,5 \\ 0 \\ 0,5 \\ 0 \end{bmatrix}.$$

Daí,

$$u_1 = \begin{bmatrix} 0,5 \\ 0 \\ 0,5 \\ 0 \end{bmatrix} + \begin{bmatrix} -2 \\ -1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -1,5 \\ -1 \\ 0,5 \\ -1 \end{bmatrix}$$

e

$$F(u_1) = \begin{bmatrix} 0 \\ 0 \\ -0,25 \\ 0 \end{bmatrix}.$$

Agora iremos decompor  $JF(u_1)$  em LU.

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \\ -0,5 & 1 & 1,5 & 1 \\ -1 & 0,5 & -1 & -1,5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -0,5 & -1 & 1 & 0 \\ -1 & -0,5 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

ou melhor,

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix} \cdot \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} = - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0,5 & 1 & 1 & 0 \\ 1 & 0,5 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ -0,25 \\ 0 \end{bmatrix}$$

e

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} = \begin{bmatrix} -0,125 \\ 0 \\ 0,125 \\ 0 \end{bmatrix}.$$

Logo,

$$u_2 = \begin{bmatrix} -0,125 \\ 0 \\ 0,125 \\ 0 \end{bmatrix} + \begin{bmatrix} -1,5 \\ -1 \\ 0,5 \\ -1 \end{bmatrix} = \begin{bmatrix} -1,625 \\ -1 \\ 0,625 \\ -1 \end{bmatrix}$$

e

$$F(u_2) = \begin{bmatrix} 0 \\ 0 \\ 0,0156 \\ 0 \end{bmatrix}.$$

Agora iremos decompor  $JF(u_2)$  em  $LU$ .

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \\ -0,625 & 1 & 1,625 & 1 \\ -1 & 0,625 & -1 & -1,625 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -0,625 & -1 & 1 & 0 \\ -1 & -0,625 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 2,25 & 0 \\ 0 & 0 & 0 & -2,25 \end{bmatrix}$$

ou melhor,

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 2,25 & 0 \\ 0 & 0 & 0 & -2,25 \end{bmatrix} \cdot \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} = - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0,625 & 1 & 1 & 0 \\ 1 & 0,625 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 0,0156 \\ 0 \end{bmatrix}$$

e

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} = \begin{bmatrix} 0,0069 \\ 0 \\ -0,0069 \\ 0 \end{bmatrix}.$$

Segue que,

$$u_3 = \begin{bmatrix} 0,0069 \\ 0 \\ -0,0069 \\ 0 \end{bmatrix} + \begin{bmatrix} -1,625 \\ -1 \\ 0,625 \\ -1 \end{bmatrix} = \begin{bmatrix} -1,618 \\ -1 \\ 0,618 \\ -1 \end{bmatrix}$$

e

$$F(u_3) = \begin{bmatrix} 0 \\ 0 \\ -0,000058 \\ 0 \end{bmatrix}.$$

Segundo o procedimento, obtemos  $a = -1,61806$ ,  $b = -1$ ,  $c = 0,61806$  e  $d = -1$ . Para

determinar as raízes, utilizamos o método hindu para resoluções de equações quadráticas (Fórmula de Bhaskara) e obtivemos os seguintes valores  $x_1 = -0,809017 + 0,5877853i$ ,  $x_2 = -0,809017 - 0,5877853i$ ,  $x_3 = 0,309017 + 0,9510565i$  e  $x_4 = 0,309017 - 0,9510565i$ .

No Apêndice (A.5) encontra-se o algoritmo desse Exemplo. Primeiramente fizemos um algoritmo para escalonar uma matriz e um para resolver um sistema triangular superior. Em seguida colocamos os dados do problema com o número de iterações qualquer, por se tratar de um software escolhemos um número de iterações como sendo  $n = 100$ , ou seja, o software repetirá os cálculos 100 vezes. Depois fizemos um algoritmo para resolver o sistema pelo Método de Newton. Para determinar as raízes, escrevemos por coordenadas, a “Fórmula de Bhaskara”.

Na expressão (3.9) desse problema, poderíamos ter resolvido o problema pelo Método de Newton modificado, onde a cada iteração substituímos apenas o valor  $F(u_n)$ , e trocamos  $JF(u_n)$  por  $JF(x_0)$ , não precisando escalonar a matriz em cada iteração.

Cabe notar que as raízes dos polinômios são autovalores da matriz companheira do polinômio  $p(x)$ ,

$$\text{Matriz Companheira} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & -1 & -1 & -1 \end{bmatrix}.$$

**Exemplo 3.2.7** Neste exemplo, vamos considerar o problema

$$\left| \begin{array}{l} \max \quad f(x, y) = e^{(xy)} \cos(x) \sin(y) \\ \text{sujeito a} \quad \begin{cases} 0 \leq x \leq \frac{\pi}{2} \\ 0 \leq y \leq \pi \end{cases} \end{array} \right. .$$

Alguns resultados deste exemplo foram obtidos com o uso do algoritmo do Apêndice (A.6).

A Figura (3.2) mostra que o valor máximo de  $f(x, y)$  ocorre no interior do domínio de  $f$ , ou seja, para  $0 \leq x \leq \frac{\pi}{2}$  e  $0 \leq y \leq \pi$ .

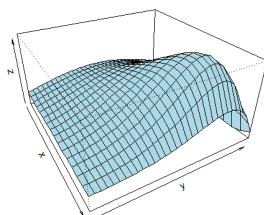


Figura 3.2: Gráfico do Exemplo (3.2.7) gerado pelo software.

Para encontrar a solução do problema vamos aplicar o Método de Newton e resolver a equação  $\nabla f(x, y) = (0, 0)$ .

Seja

$$\begin{aligned} g(x, y) &= \nabla f(x, y) \\ &= \begin{bmatrix} \frac{\partial f(x, y)}{\partial x} \\ \frac{\partial f(x, y)}{\partial y} \end{bmatrix}, \end{aligned}$$

com  $0 \leq x \leq \frac{\pi}{2}$  e  $0 \leq y \leq \pi$ . No Apêndice (A.6) chamamos de A.

Então a equação que precisamos resolver é:

$$g(x, y) = \begin{bmatrix} ye^{(xy)}\cos(x)\sen(y) - e^{(xy)}\sen(x)\sen(y) \\ xe^{(xy)}\cos(x)\sen(y) + e^{(xy)}\cos(x)\cos(y) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Pelo Método de Newton temos a sequência

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} x_n \\ y_n \end{bmatrix} - [Jg(x_n, y_n)]^{-1} \begin{bmatrix} x_n \\ y_n \end{bmatrix},$$

para  $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$  dados no domínio de  $g$ .

Cabe notar que  $Jg(x, y) = Hf(x, y)$

$$= \begin{bmatrix} \frac{\partial^2 f(x, y)}{\partial x^2} & \frac{\partial^2 f(x, y)}{\partial x \partial y} \\ \frac{\partial^2 f(x, y)}{\partial x \partial y} & \frac{\partial^2 f(x, y)}{\partial y^2} \end{bmatrix},$$

com

$$\begin{aligned} \frac{\partial^2 f(x, y)}{\partial x^2} &= e^{(xy)}y^2\sen(y)\cos(x) - e^{(xy)}2y\sen(x)\sen(y) - e^{(xy)}\sen(y)\cos(x) \\ \frac{\partial^2 f(x, y)}{\partial y^2} &= e^{(xy)}x^2\sen(y)\cos(x) + e^{(xy)}2x\cos(x)\cos(y) - e^{(xy)}\sen(y)\cos(x) \\ \frac{\partial^2 f(x, y)}{\partial x \partial y} &= e^{(xy)}xy\sen(y)\cos(x) + e^{(xy)}y\cos(x)\cos(y) - e^{(xy)}x\sen(x)\sen(y) + \\ &\quad + e^{(xy)}\sen(y)\cos(x) - e^{(xy)}\sen(x)\cos(y) \end{aligned}$$

que chamamos de B no Apêndice (A.6).

Tomando  $n = 100$ ,  $x_0 = 1$ ,  $y_0 = 2,5$  e  $u = (x, y)$  obtemos os seguintes valores

$$u_{100} = (x_{100}, y_{100}) = (1.1817391495970153, 2.4393027452826881) \quad e, \\ f(u_{100}) = 4.3764078775315118.$$

Por último, calculamos a matriz  $Hf(u_{100})$  que possui autovalores  $-31.335616$  e  $-9.569378$ . Como vimos no Capítulo 2, segue que  $f(u_{100})$  é aproximadamente valor de máximo de  $f(x)$ .



# Capítulo 4

## Problemas de otimização

No Capítulo 3 vimos uma forma de encontrar pontos críticos de uma função resolvendo a equação  $\nabla f(x) = 0$ , para  $x \in D(f)$ , para o caso em que  $D(f) \subset \mathbb{R}^m$  é um conjunto aberto.

Neste Capítulo voltamos ao problema de encontrar máximos e mínimos de funções. Porém, aqui veremos o caso em que esses pontos não estão no interior do domínio, ou seja, estão na fronteira do domínio da função. Assim, vamos tratar de casos específicos, onde o problema de otimização é dado por restrições que definem o domínio, dando uma forma particular da fronteira do domínio. Vamos estudar formas de determinar os pontos de máximo e mínimo de funções quando estão com restrições. Desse modo, vemos que quando as restrições possuem sinais de igualdade, o problema pode ser tratado com os Multiplicadores de Lagrange ou com as condições de KKT quando as restrições possuem sinais de desigualdades.

No caso em que a função  $f$  e as restrições  $g_i$  são todas lineares existem métodos eficientes para a resolução do problema e muitos softwares prontos para resolver o problema, por exemplo, o Solver do Excel, o software R e o Lindo Systems (veja por exemplo [18] para técnicas de resolução).

Vamos trabalhar com problemas de otimização da função  $f : D \rightarrow \mathbb{R}$ , onde  $D \subset \mathbb{R}^m$  é obtido por meio de equações e inequações envolvendo as funções  $g_i : \mathbb{R}^m \rightarrow \mathbb{R}$ . Logo, o problema de otimização que temos interesse é da forma

$$\left| \begin{array}{l} \text{otimizar} \\ \text{sujeito a} \end{array} \right. \begin{array}{l} f(x) \\ g_i(x) \sim 0, \quad i = 1, 2, \dots, p \end{array} \quad ,$$

onde “ $\sim$ ” seja “ $=$ ”, “ $\leq$ ” ou “ $\geq$ ” e  $x \in \mathbb{R}^m$ .

A função objetivo é a representação matemática do critério de eficiência adotado no problema de otimização e é influenciada pelas variáveis de controle do problema. Já as restrições são equações de igualdade ou desigualdade que determinam os limites de viabilidade do projeto, expressando uma condição desejável do comportamento do sistema.

Geralmente, as restrições estão relacionadas com a geometria, os esforços admissíveis, os recursos disponíveis, os custos envolvidos no projeto, dentre outros.

Um método muito utilizado e visto superficialmente nas disciplinas de Cálculo Diferencial e Integral é o Método dos Multiplicadores de Lagrange, (veja [2] e [12]).

## 4.1 Multiplicadores de Lagrange

O Método dos Multiplicadores de Lagrange se aplica ao problema de determinar os valores extremos com restrições de igualdade, por exemplo, resolver o problema

$$\left| \begin{array}{ll} \min & f(x) \\ \text{sujeito a} & g_i(x) = 0, \end{array} \right. , \quad i = 1, 2, \dots, p$$

Cabe notar que  $\min f(x) = \max -f(x)$ .

**Teorema 4.1.1** *Sejam  $f : D \rightarrow \mathbb{R}$  e  $g_i : D \rightarrow \mathbb{R}$ , com  $D \subset \mathbb{R}^m$  aberto e  $i = 1, 2, \dots, p$ . Se  $f, g_i$  são funções deriváveis com derivadas contínuas. Então o ponto  $a \in D$  é solução do problema*

$$\left| \begin{array}{ll} \min & f(x) \\ \text{sujeito a} & g_i(x) = 0, \end{array} \right. , \quad i = 1, 2, \dots, p$$

*se, e somente se, existem  $\lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R}$  tais que*

$$\nabla f(a) = \sum_{i=1}^p \lambda_i \nabla g_i(a).$$

**Demonstração:** Suponha inicialmente que  $p = 1$  e que  $f$  tem valor extremo no ponto  $a$  e que  $\nabla f(a) \neq 0$ . Seja  $S$  a superfície, de dimensão  $m - 1$ , formada pelo conjunto dos pontos  $x \in D$  tais que  $g_1(x) = 0$ . Suponha que  $a$  esteja sobre a superfície  $S$  e seja  $C$  uma curva sobre  $S$ , com equação paramétrica  $r(t) = (x_1(t), x_2(t), \dots, x_n(t))$  derivável, para  $t \in I \subset \mathbb{R}$ ,  $I$  intervalo, que passe pelo ponto  $a$ , ou seja, com  $r(t_0) = a$ , para  $t_0 \in I$ . Seja  $h(t) = (f \circ r)(t)$  que é derivável. Pela Regra da Cadeia,

$$\begin{aligned} h'(t) &= f'(r(t)) \cdot r'(t) \\ &= \nabla f(r(t)) \cdot r'(t). \end{aligned}$$

Como  $a$  é um ponto crítico de  $f$  em  $S$ , temos que  $h(t)$  tem extremo em  $t_0$ , e portanto, pelo Lema 2.1.6,

$$h'(t_0) = 0,$$

Daí,

$$\nabla f(a) \cdot r'(t_0) = 0$$

e segue que  $\nabla f(a)$  é perpendicular a  $r'(t_0)$ .

Como  $r(t) \in S$  temos que  $g_1(r(t)) = 0$  e portanto, pela Regra da Cadeia,

$$(g_1 \circ r)'(t) = \nabla g_1(r(t)) \cdot r'(t) = 0,$$

e  $\nabla g_1(a)$  é também perpendicular a  $r'(t_0)$ .

Como  $S$  tem dimensão  $m - 1$ , segue que  $\nabla f(a)$  e  $\nabla g_1(a)$  são paralelos. Logo, existe  $\lambda \in \mathbb{R}$  tal que,

$$\nabla f(a) = \lambda \nabla g_1(a).$$

Agora, sejam  $S_i = \{x \in D : g_i(x) = 0\}$  superfícies de dimensão  $m - 1$  e  $S = S_1 \cap S_2 \cap \dots \cap S_p$ . Seja  $r(t)$  uma curva em  $S$ , com  $r(t_0) = a$  e  $h(t) = (f \circ r)(t)$ . Suponhamos que  $f$  tenha um ponto de extremo em  $a \in S$  e que  $\nabla f(a) \neq 0$ .

Como,  $V_i = \{v \in \mathbb{R}^m : v \perp \nabla g_i(a)\}$  e o conjunto dos vetores de  $\mathbb{R}^m$  perpendiculares a  $\nabla g_i(a)$  e  $V = \bigcap V_i$ .

Segue

$$\begin{cases} \nabla f(a) \perp r'(t_0) \\ \nabla g_i(a) \perp r'(t_0) \end{cases} \quad (\text{isso vale se } r(t) \text{ fosse curva em } S_i)$$

Logo,  $r'(t_0) \in V_i$  e  $\nabla f(a) \perp (\bigcap V_i)$ . Assim,  $\nabla f(a)$  é perpendicular a todo vetor perpendicular a cada  $\nabla g_i(a)$ . Como,

$$\begin{aligned} \mathbb{R}^m &= [\{\nabla g_i(a)\}] \oplus [\{\nabla g_i(a)\}]^\perp \\ &= [\{g_i(a)\}] \oplus V \end{aligned}$$

(veja [3]) temos que  $\nabla f(a) \in [\{\nabla g_i(a)\}]$  e existem  $\lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R}$  tais que

$$\nabla f(a) = \sum_{i=1}^p \lambda_i \nabla g_i(a) \tag{4.1}$$

■

### 4.1.1 Método dos Multiplicadores de Lagrange

Sabemos que problemas de otimização estão relacionados a maximizar ou minimizar uma função em relação a uma variável  $x$ . Para resolver esse tipo de problema utilizando

os Multiplicadores de Lagrange é preciso que todas as restrições possuam igualdade.

Dessa forma, se houver restrições nos possíveis valores de  $x$ , o Método de Multiplicadores de Lagrange pode restringir a busca de soluções no conjunto de valores viável de  $x$ . Assim, o problema pode ser escrito da seguinte forma

$$\left\{ \begin{array}{l} \text{otimizar} \\ \text{sujeito a} \end{array} \right. \quad \begin{array}{l} f(x) \\ g_i(x) = 0, \quad i = 1, 2, \dots, p. \end{array}$$

Para resolver esse tipo de problema utilizando Multiplicadores de Lagrange, devemos primeiramente colocar a função objetivo e as restrições dadas em um único problema e como mostramos no início dessa Seção, teremos que multiplicar cada gradiente de cada restrição por um fator  $\lambda_i$  como mostra a equação (4.1). Então, o problema poderá ser visto como a equação

$$\left\{ \begin{array}{l} \nabla f(x) + \sum_{i=1}^p \lambda_i \nabla g_i(x) = 0 \\ g_i(x) = 0 \end{array} \right. \quad (4.2)$$

E podemos aplicar métodos numéricos para a aproximação da solução  $(x, \lambda)$  de (4.2), onde  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_p)$ , como por exemplo o Método de Newton.

Veja o Exemplo a seguir.

**Exemplo 4.1.2** *Resolver o problema*

$$\left\{ \begin{array}{l} \min \\ \text{sujeito a} \end{array} \right. \quad \begin{array}{l} f(x, y) = 3x + 4y \\ g(x, y) = x^2 + y^2 - 1 = 0 \end{array}$$

*utilizando Multiplicadores de Lagrange.*

*Por esse método precisamos resolver a equação*

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

*Dessa forma,*

$$\begin{aligned} (3, 4) &= \lambda(2x, 2y) \\ &= (2x\lambda, 2y\lambda) \end{aligned}$$

*e*

$$2x\lambda = 3, \quad 2y\lambda = 4.$$

*Se  $\lambda \neq 0$  segue que*

$$x = \frac{3}{2\lambda} \quad e \quad y = \frac{2}{\lambda}.$$

Substituindo  $x$  e  $y$  em  $g(x, y)$  temos

$$\left(\frac{3}{2\lambda}\right)^2 + \left(\frac{2}{\lambda}\right)^2 = 0$$

que produz  $\lambda = \pm \frac{5}{2}$ . Logo,  $x = \pm \frac{3}{5}$  e  $y = \pm \frac{4}{5}$ . Substituindo os valores de  $x$  e  $y$  na função objetivo, temos que,  $\left(-\frac{3}{5}, -\frac{4}{5}\right)$  é o ponto de mínimo.

Uma aplicação imediata desse resultado é numa outra forma de demonstrar o Teorema 2.2.2, como o que segue.

**Demonstração:** [Teorema 2.2.2] Considere o problema:

$$\left\{ \begin{array}{l} \text{Maximizar } \psi(x) = \langle Hx, x \rangle \\ \text{sujeito a } g(x) = 1 \end{array} \right.,$$

com  $g(x) = \langle x, x \rangle$  e  $H = H^*$  transformação linear. Escolhendo  $u_1$  como solução desse problema, que existe pelo Teorema 2.1.4 e usando o Método dos multiplicadores de Lagrange obtemos a igualdade  $\psi'(u_1) = \lambda_1 g'(u_1)$ . Cálculos diretos mostram que

$$\psi'(u_1)h = 2\langle Hu_1, h \rangle = 2\lambda_1 \langle u_1, h \rangle = 2\lambda_1 g'(u_1)h, \quad h \in E,$$

uma vez que  $H = H^*$ . Disso segue que

$$\langle Hu_1 - \lambda_1 u_1, h \rangle = 0,$$

e que

$$(H - \lambda_1 I)u_1 = 0,$$

ou seja,  $Hv_1 = \lambda_1 u_1$  e  $\lambda_1$  é um autovalor de  $H$ , com autovetor  $u_1$ .

Para terminar a demonstração proceda como no final da demonstração do Teorema 2.2.2.

## 4.2 Condições de Karush-Kuhn-Tucker (KKT)

As Condições de KKT são uma generalização do Método de Multiplicador de Lagrange, para problemas com restrições com desigualdades.

Com as condições de KKT, podemos encontrar os candidatos à pontos mínimos de uma função. Por questões teóricas optamos por não apresentar a demonstração desse

resultado. Uma referência introdutória do assunto é [19, p. 31] (Teorema 3.1.14).

**Definição 4.2.1** Uma função  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}$  é **convexa** quando  $D$  for convexo e, para quaisquer  $x, y \in D$  e  $\alpha, \beta \in [0, 1]$  com  $\alpha + \beta = 1$ , têm-se  $f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y)$ .

Pode-se mostrar que se  $f : D \rightarrow \mathbb{R}$  é uma função convexa, então  $f$  é contínua em  $D$  (veja [14, p. 79]). Além disso, todo ponto crítico  $a$  de uma função convexa  $f : D \rightarrow \mathbb{R}$  de classe  $C^1$  é um ponto de mínimo global, isto é,  $f(x) \geq f(a)$  para todo  $x \in D$ .

**Teorema 4.2.2** Sejam  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $g_i : D \rightarrow \mathbb{R}, i = 1, 2, \dots, p$  e  $h_j : D \rightarrow \mathbb{R}, j = 1, 2, \dots, q$ , funções convexas e diferenciáveis com derivadas contínuas. Se existem  $\lambda = (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p$  e  $\mu = (\mu_1, \dots, \mu_q) \in \mathbb{R}^q$  tais que

$$\begin{aligned} \nabla f(a) + \sum_{i=1}^p \lambda_i \nabla g_i(a) + \sum_{j=1}^q \mu_j \nabla h_j(a) = 0 \\ \lambda_i \geq 0 \\ \lambda_i g_i(a) = 0 \\ g_i(a) \leq 0 \\ h_j(a) = 0 \end{aligned} \quad (4.3)$$

então  $a$  é solução do PPNL

$$\left\{ \begin{array}{l} \min \\ \text{sujeito a} \end{array} \right. \begin{array}{l} f(x) \\ \left\{ \begin{array}{l} g_i(x) \leq 0 \\ h_j(x) = 0 \end{array} \right. \end{array} \quad (4.4)$$

A Expressão (4.3) é chamada de condição de KKT.

Existem condições em que esse resultado é escrito como condição necessária e suficiente, por exemplo quando  $h_i$  e  $g_i$  são lineares,  $g_i(x) = A_i(x) - b_i$  e  $f$  é convexa.

**Corolário 4.2.3** Seja  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}$  uma função convexa e diferenciável, com derivada contínua ( $C^1$ ),  $h_j(x) = H_j x + b_j$ ,  $g_i(x) = G_i x + c_i$ , com  $H_j$  e  $G_i$  matrizes  $m \times m$  e  $b_j, c_i$  vetores,  $i = 1, 2, \dots, p$  e  $j = 1, 2, \dots, q$ . Existem  $\lambda \in \mathbb{R}^p$  e  $\mu \in \mathbb{R}^q$  tais que

$$\left\{ \begin{array}{l} \nabla f(a) + \sum_{i=1}^p \lambda_i \nabla g_i(a) + \sum_{j=1}^q \mu_j \nabla h_j(a) = 0 \\ \lambda_i \geq 0 \\ \lambda_i g_i(a) = 0 \\ g_i(a) \leq 0 \\ h_j(a) = 0 \end{array} \right. ,$$

se, e somente se,  $a$  é solução de (4.4).

Existem condições mas gerais de KKT sem a necessidade de  $f$  ser convexa (veja [19, p. 22] Teorema 2.4.1, por exemplo).

Podemos observar que, por exemplo, se  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}$  for de classe  $C^2$  e tiver um ponto de mínimo local em  $a \in \overset{\circ}{D}$  (único numa vizinhança de  $a$ ), então

$$f(x) = f(a) + \langle Hf(a) \cdot (x - a), (x - a) \rangle + r(x - a),$$

com  $\lim_{h \rightarrow 0} \frac{r(h)}{|h|^2} = 0$ . Segue que  $Hf(x)$  é positiva, para  $x \simeq a$ , e  $f$  é convexa, em torno de  $a$ , (veja Teorema 3.1.9 em [19, p.29]).

### 4.3 Métodos Tipo Gradiente

O Método de Newton não é sempre convergente, pois precisamos de uma boa aproximação para a solução da equação, além de lidar com o problema de arredondamento e resolução de sistemas lineares. Assim, vamos trabalhar uma outra possibilidade de obtermos estimativas para o ponto de máximo ou mínimo de funções, podendo ser aplicado tanto quando o gradiente é igual a zero como no caso em que têm-se Multiplicadores de Lagrange ou Condições de KKT.

O Método dos Gradientes nos permite trabalhar com problemas que possuem vários máximos ou mínimos locais, e se tivermos uma posição e aplicamos o método, este nos direcionará para um local melhor, podendo este local não ser o ponto crítico de fato, no entanto, melhora o resultado do problema.

Um dos objetivos desta Seção é melhorar uma estimativa  $x_0$  para aproximar pontos de máximo (ou de mínimo) local.

Considere o seguinte problema a ser otimizado:

$$\left| \begin{array}{ll} \min & f(x) \\ \text{sujeito a} & g_i(x) \leq 0 \quad i = 1, \dots, p, \\ & h_j(x) = 0 \quad j = 1, \dots, q \end{array} \right.$$

com  $f$  e  $g$  diferenciáveis, com  $x \in D \subset \mathbb{R}^m$ . A ideia é que dado um  $x_0 \in D(f)$  e um valor  $f(x_0)$ , encontrar  $x_1$  de modo que  $f(x_1) < f(x_0)$ .

Assim, vamos começar supondo que o problema seja:

$$\left| \begin{array}{ll} \min & f(x) \\ \text{sujeito a} & x \in D \subset \mathbb{R}^m \end{array} \right.$$

com  $D$  aberto.

Sabemos do Cálculo Diferencial e Integral que dado  $x_0 \in D$ , a direção da maior decréscimo em  $x_0$  é dada por  $-\nabla f(x_0)$ . Isso nos leva a descobrir uma curva  $r(t)$  tal

que  $r(0) = x_0$  e  $r'(t) = -\nabla f(r(t))$ , pois esse será o caminho que leva ao menor valor de  $f(x)$  quando partimos do ponto  $x_0$ , (veja Figura 4.1). Assim,  $r(t)$  dará melhores estimativas para  $\min f(x) = f(a)$ , caso exista e caso  $f(x_0)$  não seja um mínimo local. No entanto, para discutir formas de aproximar ou determinar  $r(t)$  é necessário um pouco de teoria de Equações Diferenciais Ordinárias, com base em [20].

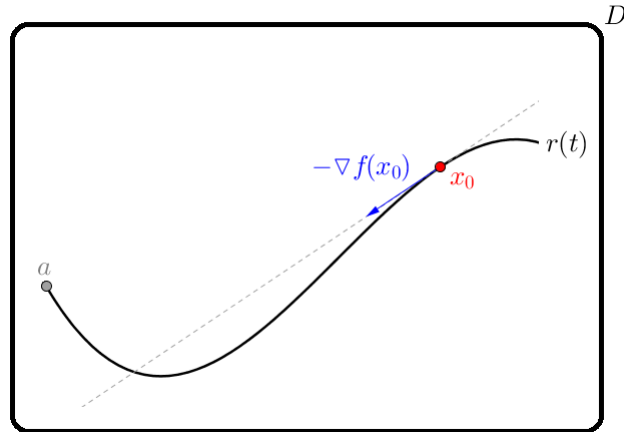


Figura 4.1: Explicação gráfica do problema.

### 4.3.1 Introdução a Equações Diferenciais Ordinárias

Vamos considerar o problema de determinar uma função  $y(t)$ , para  $t$  em um intervalo  $I$ , que contém um ponto  $t_0$  no seu interior, ou seja, uma função  $y : I \rightarrow \mathbb{R}^m$  tal que

$$\begin{cases} y' = f(t, y) \\ y(t_0) = y_0 \end{cases}, \quad (4.5)$$

onde  $f : I \times D \rightarrow \mathbb{R}^m$ ,  $D \subset \mathbb{R}^m$  é aberto e  $y(t) \in D$ . Esse problema é chamado Problema de Valor Inicial (PVI).

Se  $f(t, y)$  for contínua em  $I \times D$  então

$$y(t) = y_0 + \int_{t_0}^t f(s, y(s)) ds, \quad (4.6)$$

caso exista, por meio do Teorema Fundamental do Cálculo.

A expressão (4.6) juntamente com o Teorema de Arzelà-Ascoli permitem demonstrar o teorema a seguir, (veja [16] e [20])

**Teorema 4.3.1** (Teorema de Peano) *Seja  $f(t, y)$  definida e contínua em  $I \times D$  que contém o ponto  $(t_0, y_0)$ . Então o PVI (4.5) admite solução contínua e diferenciável, para todo  $t \in I$ .*



Cabe notar que se  $f(t, y)$  for contínua existe solução  $y(t)$  para (4.6), mas essa pode não ser única, por meio do Teorema de Peano.

Po outro lado, a unicidade é garantida adicionando ao enunciado do teorema as condições do resultado a seguir.

**Corolário 4.3.2** *Nas condições do Teorema 4.3.1, se  $|f(t, y) - f(t, z)| \leq L|y - z|$ , então  $y(t)$  é única.*

**Demonstração:** Daremos aqui apenas uma ideia da demonstração.

Se

$$y(t) = y_0 + \int_{t_0}^t f(s, y(s)) ds \quad \text{e} \quad z(t) = y_0 + \int_{t_0}^t f(s, z(s)) ds$$

forem duas soluções de (4.5), então

$$\begin{aligned} |y(t) - z(t)| &\leq \left| \int_{t_0}^t |f(s, y(s)) - f(s, z(s))| ds \right| \\ &\leq |t - t_0| L \max |y(s) - z(s)|. \end{aligned}$$

Note que, se  $t \simeq t_0$ , então  $0 \leq |t - t_0| L < 1$ . Tomando  $|t - t_0| L = L_1$  temos que existe  $\varepsilon > 0$ , tal que

$$\max_{|t-t_0|<\varepsilon} |y(t) - z(t)| \leq L_1 \max_{|t-t_0|<\varepsilon} |y(t) - z(t)|$$

e segue que  $y(t) = z(t)$ , se  $|t - t_0| < \varepsilon$ . ■

**Observação 4.3.3** *Note que, se  $f(t, y)$  for contínua no conjunto aberto  $I \times D$  e existir  $\partial f(t, y)$  também contínua, então podemos reduzir  $I \times D$  para  $I_1 \times D_1$  de tal forma que  $(t_0, y_0) \in (I_1 \times D_1)$  e*

$$L = \max_{(t,y) \in (I_1 \times D_1)} \frac{\|\partial f(t, y)\|}{\partial y} < +\infty.$$

O Corolário anterior pode ser demonstrado diretamente utilizando o Teorema do Ponto Fixo de Banach 3.1.5.

Existem métodos para determinar resoluções de casos particulares do problema (4.5). (Veja, por exemplo [20] para o assunto.)

## 4.4 Métodos Numéricos para PVI

Uma propriedade importante dos métodos computacionais para solução (4.5) é a discretização, que consiste em obter a solução aproximada do PVI em um conjunto discreto

de pontos  $\{t_n; n = 0, 1, \dots, N\} \subset [a, b]$ .

A seqüência de pontos  $t_n$  é definida por:

$$t_n = t_0 + nh; \quad n = 0, 1, \dots, N.$$

onde  $t_0 = a, t_N = b$  e  $N = \frac{b-a}{h}$ . Com o “tamanho dos passos”  $h$  de um ponto ao outro o mais próximo possível de 0.

Veremos agora métodos para solução numérica de Equações Diferenciais Ordinárias baseados em [11] e [21].

### 4.4.1 Método de Taylor

Consideremos o PVI (4.5) tal que  $f$  seja contínua e suficientemente derivável em relação a  $t$  e  $y$ . Seja  $y(t)$  a solução do PVI (4.5). A expansão em série de Taylor (veja [12, p. 151]) para  $y(t_n + h)$  em torno do ponto  $t_n$  é dada por

$$y(t_n + h) = y(t_n) + hy'(t_n) + \frac{h^2}{2!}y''(t_n) + \dots + \frac{h^p}{p!}y^{(p)}(t_n) + \frac{h^{p+1}}{(p+1)!}y^{(p+1)}(\varepsilon_n) \quad (4.7)$$

com  $t_n < \varepsilon_n < t_n + h$ , onde o último termo é o erro de truncamento local.

Note que, para as primeiras derivadas, obtemos:

$$\begin{aligned} y' &= f(t, y) \\ y'' &= f' = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} \frac{dy}{dt} = f_t + f_y f \\ y''' &= f'' = \frac{\partial f_t}{\partial t} + \frac{\partial f_t}{\partial y} \frac{dy}{dt} + \left[ \frac{\partial f_y}{\partial t} + \frac{\partial f_y}{\partial y} \frac{dy}{dt} \right] f + f_y \left[ \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} \frac{dy}{dt} \right] = \\ &= f_{tt} + 2f_{ty}f + f_{yy}f^2 + f_y f_t + f_y^2 f \\ &\vdots \end{aligned}$$

O que dificulta esse método é que, a menos que  $f(t, y)$  seja uma função simples, as derivadas totais de ordem mais elevada tornam-se cada vez mais complexas. Desse modo, é preciso limitar o número de termos em (4.7). Assim, obtemos a seguinte expressão:

$$y(t_n + h) \simeq y(t_n) + hy'(t_n) + hf(t_n, y(t_n)) + \dots + \frac{h^p}{p!}f^{(p-1)}(t_n, y(t_n)) \quad (4.8)$$

A equação (4.8) pode ser interpretada como uma relação aproximada entre os valores da solução de (4.7). Uma relação exata entre valores aproximados da solução de (4.7) pode ser obtida por:

$$y_{n+1} = y_n + hf_n + \frac{h^2}{2!}f'_n + \dots + \frac{h^p}{p!}f_n^{(p-1)} \quad (4.9)$$

que é chamado Método de Taylor de ordem  $p$ , onde  $f_n = f(t_n, y_n)$  e  $y_n \simeq y(t_n)$ .

Esse método possui muita complexidade computacional, pois é preciso calcular as derivadas parciais de  $f$  a cada passo e avaliá-las nos pontos  $(t_n, y_n)$ , mas pode ser muito útil se as derivadas de  $f$  forem de fácil implementação.

#### 4.4.2 Método de Euler

O Método de Euler foi criado por volta de 1768 e é conhecido também como Método da Reta Secante. É baseado na expansão da função  $y(t)$  em séries de Taylor, podendo-se expandir a função  $y(t)$  na vizinhança do ponto  $t_n$ , até a ordem 1, podemos então escrever esse método como:

$$y_{n+1} = y_n + hf_n. \quad (4.10)$$

O Método de Euler aproxima o valor exato da função  $y(t_n)$  através de uma reta no ponto  $(t_n, y(t_n))$  com inclinação  $f_n$ . Veja a Figura 4.2 a seguir.

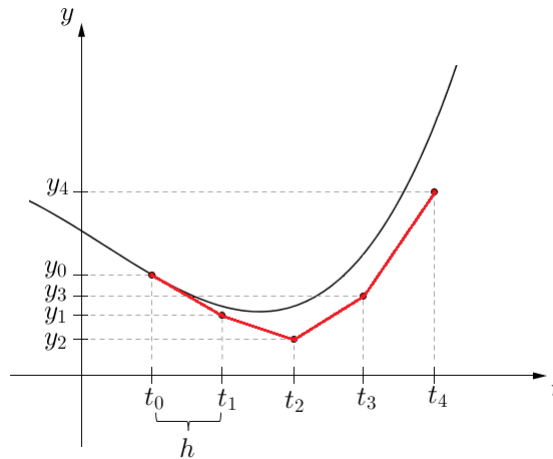


Figura 4.2: Método de Euler.

A curva em preto representa o valor exato da solução, já a curva em vermelho representa a solução aproximada pelo Método de Euler.

Veja o exemplo a seguir.

**Exemplo 4.4.1** *Resolva numericamente o Problema de Valor Inicial*

$$\begin{cases} y' = y - t \\ y(0) = 2 \end{cases} \quad (4.11)$$

em  $[0, 1]$ .

Tomemos  $N = 4$ , para exemplificar, temos  $h = \frac{1-0}{4} = 0,25$ . Ou seja,  $t_0 = 0$ ,  $t_1 = 0,25$ ,  $t_2 = 0,5$ ,  $t_3 = 0,75$  e  $t_4 = 1$ .

Sabemos que  $y_0 = 2$ . Vamos resolver esse problema com o Método de Euler, ou por passo simples.

Temos então

$$\begin{aligned} y_1 &= y_0 + hf(t_0, y_0) \\ &= 2 + 0,25 \cdot (2 - 0) \\ &= 2,5 \end{aligned}$$

$$\begin{aligned} y_2 &= y_1 + hf(t_1, y_1) \\ &= 2,5 + 0,25 \cdot (2,5 - 0,25) \\ &= 3,0625 \end{aligned}$$

$$\begin{aligned} y_3 &= y_2 + hf(t_2, y_2) \\ &= 3,0625 + 0,25 \cdot (3,0625 - 0,5) \\ &= 3,7031 \end{aligned}$$

$$\begin{aligned} y_4 &= y_3 + hf(t_3, y_3) \\ &= 3,7031 + 0,25 \cdot (3,7031 - 0,75) \\ &= 4,4684 \end{aligned}$$

Logo, temos que  $y_0 = 2$ ,  $y_1 = 2,5$ ,  $y_2 = 3,063$ ,  $y_3 = 3,703$  e  $y_4 = 4,468$ , são aproximações para  $y(0,25)$ ,  $y(0,5)$ ,  $y(0,75)$  e  $y(1)$ .

A solução analítica do PVI é  $y(t) = e^t + t + 1$ . Substituindo  $t$  nessa solução temos os valores exatos  $y_0 = 2$ ,  $y_{0,25} = 2,534$ ,  $y_{0,5} = 3,148$ ,  $y_{0,75} = 3,876$ ,  $y_1 = 4,718$ .

### 4.4.3 Método de Runge-Kutta

O Método de Runge-Kutta é um aprimoramento do Método de Euler, e consequentemente produz uma aproximação ainda melhor que a do Método de Euler.

A fórmula geral desse método de ordem é definida por:

$$y_{n+1} - y_n = h\phi(t_n, y_n, h), \quad (4.12)$$

em que,

$$\begin{aligned}
\phi(t, y, h) &= \sum_{r=1}^p c_r k_r & (4.13) \\
k_1 &= f(t, y), \\
k_r &= f\left(x + a_r h, y + h \sum_{s=1}^{r-1} b_{rs} k_s\right); \quad r = 2, 3, \dots, p, \\
a_r &= \sum_{s=1}^{p-1} b_{rs}; \quad s = 2, 3, \dots, p
\end{aligned}$$

O Método de Runge-Kutta baseia-se em, determinar constantes  $c_r$ ,  $a_r$  e  $b_{rs}$  de (4.12), comparando  $\phi(x, y, h)$  de (4.13), em potências de  $h$  com a função do Polinômio de Taylor (4.7). Assim, a ordem de Runge-Kutta é a ordem do Método de Taylor que utilizamos, mas não precisamos calcular  $f'$ ,  $f''$ ,  $\dots$ ,  $f^{(p-1)}$  (veja [11, p. 406] e [21, p. 385]).

Veja o Exemplo a seguir.

**Exemplo 4.4.2** *Resolva numericamente*

$$\begin{cases} y' = y - t \\ y(0) = 2 \end{cases} \quad (4.14)$$

em  $[0, 1]$ , pelo Método de Runge-Kutta de ordem 2.

No Apêndice (A.7) construímos primeiramente o algoritmo com as variáveis do Método de Runge-Kutta.

Assim, escrevemos os dados do problema começando com a equação, depois colocamos os extremos do intervalo chamando de  $a = 0$  e  $b = 1$  e o valor de  $y$  no ponto inicial. Criamos uma função para definir a derivada e definimos  $n$  que será a quantidade de partes que o intervalo será se dividido.

Tomando  $n = 5$  obtemos como resultado para  $y$ :  $y_0 = 2$ ,  $y_1 = 2,44$ ,  $y_2 = 2,93$ ,  $y_3 = 3,49$ ,  $y_4 = 4,12$  e  $y_5 = 4,85$ . Observe os gráficos da Figura (4.3) e note que tomando  $n = 100$ , o método faz uma aproximação mais exata da curva original.

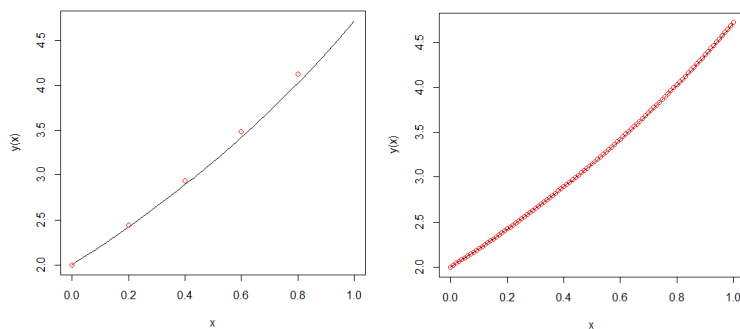


Figura 4.3: Gráfico gerado no software R com  $n = 5$  e  $n = 100$ , respectivamente.

## 4.5 Método de Declive

Como comentamos anteriormente,  $\nabla f(x)$  dá a direção do maior crescimento de  $f(x)$ , no ponto  $x$ , assim, conseqüentemente temos que  $-\nabla f(x)$  dá a direção do maior decréscimo de  $f(x)$  nesse ponto.

Sendo assim, podemos resolver o PVI

$$\begin{cases} x' = \nabla f(x) \\ x(t_0) = x_0 \end{cases}, \quad (4.15)$$

pelo Método de Euler e obter  $x_{i+1} = x_i + h\nabla f(x_i)$ , para aproximar o valor máximo de  $f(x)$  e resolver

$$\begin{cases} x' = -\nabla f(x) \\ x(t_0) = x_0 \end{cases}, \quad (4.16)$$

pelo Método de Euler e obter  $x_{i+1} = x_i - h\nabla f(x_i)$ , para aproximar o valor mínimo de  $f(x)$ , onde  $h$  é o tamanho do passo dado,  $i$  é o número de iterações e  $x_i \in \mathbb{R}^m$ .

O Método de Declive é uma maneira de aproximar máximos e mínimos locais de uma função utilizando gradientes da função.

**Observação 4.5.1** *Em um Método de Declive, a cada ponto obtido, escolhemos uma direção de descida para o gradiente, ou seja, um vetor de direção que pode proporcionar um decréscimo da função objetivo. Além disso, devemos definir a cada iteração um tamanho de passo  $h \in \mathbb{R}$  a ser dado na direção do declive.*

Este método funciona da seguinte maneira:

- 1- Primeiramente é preciso dar uma aproximação inicial  $x_0 \in D(f)$  da solução do problema e em seguida devemos resolver o gradiente da função neste ponto.
- 2- Depois, a partir do resultado encontrado em 1 é preciso determinar a solução na direção negativa ou positiva (dependendo do caso) do gradiente e repetir o processo até um dado critério de parada.

Certamente, podemos usar métodos Numéricos de ordem mais alta, como os de Runge-Kutta.

Suponhamos que queremos encontrar o mínimo de alguma função  $f(x)$ . Assim, dado um valor inicial  $x_0$  por  $x$ , podemos escolher várias direções para o declive. Como queremos encontrar um mínimo, basta tomar  $-\nabla f(x)$ .

A escolha do tamanho do passo  $h$  a ser utilizado na descida do gradiente, a partir de um ponto  $x_0$ , é outra decisão a ser tomada a cada iteração, pois nem sempre o passo que proporciona o maior decréscimo a cada iteração será a melhor escolha para resolver o

problema com um menor número de iterações. Um tamanho de passo errado  $h$  pode não alcançar a convergência, tomando  $h$  muito grande poderá divergir e muito pequeno levará muito tempo para convergir.

Neste método podemos tanto escolher o tamanho do passo fixo quanto mudar este a cada iteração. Vamos apresentar agora como fazer a escolha do tamanho máximo do passo para a convergência.

Uma condição para que esse método funcione é a Condição de Armijo (veja [19, p.101]).

**Teorema 4.5.2** (Condição de Armijo) - Sejam  $x, d \in \mathbb{R}^m$  tais que  $\nabla f(x) \neq 0, \nabla f(x) \cdot d < 0$  e  $h \in (0, 1)$ . Então existe  $\varepsilon = \varepsilon(h) > 0$  tal que

$$f(x + td) \leq f(x) + ht\nabla f(x) \cdot d,$$

para todo  $t \in (0, \varepsilon]$ .

**Demonstração:** Temos que

$$\nabla f(x) \cdot d = \lim_{t \rightarrow 0} \frac{f(x + td) - f(x)}{t}$$

Se  $\nabla f(x) \neq 0$  então temos que  $\nabla f(x) \cdot d \neq 0$ . Assim,

$$1 = \lim_{t \rightarrow 0} \frac{f(x + td) - f(x)}{t\nabla f(x) \cdot d}$$

Logo, existe  $\varepsilon > 0$  tal que para todo  $t \in (0, \varepsilon]$ ,

$$\frac{f(x + td) - f(x)}{t\nabla f(x) \cdot d} \geq h$$

Portanto, para todo  $t \in (0, \varepsilon]$ ,  $f(x + td) \leq f(x) + ht\nabla f(x) \cdot d$ . ■

Assim, o procedimento deste método é partir de um valor (suficientemente grande) do comprimento de passo  $h$  e reduzir para metade esse valor até que  $f(x_{k+1}) < f(x_k)$  seja satisfeita.

Veja o Exemplo a seguir.

**Exemplo 4.5.3** Minimizar  $f(x, y) = 3x^2 + 3xy + 2y^2 + x + y$ ,  $(x, y) \in \mathbb{R}^2$ , pelo Método de Declive com precisão de  $10^{-2}$ , dado  $(x_0, y_0) = (0, 1)$  como ponto inicial.

Primeiramente devemos calcular

$$\begin{aligned} \nabla f(x_0, y_0) &= \begin{bmatrix} 6x_0 + 3y_0 + 1 \\ 4y_0 + 3x_0 + 1 \end{bmatrix} \\ &= \begin{bmatrix} 4 \\ 5 \end{bmatrix}. \end{aligned}$$

Agora devemos determinar a solução na direção negativa do gradiente, pois queremos determinar  $\min f(x, y)$ . Desse modo, utilizando o Método de Euler, segue que

$$\begin{aligned} x_1 &= x_0 - h\nabla f(x_0) & e & & y_1 &= y_0 - h\nabla f(y_0) \\ &= -4h & & & &= 1 - 5h \end{aligned} \quad (4.17)$$

Substituindo  $x_1$  e  $y_1$  na função obtemos

$$\begin{aligned} f(x_1, y_1) &= 3x_1^2 + 3x_1y_1 + 2y_1^2 + x_1 + y_1 \\ &= 3(-4h)^2 + 3(-4h)(1 - 5h) + 2(1 - 5h) - 4h + 1 - 5h \\ &= 158h^2 - 41h + 3 \end{aligned} \quad (4.18)$$

Tomando, por exemplo,  $h = 0.1$  em (4.17) ou (4.18), verificamos que  $f(x_1, y_1) < f(x_0, y_0)$ , pois

$$f(x_0, y_0) = 3 \quad e \quad f(x_1, y_1) = 0,48 .$$

Repetimos o processo até atingir a precisão. Caso  $f(x_{n+1}) \geq f(x_n)$ , veja o Teorema 4.5.2 e diminua o valor de  $h$ .

## 4.6 Método do M-Grande

Vamos considerar um problema de programação não-linear com restrições de igualdade e desigualdade:

$$\left| \begin{array}{l} \min \quad f(x) \\ \text{sujeito a} \quad g_i(x) = 0 \end{array} \right. , \quad (4.19)$$

e  $f : \mathbb{R}^m \rightarrow \mathbb{R}$ , função e  $g = (g_1, g_2, \dots, g_p)$  função vetorial de classe  $C^1$ . A ideia é aproximar um problema de otimização com restrição por um problema de otimização sem restrições. Um modo de realizar essa aproximação é através de funções de penalidade. Em [10] a função de penalidade utilizada é

$$F(x) = f(x) + M \frac{1}{\gamma} \sum_{i=1}^p (g_i(x))^\gamma, \quad (4.20)$$

onde  $\gamma > 0$  é uma constante e  $M > 0$  é uma variável de penalidade auxiliar.

A garantia do problema (4.19) ser solução também do problema (4.20) com a condição de  $M > 0$  ser suficientemente grande, se dá pelo teorema a seguir.

**Teorema 4.6.1** *Seja  $\{x_n\}$  uma sequência gerada pelo método de penalidade. Então, qual-*



quer ponto do limite dessa sequência é uma solução para o problema original.

Usando os métodos descritos anteriormente para resolver o problema

$$\left| \begin{array}{l} \min \quad f(x) \\ \text{sujeito a} \quad x \in \mathbb{R}^m \end{array} \right.,$$

com  $M$ , grande e fixo, e  $\gamma = 2$ , temos que aproximar a solução do PVI

$$\left\{ \begin{array}{l} x' = -\nabla F(x) \\ x(t_0) = x_0 \end{array} \right., \quad (4.21)$$

onde

$$F(x) = f(x) + M \frac{1}{2} \sum_{i=1}^p (g_i(x))^2,$$

então,

$$\nabla F(x) = \nabla f(x) + M \sum_{i=1}^p g_i(x) \nabla g_i(x).$$

Assim, aproximar a solução de um problema de otimização com restrições é equivalente a resolver (4.21).

Note que, em (4.21) é simples ver que  $\min F(x) = \min f(x)$ , pois

$$\frac{M}{2} \left( \sum_{i=1}^p (g_i(x))^2 \right) \geq 0$$

e o mínimo ocorre quando a igualdade for atingida, (veja [22] para mais discussões sobre o assunto, com restrições de desigualdades e relações com as condições de KKT).

Veja o Exemplo a seguir.

**Exemplo 4.6.2** *Resolva o problema*

$$\left| \begin{array}{l} \min \quad f(x, y) = 2x^2 + 2y^2 + 2xy \\ \text{sujeito a} \quad g(x, y) = x^2 + y^2 - 1 = 0 \end{array} \right.$$

*pelo Método do M-Grande.*

*Tomando  $\lambda = 2$  e  $M = 100$  temos, da Equação (4.20)*

$$F(x, y) = 2x^2 + 2y^2 + 2xy + \frac{100}{2}(x^2 + y^2 - 1)^2.$$

*Sendo assim,*

$$\begin{aligned}
\nabla F(x, y) &= (6x + 3y + 1, 4y + 3x + 1) + 100((x^2 + y^2 - 1)(2x, 2y)) \\
&= (6x + 3y + 1, 4y + 3x + 1) + 100(2x^3 + 2xy^2 - 2x, 2y^3 + 2yx^2 - 2y) \\
&= \begin{bmatrix} 6x + 3y + 1 + 200x^3 + 200xy^2 - 200x \\ 4y + 3x + 1 + 200y^3 + 200yx^2 - 200y \end{bmatrix}
\end{aligned}$$

Podemos terminar o problema como no Exemplo 4.5.3. Porém, note que estamos resolvendo

$$\begin{cases} \min \psi(u) &= \langle Hu, u \rangle \\ \text{sujeito a } g(u) &= 1 \end{cases},$$

dado pela Demonstração 2 do Teorema 2.2.2, com

$$H = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

e  $g(u) = \langle u, u \rangle$ ,  $u \in \mathbb{R}^\alpha$

A solução desse problema, produz um autovalor e um autovetor de  $H$ , e pode ser aproximada por meio do Método numérico baseado no gradiente e no Método de Euler como

$$v_{n+1} = (I - 2hH - 2hM(g(v_n) - 1)I)v_n$$

onde  $M > 0$  é suficientemente grande,  $h$  é positivo (suficientemente próximo de zero) e  $v_0$  é um vetor qualquer de  $E$  (com norma próxima de 1). É claro que o método apresentado é o Método de Euler para aproximar numericamente a solução da equação diferencial ordinária

$$\begin{cases} v' &= -\psi'(v) - M(g(v) - 1)g'(v) \\ v(0) &= v_0 \end{cases},$$

onde  $v = v(t) \in E$ , e o valor de  $v(t)$ , para  $t \in [0, +\infty)$ , será uma aproximação de um autovetor de  $H$ , para  $t$  grande.

Pelo algoritmo do Apêndice (A.8) obtemos os valores  $n_{1000000} = -0.70356236397446625$  e  $y_{1000000} = 0.70356236397256544$ . O valor aproximado do ponto de mínimo de  $f$  foi de 0.99.

Para finalizar, calculamos os autovalores de  $H$  pela decomposição LU, pelo algoritmo do Apêndice (A.2), obtendo 3 e 1 como autovalores da Hessiana. Assim, temos que o ponto de mínimo é exatamente 1.

Para finalizar, vamos resolver o Exemplo 3.2.7 de uma outra maneira e comparar os resultados.

**Exemplo 4.6.3** Considere o problema

$$\left\{ \begin{array}{l} \max \quad f(x, y) = e^{(xy)} \cos(x) \sin(y) \\ \text{sujeito a} \quad \begin{cases} 0 \leq x \leq \frac{\pi}{2} \\ 0 \leq y \leq \pi \end{cases} \end{array} \right. .$$

Os resultados deste exemplo foram obtidos algoritmo do Apêndice (A.9).

Observe a Figura (4.4) a seguir.

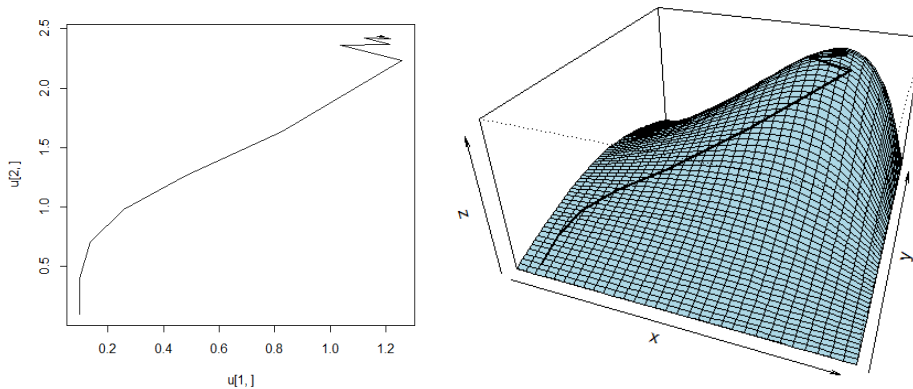


Figura 4.4: Gráfico do Exemplo (4.6.3) gerado pelo software.

Utilizando o Método dos Gradientes com Euler. Note que no primeiro gráfico da Figura (4.4) o topo, começou a oscilar, isso ocorreu porque o gradiente de  $f$  estava muito pequeno, assim, utilizamos a Condição de Armijo multiplicando  $h$  por 0.6.

Assim, tomando  $n = 100$ ,  $h_i = 0.5$ ,  $x_0 = 0.1$ ,  $y_0 = 0.1$  e  $u = (x, y)$  obtemos os seguinte valores:

$$u_{100} = (x_{100}, y_{100}) = (1.1817391494385037, 2.4393027453147207) \text{ e} \\ f(u_{100}) = 4.3764078775315127.$$

Como já vimos que os autovalores de  $Hf(u_{100})$  são negativos temos que  $f(u_{100})$  é aproximadamente valor de máximo de  $f(x)$ .

Agora vamos comparar os resultados obtidos do Exemplo (3.2.7) com o Exemplo (4.6.3).

O que difere as duas soluções obtidas nos valores de  $x$  e  $y$  são as últimas sete casas decimais e, no ponto de máximo,  $f(x, y)$ , são as três últimas casas decimais.

No entanto, note que quando solucionamos o problema pelo Método de Newton, temos que o Gradiente é aproximadamente  $10^{-15}$  enquanto que pelo Método dos Gradientes com Euler é aproximadamente  $10^{-9}$ . Isso nos leva a acreditar que o Método de Newton melhor soluciona o problema, porém, temos que o valor aproximado inicial pelo Método de Newton está bem próximo do ponto de máximo, enquanto que o valor aproximado inicial pelo Método dos Gradientes com Euler está longe do ponto de máximo.

Assim, temos que o Método de Newton é excelente quando o valor aproximado já está próximo do ponto que queremos encontrar, convergindo para este rapidamente. Desse modo, podemos encontrar uma solução ótima para o problema estudado fazendo a primeira iteração com o Método dos Gradientes com Euler e do resultado obtido, realizar mais iteração através do Método de Newton.

As Figuras (4.5) e (4.6) a seguir mostra a solução do problema obtida pelo software R.

```
> # Valor aproximado da raiz do gradiente de f(x,y)=0.
> print(u,digits=20)
[1] 1.1817391495970153 2.4393027452826881
>
> #Problema: Determinação de pontos de máximo e de mínimo de f(x,y).
>
> A(u[1],u[2])
[1] 2.294937e-15 -2.021403e-15
>
> print(f(u[1],u[2]),digits=20)
[1] 4.3764078775315118
>
> Hes=B(u[1],u[2]) # matriz do problema
> Hes
      [,1]      [,2]
[1,] -30.416901  4.376408
[2,]  4.376408 -10.488094
>
> aut(Hes,50) # Calcula autovalores e autovetores de Hes.
[1] -31.335616 -9.569378
```

Figura 4.5: Solução do Exemplo (3.2.7) gerado pelo software.

```
> print(u[,n],digits=20) # Valor aproximado do ponto de maximo.
[1] 1.1817391494385037 2.4393027453147207
>
> print(f(u[,n]),digits=20) # Valor aproximado do valor de maximo.
[1] 4.3764078775315127
>
> print(G(u[,n]),digits=20) #grad.
[1] 4.9616268423350088e-09 -1.0296750067234152e-09
```

Figura 4.6: Solução do Exemplo (4.6.3) gerado pelo software.

# Capítulo 5

## Considerações finais

Neste trabalho, discutimos um pouco sobre o problema de determinar pontos de máximo ou mínimo de funções diferenciáveis, no caso em que esse ponto é interior a um subconjunto aberto do espaço euclidiano  $\mathbb{R}^m$  ou então é um ponto de um subconjunto desse espaço, cuja fronteira é dada por superfícies de nível de outras funções também diferenciáveis. O interesse principal foi no caso em que as funções possam ser não lineares, tocando de certa forma nos chamados Problemas de Programação Não Linear.

Esse tipo de estudo é usualmente iniciado em disciplinas de Cálculo Diferencial e Integral, onde se discute a localização de pontos de máximos e mínimos de funções por meio das raízes da derivada da função ou por meio do Método dos Multiplicadores de Lagrange. Ambos os modos levam a necessidade de determinação de funções, que podem ser não lineares e de várias variáveis e assim, a dificuldade de resolver Problemas de Programação Não Linear geralmente está em alguns cálculos para encontrar ou aproximar os zeros de alguma função. Com isso, neste trabalho discutimos alguns métodos Numéricos para encontrar zeros de funções, métodos esses ligados ao Método de Newton e ao Método dos Gradientes, com auxílio de referenciais teóricos em Cálculo Diferencial e Integral, Álgebra Linear, Pesquisa Operacional e Cálculo Numérico, para encontrar os pontos de máximo e de mínimo locais de funções. Nesse processo, notamos que nem sempre existe uma forma direta para a resolução do problema e há a necessidade de uma análise detalhada de cada método e as vezes simulações, com auxílio de programação e algum software computacional, para a verificação da convergência e ou aplicabilidade do método ao problema.

Vale ressaltar que os métodos Numéricos geralmente produzem uma aproximação e não o valor exato desse ponto, no entanto, essa aproximação pode ser relativamente importante quando se trata de problemas reais, porque quando fazemos uso da Pesquisa Operacional devemos fazer um modelo para obtermos uma situação melhor que a conhecida.

Por fim, a realização deste trabalho proporcionou uma complementação na formação acadêmica. Além de desenvolver a redação científica, generalizou e consolidou teorias vistas durante a graduação.

# Referências Bibliográficas

- [1] THOMAS, G. B. *Cálculo* Vol 1. São Paulo: Pearson Addison-Wesley, 2002.
- [2] THOMAS, G. B. *Cálculo* Vol 2. São Paulo: Pearson Addison-Wesley, 2002.
- [3] BUENO, H.P. *Álgebra Linear: Um segundo curso*. Rio de Janeiro: Sociedade Brasileira de Matemática, 2006.
- [4] SAAD, Y.; VAN DER VORST, H.A. Iterative solution of linear systems in the 20th century. *Journal of Computational and Applied Mathematics*, Volume 123, Issues 12, 1 November 2000, Pages 1-33.
- [5] BRYAN, K.; LEISE, T. The R\$25,000,000,000 Eigenvector: The Linear Algebra behind Google. *SIAM Review*, 2006, Vol. 48, No. 3 : p. 569-581. <https://doi.org/10.1137/050623280>.
- [6] CANEN, A. G.; BOAVENTURA-NETTO, P. O. A SOBRAPO e os SBPO: Início, Trajetória, Atualidade, Reflexões. *XLII Sbpo*, Bento Gonçalves - RS, p.659-666, 2010. Disponível em: <http://www.din.uem.br/sbpo/sbpo2010/pdf/71320.pdf>. Acesso em: 20 abr. 2017.
- [7] MARINS, F. A. S. Introdução à Pesquisa Operacional. *São Paulo: Cultura Acadêmica: Universidade Estadual Paulista*, 2011.
- [8] KJELDSEN, T. H. A contextualized historical analysis of the Kuhn-Tucker Theorem in nonlinear programming: The impact of World War II. *Historia mathematica*, v. 27, n. 4, p. 331-361, 2000.
- [9] LACHTERMACHER, G. *Pesquisa operacional na tomada de decisões: modelagem em Excel*. Elsevier, 2004.
- [10] EVIRGEN, F. Analyze the optimal solutions of optimization problems by means of fractional gradient based system using VIM. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, v. 6, n. 2, p. 75-83, 2016.

- [11] FRANCO, N. B. *Cálculo numérico*. Pearson, 2006.
- [12] LIMA, E. L. *Curso de Análise*. Vol. 2. 8. ed. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada, 2005.
- [13] LIMA, E. L. *Curso de Análise*. Vol. 1. 8. ed. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada, 2005.
- [14] LIMA, E. L. *Análise Real*. Vol. 2. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada, 2004.
- [15] GARBI, G. G. *O romance das equações algébricas*. 4. ed., Livraria da Física: São Paulo, 2010.
- [16] OLIVEIRA, C. R. *Introdução à Análise Funcional*. 2. ed. Rio de Janeiro: IMPA, 2005.
- [17] DOMINGUES, H. H. *Espaços métricos e introdução à topologia*. Atual, 1982.
- [18] TAHA, H. A.; MARQUES, A. S.; SCARPEL, R. A. *Pesquisa operacional*. Pearson Education do Brasil, 2008.
- [19] MARTINEZ, J. M.; SANTOS, S. A. Métodos computacionais de otimização. *Colóquio Brasileiro de Matemática, Apostilas*, v. 20, 1995.
- [20] ZILL, D. G. *Equações diferenciais*. 3. ed., São Paulo: Makron Books, 2001.
- [21] ZILL, D. G. *Equações diferenciais com aplicações em modelagem*. 3. ed., São Paulo: Cengage Learning, 2016.
- [22] ZHOU, L. et al. Convergence analysis of a differential equation approach for solving nonlinear programming problems. *Applied mathematics and computation*, v. 184, n. 2, p. 789-797, 2007.

## Apêndice A

# APÊNDICE - Algoritmos em linguagem R

### A.1 Método de Newton do Exemplo (3.1.7)

```
# f(x)=x ^ 3 + 3x ^ 2 + 2*x - 1
f<-function(x){x^ 3 + 3*x ^ 2 + 2*x - 1}
curve(f,-4,2) # gráfico de f(n), para -4<= x<= 2
df<-function(x){3*x ^ 2+6*x+2}
n=1000
x=1
g<-function(x){
  x-f(x)/df(x)
}
for (i in 1:n){
  x=g(x)
}
# Valor aproximado da raiz de f(x)=0.
x
# Teste da estimativa.
f(x)
```

### A.2 Cálculo de autovalores por Decomposição LU do Exemplo (3.2.4)

```
# Autovalores por LU
```



```

aut<-function(A,m){
n=length(A[1,])
for (i in 1:m){
U<-matrix(0, nrow=nrow(A),ncol=ncol(A),byrow=TRUE)
L<-diag(nrow=nrow(A))
for ( i in 1:nrow(A)){ # LU rápido
for ( j in 1:ncol(A)){
if (i<=j){
U[i,j]<-A[i,j]-(L[i,c(1:(i-1))])%*%U[c(1:(i-1)),j] }
else{
L[i,j]<-(A[i,j]-(L[i,c(1:(j-1))])%*%U[c(1:(j-1)),j])/U[j,j] }
} }
A=U%*% L
}
b=A[1,]
for (i in 1:n){b[i]=A[i,i]} # aproximação dos autovalores
b }
#-----
# Teste
A=matrix(c(54,2,5,4,1,2,34,6,0,9,5,6,11,1,1,4,0,1,81,22,1,9,1,22,78),5,5)
A
aut(A,50) # aproximação dos autovalores.

```

### A.3 Algoritmo de Eliminação de Gauss

```

GaussElim<-function(a,b)
{
n=length(b)
A=matrix(a,n,n)
L=0*A
for (i in 1:n){L[i,i]=1}
c=b
# iterações escalonamento
for ( k in 1:(n-1) ) {
for ( i in (k+1):n )
{
m=A[i,k]/A[k,k]
A[i,k]=0

```

```

        L[i,k]=m
        for ( j in (k+1): n )
        {
            A[i,j]=A[i,j]-m*A[k,j]
        }
        b[i]=b[i]-m*b[k]
    }
}
B=matrix(0,n,2*n+1)
B[1:n,1:n]=L
B[1:n,(n+1):(2*n)]=A
B[,2*n+1]=b
p=B
p
}

#-----
# Resolver o sistema triangular superior
Triansup<- function(a,b)
{
n=sqrt(length(a))
# iterações
x=0*(1:n)
x[n]=b[n]/a[n,n]
for ( k in (n-1):1 )
{
s=0
for ( j in (k+1): n )
{
s=s+a[k,j]*x[j]
}
x[k]=(b[k]-s)/a[k,k]
}
x
}

```

## A.4 Algoritmo de Eliminação de Gauss com Pivota- mento Parcial

```

# algoritmo para Eliminacao de Gauss com pivotamento parcial
# entrada dos valores da matriz a T
A=matrix(c(54,2,5,4,1,2,34,6,0,9,5,1,11,1,1,3,5,7,81,22,4,4,8,2,78),n,n)
A
b=c(1,2,3,4,5)
b
#-----
# Início do escalonamento
n=length(b)
n
for ( k in 1:(n-1))
  {
  # teste pivo
  c=0*b
  for (i in k:n)
    {
    c[i]=abs(A[i,k])
    }
  c1=max(c)
  if (c1 > 0)
    {
    i=k
    while ( abs(A[i,k])< c1 )
      {
      # print("trocando linha")
      i=i+1
      }
    for ( j in 1:n )
      {
      c[j]=A[k,j]
      A[k,j]=A[i,j]
      A[i,j]=c[j]
      }
    d=b[k]
    b[k]=b[i]
  }
}

```

```

    b[i]=d
  } else { stop("o sistema nao possui solucao unica")}
#-----
# Escalonar a proxima linha
for ( i in (k+1):n )
{
  m=A[i,k]/A[k,k]
  A[i,k]=0
  for ( j in (k+1): n )
  {
    A[i,j]=A[i,j]-m*A[k,j]
  }
  b[i]=b[i]-m*b[k]
}
}
#-----
A
b
#-----
# resolucao do sistema triangular
a=0*(1:n)
a[n]=b[n]/A[n,n]
for ( k in (n-1):1 )
{
  s=0
  for ( j in (k+1): n )
  {
    s=s+A[k,j]*a[j]
  }
  a[k]=(b[k]-s)/A[k,k]
} a

```

## A.5 Método de Newton do Exemplo (3.2.6)

```

# Dados
f3 <- function(a,b,c,d) {a+c+1}
f0 <- function(a,b,c,d) {b*d-1}
f2 <- function(a,b,c,d) {b+d-a*c+1}

```

```

f1 <- function(a,b,c,d) {a*b+b*c-1}
f <- function(a,b,c,d) {c(f3(a,b,c,d),f0(a,b,c,d),f2(a,b,c,d),f1(a,b,c,d))}
# forma de vetor para f(u)
B <- function(a,b,c,d) {matrix(c(1,0,1,0,0,d,0,b,-c,1,-a,1,d,c,b,a),4,4,
byrow = TRUE)} # matriz do jacobiano de f(u)
n=10 # chute inicial para iterações
x=-2 # Chute inicial par u_0
y=-1
z=0
w=-1
u=c(x,y,z,w)
#-----
# Método de Newton
for (i in 1:n) {
  G=GaussElim(B(u[1],u[2],u[3],u[4]),-f(u[1],u[2],u[3],u[4])) # Escalona
  Bv=-F(u)
  G
  q=length(u)
  A1=G[1:q,(q+1):(2*q)]
  A1
  b1=G[,2*q+1]
  b1
  # Resolvendo o sistema escalonado (triangular superior)
  v=Triansup(A1,b1) # Resolve o sistema tringular Bv=w.
  u=v+u
  u}
#-----
# Valor aproximado da raiz de f(a,b,c,d)=0.
print(u,digits=22)
f(u[1],u[2],u[3],u[4]) # Teste do resultado
#-----
# As raízes de p(x)=(x ^ 2-ax-b)(x ^ 2-cx-d)
x1=(u[1]+sqrt(u[1] ^ 2+4*u[2]+0*1i))/2
  x1
x2=(u[1]-sqrt(u[1] ^ 2+4*u[2]+0*1i))/2
  x2
x3=(u[3]+sqrt(u[3] ^ 2+4*u[4]+0*1i))/2
  x3
x4=(u[3]-sqrt(u[3] ^ 2+4*u[4]+0*1i))/2

```

```

      x4
#-----
# teste
p<-function(x) { p=x ^ 4+x ^ 3+x ^ 2+x+1
                p }
p(x1)
p(x3)

```

## A.6 Método de Newton do Exemplo (3.2.7)

```

# f(x,y)=exp(x*y)*cos(x)*sin(y)
x <- seq(0, pi/2, by=0.1)
y <- seq(0,pi,by=0.1)
f<-function(x,y){exp(x*y)*cos(x)*sin(y)}
#-----
z <- outer(x, y, f) # gráfico de f(x,y)
z[is.na(z)] <- 4
op <- par(bg = "white")
persp(x, y, z, theta = 60, phi = 30, expand = 0.5, col = "lightblue")
#-----
dfx <- function(x,y) {sin(y)*(exp(x*y)*y*cos(x)-exp(x*y)*sin(x))}
dfy <- function(x,y) {cos(x)*(exp(x*y)*x*sin(y)+cos(y)*exp(x*y))}
ddfx <- function(x,y)
  {sin(y)*(exp(x*y)*y*y*cos(x)-2*exp(x*y)*y*sin(x)-exp(x*y)*cos(x))}
ddfy <- function(x,y)
  {cos(x)*(exp(x*y)*x*x*sin(y)+2*exp(x*y)*x*cos(y)-exp(x*y)*sin(y))}
ddfxy <- function(x,y) {cos(y)*(exp(x*y)*y*cos(x)-exp(x*y)*sin(x))+
  +sin(y)*(cos(x)*(x*exp(x*y)*y+exp(x*y))-sin(x)*exp(x*y)*x)}
A <- function(x,y){c(dfx(x,y),dfy(x,y))}
B <- function(x,y){
matrix(c(ddfx(x,y),ddfxy(x,y),ddfxy(x,y),ddfy(x,y)),
2,2)}
#-----
n=100
x=1
y=2.5
u=c(x,y)

```

```

for (i in 1:n){
  v=solve(B(u[1],u[2]),-A(u[1],u[2])) # Resolve o sistema Bv=-f(u).
  u=v+u
  u}

#-----
# Valor aproximado da raiz de f(x,y)=0.
print(u,digits=22)
#-----
#Problema: Determinação de pontos de máximo e de mínimo de f(x,y).
A(u[1],u[2])
f(u[1],u[2])
Hes=B(u[1],u[2])
aut(Hes,50) # Calcula autovalores e autovetores de Hes.

```

## A.7 Runge-Kutta do Exemplo (4.4.2)

```

# EDO: método de Runge-Kutta de ordem 2
RungK2<-function(f,u0,a,b,n){
  h=(b-a)/n
  t=seq(a,b,by=h)
  m=length(u0)
  Y=matrix(0,m,n+1)
  Y[,1]=u0
# O método Runge Kutta 2
  for (i in 1:n)
  {
    Y[,i+1]=Y[,i]+( f(t[i],Y[,i])+
    +f( t[i],Y[,i]+h*f(t[i],Y[,i]) ) ) *h/2
  }
  Y
}

#-----
# Equação  $y' = y - t$ 
a=0
b=1
u0=2
#-----
# Dados

```

```
f<-function(s,u) # Definição de f(t,Y)=Y'.
{
  p=u-s
  p
}
n=10
Y=RungK2(f,u0,a,b,n)
h=(b-a)/n
t=seq(a,b,by=h)
y<-function(s)exp(s)+s+1 # Solução exata.
curve(y,a,b)
points(t,Y,col = "red") # Gráfico de y(t)
```

## A.8 Método de Euler do Exemplo (4.6.2)

```
# Metodo de Euler para resolução numérica de EDO's PVI's, para problema de
otimização.
# Problema: Minimizar  $f(x,y)=2x^2+2y^2+2xy$ , com restrição  $g(x,y)=x^2+y^2=1$ .
# Equação  $u'(t) = -\text{Grad}(f(x,y)) - \text{Mg}(x,y) \text{ Grad}(f(x,y))$  e  $u(0)=(0.5,0)$ 
#-----
# Dados
n=1000000          # Número de passos
M=100              # M grande
ui=c(-1,0)         # u(0) inicial
ti=0               # tempo inicial
tf=20              # tempo final
h=(tf-ti)/n        # Tamanho do passo
h
A=matrix(c(2,1,1,2),2,2)
I= matrix(c(1,0,0,1),2,2)
g<-function(u)u[1]^ 2+u[2]^ 2
t=seq(ti,tf+h,by=h)
#-----
# O método dos gradientes com M grande.
u=matrix(0,2,n+1)
u[,1]=ui
for (i in 1:n)
```



```

    {
      u[,i+1]=(I-2*h*A-2*h*M*I*( g(u[,i])-1 ) )%*%u[,i]
    }
#-----
print(u[,n],digits=20) # Valor aproximado do ponto de mínimo do autovetor.
f<-function(u){p=2*u[1]^ 2+2*u[2]^ 2+2*u[1]*u[2]
  p}
print(f(u[,n]),digits=20) # Valor aproximado do ponto de mínimo do autovalor.
x=u[1,]
y=u[2,]
plot(x,y,'l')      # y(x) aproximado.
teste=matrix(c(2,1,1,2),2,2)
#-----
Usando o Algoritmo A.2
# Teste-----
H=matrix(c(2,1,1,2),2,2) # matriz do problema
aut(H,50) # aproximação dos autovalores.

```

## A.9 Método de Euler do Exemplo (4.6.3)

```

# Metodo de Euler para resolução numérica de EDO's PVI's, para problema de
# otimização.
# Com condição de Armijo.
# Problema: Maximizar  $f(x,y)=\exp(xy)\cos(x)\sin(y)$ , com  $0 < x < \pi/2$  e  $0 < y < \pi$ .
f<-function(u){
  x=u[1]
  y=u[2]
  p=exp(x*y)*cos(x)*sin(y)
  p}
# Equação diferencial  $u'(t) = \text{Grad}(f(u))$ , para  $u(0)=(0.1,0.1)$ .
#-----
# Dados
ui=c(0.1,0.1) # u(0) inicial
hi= 0.5 # Tamanho do passo inicial
#  $G(x,y)=\text{Grad}(f(x,y))$ .
G<-function(u){
  x=u[1]
  y=u[2]

```

```

    dfx = sin(y)*(exp(x*y)*y*cos(x)-exp(x*y)*sin(x))
    dfy = cos(x)*(exp(x*y)*x*sin(y)+cos(y)*exp(x*y))
    p=c( dfx, dfy )
    p
}
n=100 # Número de passos
#-----
# 0 método dos gradientes com Euler e Armijo.
u=matrix(0,2,n+1)
u[,1]=ui
for (i in 1:n)
{
    h=hi
    if (sqrt( sum( G( u[,i] )^ 2 ) ) < 10{-5} ){h=hi*10{5} }
    arm=1
    while ( arm > 0 ){
        h=0.6*h
        arm=f(u[,i])- f(u[,i]+h*G(u[,i])) # Teste de Armijo.
    }
    u[,i+1]=u[,i]+h*G(u[,i])
}
#-----
plot(u[1,],u[2,],'l') # Gráfico 2D para u(t).
# Grafico 3D para f e f(u(t)).
x <- seq(0, pi/2, by=0.05)
y <- seq(0,pi,by=0.05)
fa<-function(x,y){exp(x*y)*cos(x)*sin(y) # fa(x,y)=f(x,y); auxiliar para gráfico
}
require(grDevices) # for trans3d
z <- outer(x, y, fa)
z[is.na(z)] <- 4
op <- par(bg = "white")
persp(x, y, z, theta = 20, phi = 30, expand = 0.5, col = "lightblue")->
res
lines(trans3d(u[1,],u[2,], fa(u[1,],u[2,]), res), col = "black", lwd = 2)
#-----Solução-----
print(u[,n],digits=20) # Valor aproximado do ponto de maximo.
print(f(u[,n]),digits=20) # Valor aproximado do valor de maximo.
print(G(u[,n]),digits=20) #grad.

```