

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE MATEMÁTICA

ROBERTA MARCELINO DE ALMEIDA ALVES

**MINIMIZAÇÕES DE FUNÇÕES CONVEXAS VIA MÉTODO DE
NEWTON**

TRABALHO DE CONCLUSÃO DE CURSO

CORNÉLIO PROCÓPIO

2016

ROBERTA MARCELINO DE ALMEIDA ALVES

**MINIMIZAÇÕES DE FUNÇÕES CONVEXAS VIA MÉTODO DE
NEWTON**

Trabalho de Conclusão de Curso apresentada ao Departamento de Matemática da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do grau de "Licenciado em Matemática".

Orientador: Prof.^a Dr.^a Elenice Weber Stiegemeier

CORNÉLIO PROCÓPIO

2016

TERMO DE APROVAÇÃO

ROBERTA MARCELINO DE ALMEIDA ALVES

MINIMIZAÇÕES DE FUNÇÕES CONVEXAS VIA MÉTODO DE NEWTON

Trabalho de Conclusão de Curso apresentada ao Departamento de Matemática da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do grau de "Licenciado em Matemática".

BANCA EXAMINADORA

Prof.^a Dr.^a Elenice Weber Stiegelmeier
Universidade Tecnológica Federal do Paraná

Prof.^a M.^a Daniele Costa Silva
Universidade Tecnológica Federal do Paraná

Prof.^a Dr.^a Luciana Casacio
Universidade Tecnológica Federal do Paraná

Cornélio procópio, 16 de Junho de 2016.

A Folha de Aprovação assinada encontra-se na Coordenação do Curso.

Dedico este trabalho aos meus pais, irmãos, namorado, meus amigos e a toda minha família que, com muito carinho e apoio, não mediram esforços para que eu chegasse até esta etapa de minha vida. Agradeço também a todos os professores que me acompanharam durante a graduação e que me proporcionaram um grande amadurecimento intelectual e caráter ao longo desta jornada.

AGRADECIMENTOS

Quando tudo parecia desabar, quando decidi rir ou chorar, ir ou ficar, desistir ou lutar algumas pessoas estiveram ao meu lado e percorreram este caminho incentivando a busca dos meus sonhos.

Agradeço a Deus, em primeiro lugar, por ouvir as minhas orações, me confortou e deu forças para chegar onde eu estou.

Agradeço a minha família que não só neste momento, mas em toda a minha vida estiveram ao meu lado me ajudando e indicando o caminho correto.

Agradeço a mulher da minha vida e mãe Marilena que me ensinou a ser uma mulher íntegra, com caráter, coragem e dignidade para enfrentar a vida.

Agradeço ao meu pai Eziquiel que me incentivou a estudar, me ensinou a batalhar e buscar meus objetivos.

Agradeço a minha irmã Mariana por sempre estar ao meu lado, me apoiando.

Agradeço ao meu namorado Renan que esteve presente nesse momento aguentando minhas crises de estresse, que teve paciência comigo, pelo amor, incentivo e por me ajudar dando dicas para o desenvolvimento desse trabalho.

Agradeço a minha orientadora Elenice pela oportunidade e por ter me ajudado e guiado no decorrer desse trabalho, me dando o suporte necessário.

Agradeço também aos meus amigos da Universidade, Marila, Pedro, Nayara, Ingrid, Lilian, Lissa, Maria Carolina e Michelly que hoje são considerados como melhores amigas porque foram a minha família ao longo desses anos.

Por último, mas não menos importante, agradeço a todos os meus professores por me incentivar na busca do conhecimento e amar a matemática.

O amor é paciente, o amor é bondoso. Não inveja, não se vangloria, não se orgulha. Não maltrata, não procura seus interesses, não se ira facilmente, não guarda rancor. O amor não se alegra com a injustiça, mas se alegra com a verdade. Tudo sofre, tudo crê, tudo espera, tudo suporta.

1 Coríntios 13:4-7

RESUMO

ALVES, Roberta M. A. MINIMIZAÇÕES DE FUNÇÕES CONVEXAS VIA MÉTODO DE NEWTON. 51 f. TRABALHO DE CONCLUSÃO DE CURSO – Departamento Acadêmico de Matemática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2016.

Neste trabalho será abordado o estudo de problemas de otimização não linear com foco em minimização de funções convexas. Serão apresentadas definições de funções convexas, os problemas de minimização irrestritos juntamente com a existência de soluções e as condições de otimalidade em situações convexas. Além disso, são apresentados resultados de simulações numéricas utilizando o método de Newton para minimização de funções convexas. Para validar os resultados obtidos utiliza-se a rotina de minimização `fminunc` presente no Optimization Toolbox do MATLAB.

Palavras-chave: Programação Não Linear, Otimização Irrestrita, Funções Convexas, Método de Newton.

ABSTRACT

ALVES, Roberta M. A. CONVEX FUNCTIONS MINIMIZATION BY WAY OF NEWTON'S METHOD. 51 f. TRABALHO DE CONCLUSÃO DE CURSO – Departamento Acadêmico de Matemática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2016.

This work aimed to study nonlinear optimization problems with focus on convex functions minimization. The Convex function definitions, unconstrained minimization problems with the existence of solutions and optimality conditions in convex situations are presented. In addition, results of numerical simulations using Newton's method for minimizing convex functions are presented. In order to validate the results, the routine `fminunc` of MATLAB Optimization Toolbox is used.

Keywords: Nonlinear Programming, Unconstrained Optimization, Convex Functions, Newton Method

LISTA DE FIGURAS

FIGURA 1	– Fases da Modelagem.	13
FIGURA 2	– Teorema de Weierstrass.	19
FIGURA 3	– Representação gráfica do exemplo 1.	20
FIGURA 4	– Aproximação polinômio de Taylor.	21
FIGURA 5	– Conjunto convexo e não convexo	23
FIGURA 6	– Função convexa e não convexa.	25
FIGURA 7	– Representação geométrica do Teorema 10.	27
FIGURA 8	– Ilustração do Corolário 2.	27
FIGURA 9	– Superfície de $f(x,y) = \frac{x^2}{y}$	32
FIGURA 10	– Uma iteração do Método de Newton.	35
FIGURA 11	– Gráfico da função $f(x) = 2(x-3)^2 + (x+2)^2$	36
FIGURA 12	– Análise gráfica da função Booth's.	44
FIGURA 13	– Análise gráfica da função Parabolic.	45
FIGURA 14	– Análise gráfica da função Rosenbrock.	45
FIGURA 15	– Análise gráfica da função Matyas.	46
FIGURA 16	– Análise gráfica da Freudenstein Function	46
FIGURA 17	– Análise gráfica da função Beales.	47

LISTA DE TABELAS

TABELA 1	– Resultados obtidos com a implementação do Algoritmo 1.	41
TABELA 2	– Funções utilizadas para avaliação do Algoritmo 3.	42
TABELA 3	– Resultados obtidos com a implementação do Algoritmo 3.	42
TABELA 4	– Resultados obtidos com a rotina <i>fminunc</i> presente no MATLAB.	42
TABELA 5	– Comparação dos resultados	43

SUMÁRIO

1	INTRODUÇÃO	11
2	FERRAMENTAS MATEMÁTICAS	16
2.1	TOPOLOGIA NO \mathbb{R}^N	16
2.2	RESULTADOS DA ÁLGEBRA LINEAR	18
2.3	RESULTADOS DO CÁLCULO DIFERENCIAL	18
3	CONVEXIDADE	23
3.1	CONJUNTO CONVEXO	23
3.2	FUNÇÕES CONVEXAS	25
3.3	FUNÇÕES CONVEXAS DIFERENCIÁVEIS	26
4	O PROBLEMA DE MINIMIZAÇÃO CONVEXA	29
4.1	FORMULAÇÃO MATEMÁTICA	29
4.2	EXISTÊNCIA DE SOLUÇÕES	30
4.3	CONDIÇÕES DE OTIMALIDADE	31
5	MÉTODO DE NEWTON	33
5.1	MÉTODO DE NEWTON NA RESOLUÇÃO DE SISTEMAS NÃO LINEARES	33
5.2	MÉTODO DE NEWTON NA OTIMIZAÇÃO DE FUNÇÕES DE UMA ÚNICA VARIÁVEL	34
5.3	MÉTODO DE NEWTON NA OTIMIZAÇÃO DE FUNÇÕES DE MÚLTIPLAS VARIÁVEIS	37
5.4	MÉTODO DA SEÇÃO ÁUREA	38
5.5	CONVERGÊNCIA DO MÉTODO DE NEWTON	40
6	SIMULAÇÕES NUMÉRICAS	41
6.1	ANÁLISE DOS RESULTADOS	41
7	CONCLUSÃO	48
	REFERÊNCIAS	50

1 INTRODUÇÃO

A história da matemática mostra que a busca por modelos matemáticos para os problemas de otimização não é tão recente. Existe uma lenda da Princesa Dido da Fenícia (século IX a.C), contado pelo poeta Virgílio (70 à 19 anos a.C) em Eneida que mostra a procura de formas ótimas.

De acordo com Vieira (apud Valentino, 2009, p. 5):

"Dido teve seu marido, sacerdote Arquebas, assassinado pelo seu irmão Pigmalião, para subtrair-lhe seus tesouros. Temendo sua própria morte, Dido fugiu em um navio com um grande número de seguidores dispostos a fundar uma nova cidade chamada Cartago. Tentou comprar terras do rei Jarbas da Numídia, para que pudessem se estabelecer. O que conseguiu foi que, só teria em terras o que pudesse abranger com a pele de um boi. Dido e seu povo então decidiram cortar a pele do boi em tiras bem finas, emendar todas e englobar num semicírculo em terreno às marges do Mediterrâneo."

Além desse exemplo, Valentino (2009) apresenta outra busca pela otimalidade realizada pelo Cartógrafo-mor Pedro Nunes do rei D. João III em 1537, que lecionou aulas de matemática na Universidade de Coimbra. Pedro foi o primeiro a perceber que se um navio se mantivesse em rota fixa não conseguiria navegar através de uma circunferência, que é o caminho ótimo numa esfera, este deveria então seguir uma rota em espiral chamada loxodrômica, conhecida como Tratado da Esfera de 1537. Por volta de 1629, Fermat fez uma grande descoberta: um método para achar máximos e mínimos de funções polinomiais, que foi publicada após a sua morte.

Os problemas de máximos e mínimos de Fermat ajudaram Newton e Leibiz desenvolver o que hoje é conhecido como Cálculo. Segundo Valentino (2009) "[...] Fermat, quando descobriu o método para encontrar máximos, mínimos e tangentes em curvas algébricas, ainda não tinha conhecimento sobre limite, mas com base em suas ideias junto das definições e notações de Newton e Leibniz [...]", enunciou a Condição Necessária de Otimalidade.

Porém, foi durante a Segunda Guerra Mundial, 1939 a 1945, que um grupo de militares ingleses se deparou com o problemas de recursos militares limitados . A partir desse problema,

formou-se um grupo multidisciplinar composto por matemáticos, físicos, engenheiros e cientistas sociais na Inglaterra para explorar os problemas de estratégia e de tática para defesa do país (GASS, 2002). Assim, surgiu pela primeira vez o termo Pesquisa Operacional.

Segundo o Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa em Engenharia (COOPE) o termo pesquisa operacional consistem em:

"[...] estudar problemas reais envolvendo situações de tomada de decisão, através de modelos matemáticos habitualmente processados computacionalmente. Ela aplica conceitos e métodos de outras disciplinas científicas na concepção, no planejamento ou na operação de sistemas para atingir seus objetivos. Procura, assim, introduzir elementos de objetividade e racionalidade nos processos de tomada de decisão, sem descuidar, no entanto, dos elementos subjetivos e de enquadramento organizacional que caracterizam os problemas[...] (COOPE, 2015)".

Com o fim da Segunda Guerra Mundial, a utilização de técnicas de PO atraiu o interesse de diversas outras áreas. Nesse contexto, começou a serem elaborados modelos matemáticos que permitiam soluções de caráter racional e objetivo junto com a simulação da realidade, com vistas à discussão e avaliação de diferentes estratégias ou alternativas de ação. Logo, esses modelos foram absorvidos pelas empresas que precisavam de ferramentas que ajudassem na tomada de decisão em mercados crescentes e competitivos. Estes modelos matemáticos consistem na representação de um problema real, que surgem com a necessidade de compreender os fenômenos físicos e econômicos da sociedade.

Segundo Lachtermacher (2007) com o advento da pesquisa operacional surgem modelos matemáticos associados a diferentes contextos, tais como:

1. Modelos Icômicos: os quais se referem a mudanças de escalas;
2. Modelos Analógicos: usam um conjunto de propriedades para representar um outro conjunto;
3. Modelos Simbólicos: faz uso de letras, números ou símbolos para representar variáveis e suas relações.

A partir destes modelos, busca-se a solução do problema, o qual é obtido através da validação do modelo na representação do sistema real. A diferença entre a solução real e a solução proposta pelo modelo utilizado depende diretamente da precisão do modelo em descrever o comportamento original do sistema.

A Figura 1 apresenta um diagrama exemplificando o processo que pode ser adotado para a modelagem dos problemas reais.

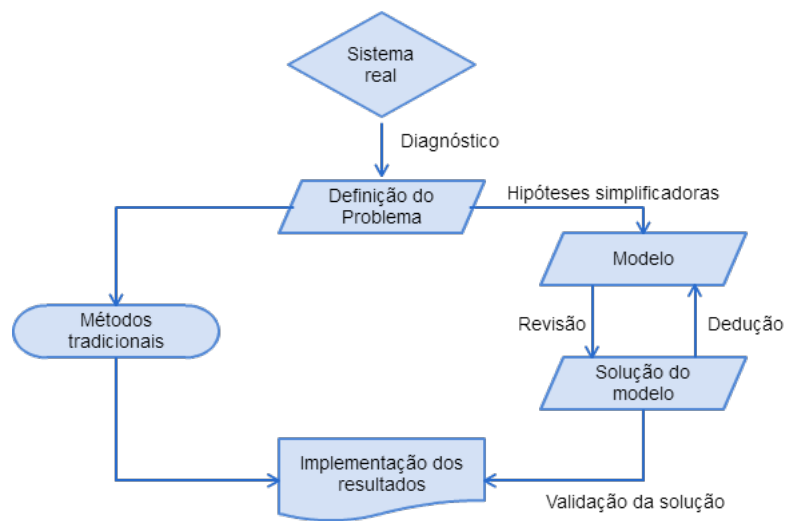


Figura 1: Fases da Modelagem.

Fonte: Arenales et al. (2007).

No caso dos problemas de otimização, os modelos matemáticos são compostos por três requisitos (LACHTERMACHER, 2007):

1. **Função Objetivo:** é uma função matemática que define a qualidade da solução em função das variáveis de decisão;
2. **Variáveis de decisão e parâmetros:** variáveis de decisão são as incógnitas a serem determinadas pela solução do modelo. Parâmetros são valores fixos no problema;
3. **Restrições:** de modo a levar em conta as limitações físicas do sistema, o modelo deve incluir restrições que limitam as variáveis de decisão a seus valores possíveis (ou viáveis).

Dentro dessa perspectiva, os modelos matemáticos para otimização em geral são formulados com o critério de minimizar (*min*) ou maximizar (*max*) uma função objetivo, satisfazendo um conjunto de restrições. As restrições podem ser escrita na forma de equações ou inequações e as vezes podem não existir, neste caso têm-se os problemas chamados irrestritos. Esses modelos podem ser classificados como problemas de programação linear, programação não linear, programação inteira e problemas de controle ótimo, dependendo da forma da função objetivo e suas restrições (LACHTERMACHER, 2007).

Nos modelos de programação linear, como o próprio nome já diz, a função objetivo e as restrições são lineares. Os modelos de programação não lineares (PNL) ou do inglês “*nonlinear programming*” (NLP) apresentam funções não lineares na função objetivo ou em uma das restrições envolvidas.

Os modelos não lineares estão presentes em diversas áreas do conhecimento, podendo citar o campo da agricultura, finanças, transporte, processos químicos, produtivos, recursos naturais, ambientais e energéticos, entre outros (ARENALES et al., 2007; RIBEIRO; KARAS, 2013).

Diversos autores têm estudado e proposto modelos de otimização com o objetivo de reduzir os custos de produção. Em Stiegelmeier (2012) foi proposto um modelo de otimização dinâmico não linear com o objetivo de minimizar os custos de produção em uma lavoura a partir da aplicação de controle químico. Valentino (2009) apresenta um modelo matemático de controle ótimo linear a não linear, com interesse de maximizar a produção, minimizar o custo e a quantidade de fungicida usado no feijoeiro.

A solução dos problemas de otimização não linear é obtida a partir das condições de otimalidade. As condições de otimalidade para problemas de otimização não linear têm o objetivo de caracterizar minimizadores locais. Porém, reconhecer se um minimizador local também é global não é fácil, a menos que a função objetivo tenha características especiais. O caso mais simples é o de funções convexas (FRIEDLANDER, 1994).

Em problemas de otimização não linear convexa a função objetivo e as restrições expressas por desigualdade devem ser convexas e as restrições expressas por igualdade devem ser afins. Devido a complexidade em se obter a solução analítica do problema faz-se necessário a utilização de métodos numéricos e, com isso, obter a solução aproximada do problema. No caso de problemas convexas existem algoritmos eficientes que podem resolver os problemas de otimização convexa com uma boa aproximação (IZMAILOV; SOLODOV, 2009; RIBEIRO; KARAS, 2013).

Os métodos numéricos mais utilizados na resolução de problemas de otimização são método do Gradiente, método de Newton, método Quase-Newton, método de direções conjugadas e estratégias de globalização. No presente trabalho será utilizado o método de Newton, o qual consiste em um processo iterativo que minimiza, a cada iteração a aproximação quadrática da função objetivo. Se o ponto inicial x_0 estiver próximo de um minimizador, a sequência gerada pelo método de Newton será convergente para a solução ótima.

O método de Newton é um dos métodos numéricos mais eficiente para a resolução de equações não lineares e minimização de funções. O método de Newton surgiu com os estudos de Issac Newton em *De analysi per aequationes numero terminorum infinitas* (Análise de equações com um número infinito de termos), escrito em 1669, e publicado em 1711 por William Jones em *De methodis fluxionum et serierum infinitarum* (O método de série infinita e Fluxionum) (MACHADO; ALVES, 2013).

Para o desenvolvimento do presente trabalho serão estudados problemas onde todas as funções usadas são não lineares e o caso abordado é o problema irrestrito, isto é, problemas de otimização sem restrição. E ainda, serão descritas algumas situações que garantem a existência de um minimizador e, em seguida, discutidas as condições de otimalidade para o problema de otimização irrestrita convexa. Os resultados numéricos serão obtidos via implementação do método de Newton utilizando o *Software* matemático MATLAB.

O presente texto está dividido 7 capítulos. O Capítulo 2, seguinte a essa introdução, apresenta uma revisão sobre os temas relevantes relacionados a este trabalho. Nele são descritos alguns resultados da Análise no \mathbb{R}^n , Álgebra Linear e do Cálculo Diferencial. O Capítulo 3 introduz a teoria de convexidade, onde são descritos os conjuntos e funções convexas e sua diferenciabilidade utilizada na otimização irrestrita. No Capítulo 4 é descrita a formulação do problema de minimização convexa, otimização irrestrita e condições de otimalidade para problemas de minimização irrestrito convexo. O Capítulo 5 descreve o Método de Newton para solução de funções não lineares. Serão abordados os casos de minimização irrestrita de funções de uma única variável e minimização irrestrita de funções de múltiplas variáveis, utilizado para a obtenção da solução numérica dos problemas de minimização convexa. Além disso, será apresentado o método da Seção Áurea o qual é utilizado na busca do passo ótimo no caso do método de Newton. E, no Capítulo 6, são descritos os resultados da minimização de funções de uma única variável e para funções de múltiplas variáveis através da busca exata. A validação dos resultados numéricos será obtida através da comparação dos resultados obtidos pelo método implementado e a rotina de minimização *fminunc* presente no *Optimization Toolbox* do MATLAB.

2 FERRAMENTAS MATEMÁTICAS

Neste capítulo são apresentados conceitos que fundamentam o desenvolvimento deste trabalho. Os conceitos da Topologia no \mathbb{R}^n , da Álgebra Linear e do Cálculo Diferencial são considerados primordiais para este estudo, para mais detalhes veja Lima (1999, 2006), Boldrini et al. (1980), Guidorizzi (2001).

2.1 TOPOLOGIA NO \mathbb{R}^N

A seguir, são descritos alguns conceitos topológicos da teoria de conjuntos de fundamental importância para o desenvolvimento dos capítulos seguintes.

Definição 1. Um ponto a é interior ao conjunto $X \subset \mathbb{R}^n$ se existe $\varepsilon > 0$ tal que $(a - \varepsilon, a + \varepsilon) \in X$.

Definição 2. O conjunto dos pontos interiores de X é chamado de interior do conjunto X e denotado por $\text{int } X$.

Definição 3. Um conjunto $A \subset \mathbb{R}^n$ é dito aberto se $A = \text{int}A$, ou seja, todos os pontos de A são interiores a A .

Definição 4. Um ponto a é ponto aderente do conjunto $X \subset \mathbb{R}^n$ quando a é limite de alguma sequência de pontos $x_n \in X$.

Definição 5. O fecho de um conjunto X é o conjunto \bar{X} formado por todos os pontos aderentes do conjunto X .

Definição 6. Um conjunto X é dito fechado quando $X = \bar{X}$, ou seja, todo ponto aderente de X pertence a X .

Definição 7. Um conjunto $X \subset \mathbb{R}^n$ chama-se compacto quando é limitado e fechado.

Definição 8. Uma cobertura do conjunto $X \subset \mathbb{R}^n$ é uma família $C = (C_\lambda)_{\lambda \in L}$ de conjuntos $C_\lambda \subset \mathbb{R}^n$, tais que $X \subset \bigcup_{\lambda \in L} C_\lambda$, isto é, para todo $x \in X$ existe algum $\lambda \in L$ tal que $x \in C_\lambda$.

Definição 9. Uma função $f : X \rightarrow \mathbb{R}$ definida no conjunto $X \subset \mathbb{R}$ é contínua no ponto $a \in X$ quando, para todo $\varepsilon > 0$ arbitrário, pode-se obter $\delta > 0$ tal que $x \in X$ e $|x - a| < \delta$ impliquem $|f(x) - f(a)| < \varepsilon$.

Exemplo 1. Todo polinômio $p : \mathbb{R} \rightarrow \mathbb{R}$ é uma função contínua. Toda função racional $\frac{p(x)}{q(x)}$ (quociente de dois polinômios) com $q(x) \neq 0$ é contínua no seu domínio.

Definição 10. Uma função $f : X \rightarrow \mathbb{R}$ e $a \in X \cap X'$, é derivável no ponto a quando

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

Corolário 1. Uma função é contínua nos pontos em que é derivável.

Definição 11. Seja $f : U \rightarrow \mathbb{R}$ um função definida no aberto $U \subset \mathbb{R}^n$. Para cada $i = 1, \dots, n$, a i -ésima derivada parcial de f no ponto $a = (a_1, \dots, a_n) \in U$ é o número

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a + te_i) - f(a)}{t} = \lim_{t \rightarrow 0} \frac{f(a_1, \dots, a_i + ta_i, \dots, a_n) - f(a)}{t},$$

caso esse limite exista.

Definição 12. Seja $f : U \rightarrow \mathbb{R}$ uma função que possui as n derivadas parciais em todos os pontos de $U \subset \mathbb{R}$. Ficam então definidas n funções

$$\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} : U \cup \mathbb{R}, \text{ onde } \frac{\partial f}{\partial x_i} : x \mapsto \frac{\partial f}{\partial x_i}(x).$$

Se estas funções forem contínuas em U , se f é uma função de classe C^1 e escreve-se $f \in C^1$.

Definição 13. A função $f : U \rightarrow \mathbb{R}$ é de classe C^2 quando existem as derivadas parciais até a segunda ordem em todos os pontos de U . f é uma função de classe C^n , e escreve-se $f \in C^n$, quando f é n vezes derivável e além disso, a função $f^{(n)} : U \rightarrow \mathbb{R}$ é contínua. Quando $f \in C^n$ para todo $n \in \mathbb{N}$, diz-se que f é de classe C^∞ e escreve-se $f \in C^\infty$. É conveniente considerar f como a sua própria "derivada de ordem zero" e escrever $f^{(0)} = f$. Assim, $f \in C^0$ significa que f é uma função contínua.

Definição 14. Uma função $f : U \rightarrow \mathbb{R}$ é dita suave quando $f \in C^\infty$.

Exemplo 2. A função dada por

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{se } x \neq 0 \\ 0 & \text{se } x = 0 \end{cases}$$

é um exemplo de função suave.

Mais detalhes sobre as definições acima podem ser encontradas em Lima (1999).

2.2 RESULTADOS DA ÁLGEBRA LINEAR

Os conceitos de Álgebra Linear estão presente no desdobramento da teoria de otimização vinculados aos conceitos de matriz simétrica e matriz definida positiva, que são usualmente utilizadas para definir as condições de otimalidade, conforme apresentado a seguir.

Definição 15. Considere a matriz real $A = (a_{ij})$ de ordem $m \times n$. Denomina-se transposta de A e indica-se por A^t a seguinte matriz $n \times m$: $A^t = (b_{ji})$ em que $b_{ji} = a_{ij}$, $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n$.

Definição 16. Uma matriz quadrada é dita simétrica se $A = A^t$, ou seja, quando $a_{ij} = a_{ji}$, para todo i e $j = 1, \dots, m$.

Exemplo 3. Exemplos de matriz simétrica.

$$\begin{bmatrix} 4 & 3 & -1 \\ 3 & 2 & 0 \\ -1 & 0 & 5 \end{bmatrix} \text{ e } \begin{bmatrix} a & b & c & d \\ b & e & f & g \\ c & f & h & i \\ d & g & i & k \end{bmatrix}$$

Definição 17. O núcleo de uma matriz quadrada $A_{n \times n}$, denotado por $N(A)$, é um subconjunto de \mathbb{R}^n formado por todas as soluções do sistema homogêneo $Ax = 0$, ou seja, $N(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$.

Definição 18. Seja $A_{n \times n}$ uma matriz simétrica. Diz que A é definida positiva quando $x^T Ax > 0$, para todo $x \in \mathbb{R}^n - \{0\}$. Tal propriedade é denotada por $A > 0$. Se $x^T Ax \geq 0$, para todo $x \in \mathbb{R}^n$, A é dita semi definida positiva, fato este denotado por $A \geq 0$.

Definição 19. Dada uma matriz $A_{n \times n}$ simétrica, a forma quadrática é definida como $Q_A(x) := x^T Ax$, para $x \in \mathbb{R}^n$. Se $\forall x \neq 0$, $Q^A(x) > 0$, então a forma quadrática é chamada positiva definida.

Definição 20. Uma matriz é chamada positiva definida se a forma quadrática associada a A , $Q_A(x) > 0$, $\forall x \neq 0$.

Definição 21. Uma matriz tem posto completo se $\text{posto}(A) = \min(m, n)$, isto é, se o posto é o maior valor possível.

2.3 RESULTADOS DO CÁLCULO DIFERENCIAL

O estudo de problemas de otimização requerem conhecimentos de Cálculo Diferencial, para determinação dos valores extremos de uma função, isto é, o valor máximo ou mínimo que

uma função pode assumir em um intervalo dado. Assim, destacam-se conceitos relacionados ao gradiente e matriz Hessiana, fundamentais para o desenvolvimento dos métodos numéricos (GUIDORIZZI, 2001).

Teorema 1 (Teorema de Weierstrass). *Seja $f : X \rightarrow \mathbb{R}$ uma função contínua no conjunto compacto $X \subset \mathbb{R}$. Existem x_1 e $x_2 \in X$ tais que $f(x_1) \leq f(x) \leq f(x_2)$ para todo $x \in X$.*

Demonstração: Como $f(x)$ é compacto, logo é fechado e limitado, ou seja, se $\alpha = \inf f(x)$ e $\beta = \sup f(x)$, então $-\infty < \alpha \leq f(x) \leq \beta < +\infty$, para todo $x \in X$.

Como f é fechado, existe $x_1, x_2 \in X$ tal que $f(x_1) = \alpha$ e $f(x_2) = \beta$, então

$$f(x_1) \leq f(x) \leq f(x_2).$$

A Figura 2 ilustra a representação geométrica do Teorema 1.

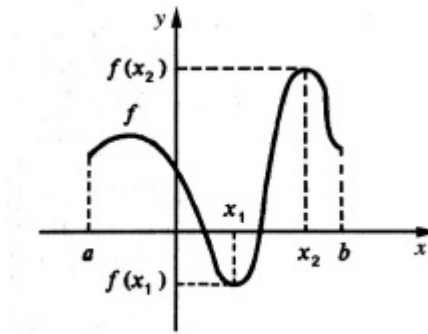


Figura 2: Teorema de Weierstrass.

Fonte: Guidorizzi (2001, p.122).

Definição 22. *O gradiente de uma função diferenciável $f : U \rightarrow \mathbb{R}$ no ponto $a \in U$ é o vetor*

$$\text{grad } f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right) \quad (1)$$

O vetor gradiente pode ser denotado por ∇f ou $\text{grad } f$.

Exemplo 4. *Seja $f(x, y) = x^2 + y^2$. Calcule $\nabla f(1, 1)$ e represente-o geometricamente.*

Solução:

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right) = (2x, 2y).$$

Assim,

$$\nabla f(1, 1) = (2, 2) = 2\vec{i} + 2\vec{j}.$$

A representação geométrica é expressada na Figura 3.

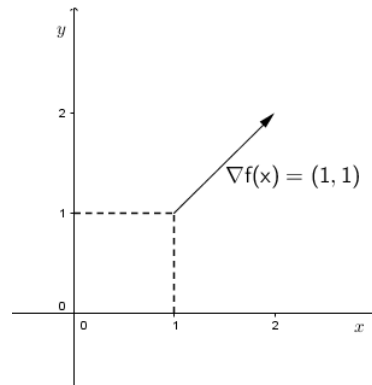


Figura 3: Representação gráfica do exemplo 1.

O gradiente de uma função possui algumas propriedades que justificam a sua importância, o qual dá interessantes informações sobre o comportamento da função. Para isto, fixa-se $a \in U$ e suponha que $\text{grad } f(a) \neq 0$. Então:

- 1 O gradiente aponta para uma direção segundo a qual a função é crescente;
- 2 Dentre todas as direções ao longo das quais a função cresce, a direção do gradiente é a de crescimento mais rápido;
- 3 O gradiente de f no ponto a é ortogonal ao conjunto de nível de f que passa por a .

Detalhes das propriedades podem ser encontradas em Lima (2006).

Definição 23. Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de n variáveis reais. Se todas as derivadas parciais de f existem, a matriz hessiana de f é dada por:

$$H[f(x_1, x_2, \dots, x_n)] = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{vmatrix}.$$

Definição 24. Seja $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ uma função vetorial de várias variáveis, com

$F(X) = (f_1(X), \dots, f_n(X))$ onde $X = [x_1, x_2, \dots, x_m]$. A representação matricial da derivada, quando existe, é denominada de Matriz Jacobiana e definida como:

$$J(x_1, \dots, x_m) = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_m} \end{bmatrix}.$$

A Matriz Jacobiana foi denominada pelo matemático alemão Carl Gustav Jakob Jacobi, é formada pelas derivadas parciais de primeira ordem de uma função vetorial.

Exemplo 5. *Obter a matriz jacobiana de*

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

Solução:

$$J(x, y) = \begin{bmatrix} x^2y \\ \frac{x}{y} \\ x - y \end{bmatrix} = \begin{bmatrix} 2xy & x^2 \\ \frac{1}{y} & \frac{-x}{y^2} \\ 1 & -1 \end{bmatrix}.$$

Teorema 2. (Taylor de primeira ordem) *Seja f uma função diferenciável em x_0 . A aproximação T do polinômio de Taylor é:*

$$T(x) = f(x_0) + f'(x_0)(x - x_0).$$

O gráfico de T , representado na Figura 4, é a reta tangente à f em $(x_0, f(x_0))$.

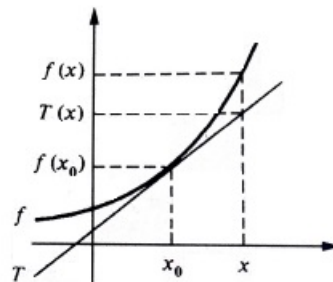


Figura 4: Aproximação polinômio de Taylor.

Fonte: Guidorizzi (2001, p. 465).

Assim, T é o polinômio de grau no máximo 1 que melhor aproxima localmente a f em volta de x_0 .

O polinômio

$$P(x) = f(x_0) + f'(x_0)(x - x_0)$$

denomina-se polinômio de Taylor de ordem 1 de f em volta de x_0 .

Teorema 3. (Taylor de segunda ordem)

Se $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função duas vezes diferenciável em x_0 , então

$$P_2(x) = f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2$$

é denominado polinômio de Taylor de ordem 2 de f em torno de x_0 .

Proposição 1. *Analogamente, o polinômio*

$$P_2(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2$$

é uma aproximação quadrática para f em torno do ponto x_0 e é chamado polinômio de Taylor de ordem 2 da função.

Dentre todos os polinômios de grau menor ou igual a 2, ele é o único que satisfaz

$$p(x_0) = f(x_0), p'(x_0) = f'(x_0)$$

e

$$p''(x_0) = f''(x_0).$$

3 CONVEXIDADE

Neste capítulo são descritos resultados relacionados à Análise Convexa, onde serão abordados temas relacionados a conjuntos convexos, funções convexas e funções convexas diferenciáveis (NOCEDAL; WRIGHT, 2006; LIMA, 1999).

3.1 CONJUNTO CONVEXO

Definição 25. Um conjunto $C \subset \mathbb{R}^n$ é dito convexo quando dados $x, y \in C$, o segmento $[x, y] = \{(1-t)x + ty \mid t \in [0, 1]\}$ estiver inteiramente contido em C .

Definição 26. Um conjunto $C \subset \mathbb{R}^n$ é dito não convexo, ou seja côncavo, quando dados $x, y \in C$, o segmento $[x, y] = \{(1-t)x + ty \mid t \in [0, 1]\}$ não estiver inteiramente contido em C .

A Figura 5 explicita a definição de conjunto convexo e conjunto não convexo.

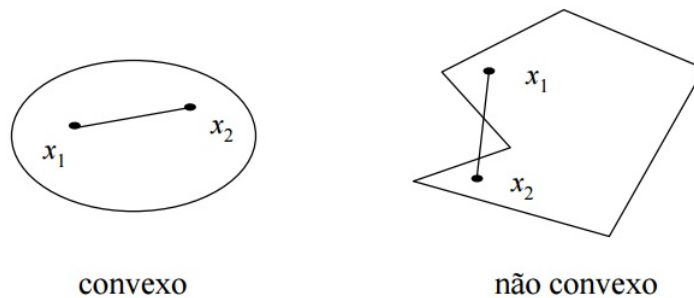


Figura 5: Conjunto convexo e não convexo

Exemplo 6. Pode-se tomar os seguintes conjuntos como convexas: qualquer subespaço de \mathbb{R}^n e o conjunto vazio.

Lema 1. Seja $C \subset \mathbb{R}^n$ um conjunto fechado não vazio. Dado $z \in \mathbb{R}^n$, existe um $\bar{z} \in C$ tal que

$$\|z - \bar{z}\| \leq \|z - x\|,$$

para todo $x \in C$.

Lema 2. Seja $C \subset \mathbb{R}^n$ um conjunto não vazio, convexo e fechado. Dado $z \in \mathbb{R}^n$ existe um único $\bar{z} \in C$ tal que

$$\| (z - \bar{z}) \|_2 \leq \| (z - x) \|_2,$$

para todo $x \in C$.

Teorema 4. Seja $C \subset \mathbb{R}^n$ um conjunto não vazio, convexo e fechado, $z \in \mathbb{R}^n$ e $\bar{z} = \text{proj}_C(z)$. Então

$$(z - \bar{z})^T (x - \bar{z}) \leq 0,$$

para todo $x \in C$.

Lema 3. Seja $C \subset \mathbb{R}^n$ um conjunto não vazio, convexo e fechado e $z \in \mathbb{R}^n$. Se satisfaz

$$(z - \bar{z})^T (x - \bar{z}) \leq 0$$

para todo $x \in C$, então $\bar{z} = \text{proj}_C(z)$.

Teorema 5. Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável e $C \subset \mathbb{R}^n$ convexo e fechado. Se $x^* \in C$ é minimizador local de f em C , então

$$\text{proj}_C(x^* - \alpha \nabla f(x^*)) = x^*,$$

para todo $\alpha \geq 0$.

Teorema 6. Seja Ω e Γ conjuntos convexos do \mathbb{R}^n . Então também são conjuntos convexos:

1. $\beta\Gamma := \{y : y = \beta x, x \in \Gamma, \beta \in \mathbb{R}\}$
2. $\Omega + \Gamma$
3. $\Omega \cap \Gamma$

Proposição 2. Seja $D_i \subset \mathbb{R}^n, i = 1, 2$, conjuntos convexos fechados. Se um deles também é limitado (logo, compacto), então o conjunto $D_1 + D_2$ é convexo e fechado.

Definição 27. Seja $D \subset \mathbb{R}^n$ um conjunto qualquer. O fecho convexo de D , denotado $\text{conv } D$, é o menor conjunto convexo em \mathbb{R}^n que contém D ou, equivalentemente, a interseção de todos os conjuntos convexos em \mathbb{R}^n que contém D .

Proposição 3. Seja $D \subset \mathbb{R}^n$ um conjunto compacto. Então D é compacto.

3.2 FUNÇÕES CONVEXAS

Funções convexas são importantes no contexto da otimização uma vez que caracterizam pontos globais de uma função (LIMA, 2006; NOCEDAL; WRIGHT, 2006; FRIEDLANDER, 1994).

Definição 28. Seja $C \subset \mathbb{R}^n$ um conjunto convexo. A função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é convexa em C quando para quaisquer $x \in C, y \in C$ e $t \in [0, 1]$, tem-se que

$$f((1-t)x+ty) \leq (1-t)f(x) + tf(y).$$

Definição 29. Se $C \subset \mathbb{R}^n$ um conjunto não convexo, diz-se que a função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é não convexa em C quando para quaisquer $x \in C, y \in C$ e $t \in [0, 1]$, tem-se que

$$f((1-t)x+ty) \geq (1-t)f(x) + tf(y).$$

Geometricamente, pode-se dizer que uma função convexa está sempre abaixo do segmento que liga as extremidades. Na Figura 6 são apresentados exemplos de função convexa e não convexa.

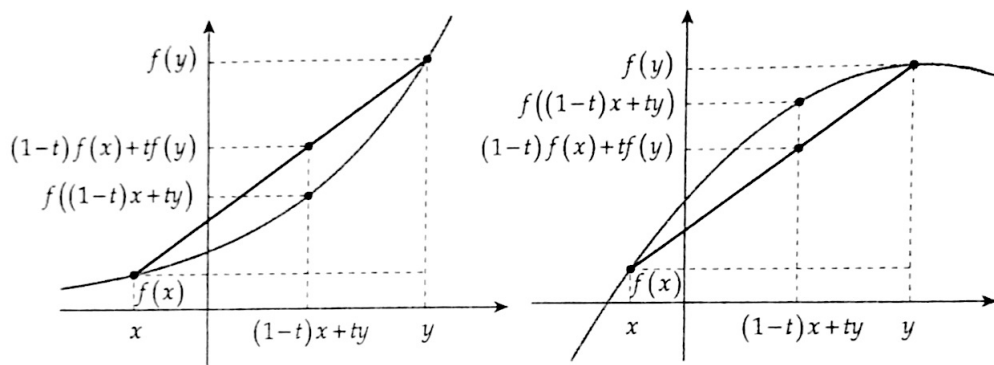


Figura 6: Função convexa e não convexa.

Fonte: Ribeiro e Karas (2003, p. 56).

Definição 30. A função f é estritamente convexa quando a desigualdade acima é estrita para todos $x \neq y$ e $t \in (0, 1)$. A função f diz-se fortemente convexa com módulo $\gamma > 0$, quando para quaisquer $x \in C, y \in C, t \in (0, 1)$, tem-se

$$f((1-t)x+ty) \leq (1-t)f(x) + tf(y) - \gamma t(1-t)\|x-y\|^2$$

Uma função fortemente convexa é estritamente convexa e uma função estritamente convexa é convexa.

Exemplo 7. Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ uma função estritamente convexa e fortemente convexa.

1. Note que $f(x) = x^2$ é um exemplo de função fortemente convexa, pois

$$\begin{aligned} (x + t(y-x))^2 &= x^2 + 2tx(y-x) + t^2(y-x)^2 \\ &\leq x^2 + 2tx(y-x) + t(y-x)^2 \\ &= x^2 + t(y^2 - x^2). \end{aligned}$$

2. Note que $f(x) = e^x$ é estritamente (mas não fortemente) convexa. Então, seja $z = (1-t)x + ty$ e $e^d \geq 1 + d$, $\forall d \in \mathbb{R}$, tomando $d = x - z$ e depois $d = y - z$, obtêm-se

$$e^x \geq e^z + e^z(x-z)$$

e

$$e^y \geq e^z + e^z(y-z).$$

Portanto, multiplicando a primeira desigualdade por $(1-t)$ e a segunda por t , concluí-se que $e^z \leq (1-t)e^x + te^y$, donde segue a convexidade de $f(x)$.

Teorema 7. Se $f : C \rightarrow \mathbb{R}$ e $g : D \rightarrow \mathbb{R}$ são funções convexas nos conjuntos convexos C e D , então

i) tf é convexa em C para todo $t \geq 0$;

ii) $f + g$ é convexa em $C \cap D$.

3.3 FUNÇÕES CONVEXAS DIFERENCIÁVEIS

Teorema 8. Seja $C \subset \mathbb{R}^n$ um conjunto convexo e $f : C \rightarrow \mathbb{R}$ uma função convexa. Então f é contínua no interior de C .

Exemplo 8. Observe uma função $f : C \rightarrow \mathbb{R}$ convexa pode ser descontínua na **fronteira** de C .
 $C = [-1, 1]$ e

$$Q(x) = \begin{cases} x^2 & \text{se } |x| < 1 \\ 2 & \text{se } |x| = 1 \end{cases}$$

Teorema 9. Seja $C \subset \mathbb{R}^n$ um conjunto convexo e $f : C \rightarrow \mathbb{R}$ uma função convexa. Então

1. O conjunto $D \subset C$ no qual f atinge seu mínimo é convexo;

2. Qualquer mínimo local é mínimo global f .

Teorema 10. Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável e $C \subset \mathbb{R}^n$ um conjunto convexo. Então f é convexa em C se, e somente se, para todos $x, y \in C$,

$$f(x) \geq f(y) + \nabla f(y)^T(x - y).$$

Na Figura 7 é ilustrado a diferenciabilidade de f conforme o Teorema 10.

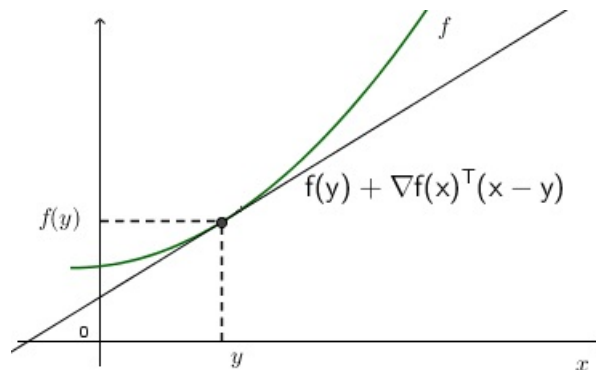


Figura 7: Representação geométrica do Teorema 10.

Teorema 11. Seja $C \subset \mathbb{R}^n$ um conjunto convexo com interior não-vazio e $f : C \rightarrow \mathbb{R}$. Então f é convexa em C se, e somente se, f for semi-definida positiva em C .

Corolário 2. Seja $C \subset \mathbb{R}^n$ um conjunto convexo e $f : C \rightarrow \mathbb{R}$ uma função convexa. Se existe um ponto $x^* \in C$ tal que para todo $x \in C$

$$\nabla f(x^*)^T(x - x^*) \geq 0,$$

então x^* é um mínimo global de f em C . Em particular, todo ponto estacionário é minimizador global como pode ser visto na Figura 8.

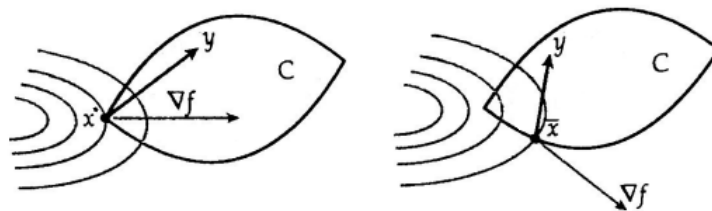


Figura 8: Ilustração do Corolário 2.

Fonte: Ribeiro (2013, p. 60).

Teorema 12. *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa diferenciável e $C \subseteq \mathbb{R}^n$ convexo e fechado.*

Se

$$\text{proj}_C(x^* - \nabla f(x^*)) = x^*,$$

então $x^ \in C$ é um minimizador global de f em C .*

Para demonstração dos Teoremas acima veja Guidorizzi (2001), Lima (2006) e Nocedal e Wright (2006).

4 O PROBLEMA DE MINIMIZAÇÃO CONVEXA

Neste capítulo é apresentada a formulação matemática de problemas de otimização convexos, destacando-se os aspectos teóricos relacionados aos problemas de otimização irrestrito, a existência de soluções e as condições necessárias e suficientes de otimalidade.

4.1 FORMULAÇÃO MATEMÁTICA

Problemas de otimização podem ser formulados por

$$\begin{aligned} & \min f(x) \\ & \text{sujeito a} \\ & g_i(x) \geq 0, \quad i = 1, 2, \dots, m \\ & h_j(x) = 0, \quad j = 1, 2, \dots, p, \end{aligned} \tag{2}$$

onde x um vetor n -dimensional de variáveis, $f(x)$ é a função objetivo que deve ser maximizada ou minimizada, $g_i(x)$ representam as restrições de desigualdade, $h_j(x)$ as restrições de igualdades e i, j é o conjunto de índices. Note que um problema de maximizar f pode ser substituído por minimizar $-f$.

Segundo Nocedal e Wright (2006), o problema de minimização convexo apresenta três requisitos:

- a função objetivo deve ser convexa;
- as restrições expressas por desigualdades devem ser convexas;
- as restrições expressas por igualdade devem ser afim.

Problemas de otimização, em geral, podem ser escritos usando a formulação (2), chamada de formulação padrão. Uma distinção importante entre os problemas do tipo (2) é conforme a forma do funcional objetivo $f(x)$ e suas restrições.

Neste trabalho será abordado o caso de programação não linear irrestrito onde o funcional objetivo é uma função convexa.

4.2 EXISTÊNCIA DE SOLUÇÕES

A seguir são estudadas algumas situações que garantem a existência de um minimizador e, em seguida, serão discutidas as condições de otimalidade para o problema de otimização irrestrita.

Para a obtenção da solução de um problema de programação não linear irrestrito, utiliza-se uma condição necessária de otimalidade, a qual permite a obtenção de uma candidata à solução analítica para o problema e, portanto, exata. Porém, em muitos casos, se torna difícil e até mesmo impraticável obter uma candidata a solução analiticamente devido à complexidade das equações envolvidas. Em casos como esses, é necessário utilizar métodos numéricos, com os quais é possível a obtenção de uma solução aproximada.

Dados um conjunto $D \subset \mathbb{R}^n$ e uma função $f : D \rightarrow \mathbb{R}$, tem-se por objetivo achar um minimizador de f no conjunto D , ou seja, o ponto no qual o valor de f é o menor possível. Este problema pode ser descrito como:

$$\min f(x), x \in D. \quad (3)$$

O conjunto D é designado conjunto viável do problema e os pontos de D correspondem aos pontos viáveis, ou seja, o conjunto viável de soluções.

Teorema 13 (Teorema da Existência). *Um problema de otimização tem solução se a função objetivo é contínua e o conjunto viável é não-vazio, fechado e limitado.*

Definição 31. *Um ponto $x^* \in D$ é um minimizador global de (3), se*

$$f(x^*) \leq f(x) \quad \forall x \in D.$$

Observe que x varia em todo \mathbb{R}^n ou pelo menos no domínio de interesse.

Muitas vezes o minimizador global pode ser difícil de ser encontrado, pois tem-se apenas o conhecimento local de f .

Teorema 14 (Teorema da minimização convexa). *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $f : D \rightarrow \mathbb{R}$ é uma função convexa em D . Então, todo minimizador local de (3) é global. Além disso, o conjunto minimizador é convexo. Se f é estritamente convexa, o minimizador é único.*

Pelas definições apresentadas, pode-se concluir que a única maneira de determinar se um ponto x^* é um minimizador local é examinando todos os pontos em uma vizinhança U a fim de verificar se algum deles possui valor de função menor.

No entanto, quando uma função é suave, existem maneiras mais práticas e eficientes para se identificar um minimizador local. Em particular, se f possui segunda derivada contínua, é possível dizer se um ponto x^* é minimizador local examinando apenas o gradiente $\nabla f(x^*)$ e a Matriz Hessiana $\nabla^2 f(x^*)$.

Teorema 15 (Teorema da Unicidade da solução). *Em um problema de otimização têm-se a solução única se o conjunto viável é estritamente convexo ou a função objetivo é estritamente quase-côncava ou ambos.*

4.3 CONDIÇÕES DE OTIMALIDADE

No caso de problemas de minimização irrestrita o objetivo é minimizar uma função de variáveis reais a um valor real sem que hajam restrições nos valores das variáveis, ou seja:

$$\min f(x), \quad x \in \mathbb{R}^n, \quad (4)$$

em que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função suave.

Para a obtenção da solução do problema descrito em (4), pode-se utilizar uma condição necessária de otimalidade, a qual fornece uma candidata à solução analítica para o problema e, portanto, exata.

Teorema 16. (Condições de primeira ordem) *Seja $C \subset \mathbb{R}^n$ um conjunto convexo e $f : C \rightarrow \mathbb{R}$. Então f é estritamente convexa em C se, e somente se, para todo x e $x^* \in C$ tal que $x \neq x^*$,*

$$f(x) > f(x^*) + \nabla f(x^*)^T (x - x^*)$$

Note que se $\nabla f(x^)^T (x - x^*) \geq 0$ para todo $x \in C$, então x^* é mínimo global estrito de f em C .*

Teorema 17. *Seja $C \subset \mathbb{R}^n$ um conjunto convexo com interior não-vazio e $f : C \rightarrow \mathbb{R}$ definida em C . Se $f(x) > 0$ para todo $x \in C$, então f é estritamente convexa em C .*

A seguir, são descritas as condições de segunda ordem.

Teorema 18. (Condições de segunda ordem) *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe C^2 e $C \subset \mathbb{R}^n$ convexo não vazio. Então, f é convexa se, somente se, a sua matriz Hessiana é semidefinida positiva em todo o seu domínio, ou seja, $\nabla^2 f(x) \geq 0$, para todo $x \in C$.*

Exemplo 9.

a) Considere a função quadrática $f: \mathbb{R}^n \rightarrow \mathbb{R}$, com $\text{dom } f = \mathbb{R}$ dado por

$$f(x) = \left(\frac{1}{2}\right)x^T P x + q^T x + r$$

sendo $P, q \in \mathbb{R}^n$ e $r \in \mathbb{R}$. Seja $\nabla^2 f(x) = P$ para todo x , f é convexa se e somente se $P \succeq 0$. Para funções quadráticas, f é estritamente convexa se e somente se $P \succ 0$.

b) Na Figura 9 têm-se um exemplo de função convexa.

$$f(x, y) = \frac{x^2}{y}, \text{ com}$$

$$\text{dom}(f) = \mathbb{R} \times \mathbb{R}_+ = \{(x, y) \in \mathbb{R}^2 | y > 0\}.$$

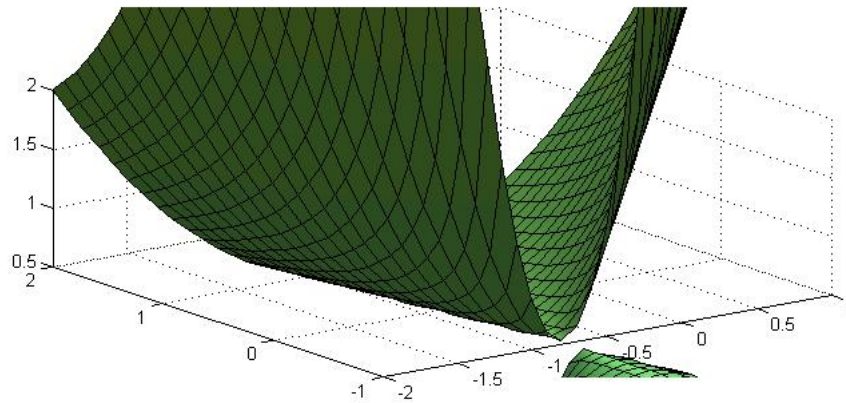


Figura 9: Superfície de $f(x, y) = \frac{x^2}{y}$.

Mais detalhes sobre condições de otimalidade para problemas convexos podem ser encontrados em Nocedal e Wright (2006),

5 MÉTODO DE NEWTON

Este Capítulo aborda o estudo relativo ao método de Newton e suas aplicações na solução numérica de problemas de sistema de equações não lineares e de problemas de otimização convexa. Além disso, é descrito o método da seção Áurea para a determinação do passo ótimo em problemas de minimização de funções.

Os resultados descritos neste Capítulo podem ser encontrados em Izmailov e Solodov (2012), Ribeiro e Karas (2013), Nocedal e Wright (2006).

5.1 MÉTODO DE NEWTON NA RESOLUÇÃO DE SISTEMAS NÃO LINEARES

Inicialmente o método de Newton foi proposto para encontrar as raízes de polinômios e posteriormente, foi utilizado na resolução de problemas de otimização.

O método de Newton é um processo iterativo importante na resolução de problemas de programação não linear, visando encontrar um minimizador para a função f . Utilizando as condições de otimalidade, serão resolvidos sistemas de n equações e n incógnitas dado por $\nabla f(x) = 0$.

A solução de sistema de equações não lineares do método de Newton é formulado como:

Encontre $x \in \mathbb{R}^n$, tal que $F(x) = 0$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$,

onde $F(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix}$.

A ideia é aproximar F por um polinômio de Taylor de primeira ordem. Dada uma estimativa x_k , considere o sistema linear:

$$F(x_k) + J(x_k)(x - x_k) = 0, \quad (5)$$

onde J representa a matriz jacobiana de F .

Caso $J(x_k)$ seja inversível, da equação (5) obtêm-se:

$$x_{k+1} = x_k - (J(x_k))^{-1}F(x_k). \quad (6)$$

A equação (6) corresponde ao método de Newton para a determinação de sistemas de equações não lineares.

Maiores detalhes do método de Newton para resolução de sistema de equações não lineares podem ser encontrados em Martinez e Santos (1995).

5.2 MÉTODO DE NEWTON NA OTIMIZAÇÃO DE FUNÇÕES DE UMA ÚNICA VARIÁVEL

O Método de Newton, para resolver problemas de minimização irrestrita, é o mesmo utilizado para resolver o sistema de equações (MARTINEZ; SANTOS, 1995).

A solução do problema de minimização irrestrita é formulado por:

$$\min f(x), x \in \mathbb{R}, \quad (7)$$

onde $f : \mathbb{R} \rightarrow \mathbb{R}$. Os pontos estacionários deste problema são caracterizados pela equação $f'(x) = 0$.

Seja $x_k \in \mathbb{R}^n$ uma aproximação de um ponto estacionário do problema (7). O método inicia utilizando uma aproximação do polinômio de Taylor de segunda ordem:

$$f(x_{k+1}) = f(x_k) + f'(x_k)[x_{k+1} - x_k] + \frac{1}{2}f''(x_k)[x_{k+1} - x_k]^2 \quad (8)$$

$$f'(x_k) + f''(x_k)[x_{k+1} - x_k] = 0 \Rightarrow \quad (9)$$

$$x_{k+1} = x_k - [f''(x_k)]^{-1}f'(x_k). \quad (10)$$

A equação (10) corresponde ao processo iterativo do método de Newton para minimização de uma única variável. A Figura 10 ilustra uma iteração do método de Newton para os problemas de otimização convexa.

O Algoritmo 1 apresenta o processo iterativo do método de Newton para minimização de funções de uma única variável considerando o passo unitário.

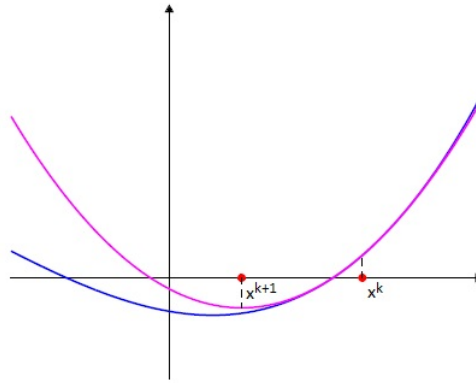


Figura 10: Uma iteração do Método de Newton.

Algoritmo 1: Método de Newton puro (versão para minimização).

Entrada: $f(x), f'(x), f''(x), \varepsilon > 0, x_0$

Saída: Solução ótima x^*

- 1 Inicialize $k = 0$;
 - 2 REPITA enquanto $|f'(x)| \geq \varepsilon$;
 - 3 FAÇA
 - 4 $x_{k+1} = x_k - (f''(x))^{-1}f'(x)$;
 - 5 $k = k + 1$;
 - 6 **retorna**;
 - 7 Solução ótima = x_k^*
-

O exemplo a seguir ilustra o processo iterativo do método de Newton para o problema de minimização.

Exemplo 10. Considere o problema de minimização:

$$\min f(x) = 2(x - 3)^2 + (x + 2)^2.$$

Solução:

Primeira iteração:

Tomando a precisão $\varepsilon = 0,001$ e o ponto inicial $x_0 = 2$, tem-se:

$$x_1 = x_0 - [f''(x_0)]^{-1}f'(x_0)$$

$$x_1 = 2 - [0,1667]4$$

$$x_1 = 2 - 0,6668$$

$$x_1 = 1,3332.$$

Como $f'(x_1) > \varepsilon$ o processo continua.

Segunda iteração:

$$x_2 = x_1 - [f''(x_1)]^{-1} f'(x_1)$$

$$x_1 = 1,3332 - [0,1667] \cdot -0,0008$$

$$x_1 = 1,3332 + 0,0001$$

$$x_1 = 1,3333.$$

Como $f'(x_1) = 0,0008 < \varepsilon$, então, com duas iterações, o ponto ótimo $x^* = 1,3333$ foi alcançado com a precisão desejada.

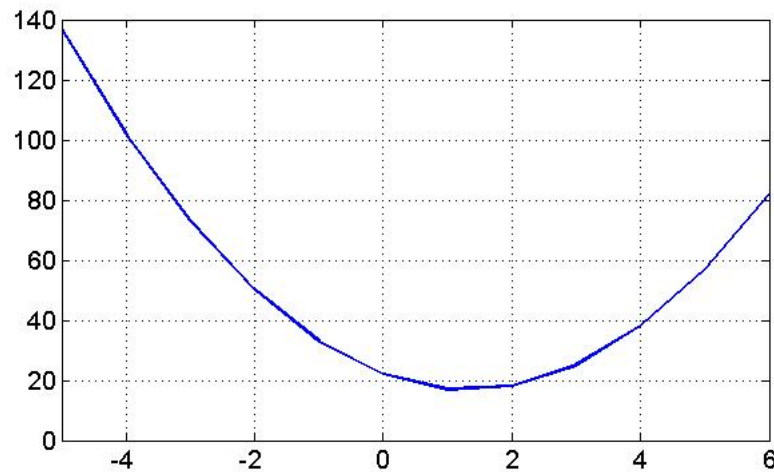


Figura 11: Gráfico da função $f(x) = 2(x-3)^2 + (x+2)^2$.

A Figura 11 ilustra o comportamento da função $f(x)$ analisada. Note que o ponto $x^* = 1,333$ é o ponto de mínimo da função.

Portando, o método de Newton gera uma sequência de soluções x_0, x_1, \dots, x_k que converge para um ponto estacionário x^* , ou seja, o ponto onde $f'(x^*) = 0$. No entanto, x^* não é necessariamente um ponto de mínimo, podendo também definir um ponto de inflexão ou um ponto de máximo.

Para definir em qual caso o ponto ótimo x^* se encontra, é necessário verificar as condições necessárias e suficientes para mínimo local.

1. Condições necessárias para mínimo local: $f'(x^*) = 0$ e $f''(x^*) \geq 0$
2. Condições suficientes para mínimo local: $f'(x^*) = 0$ e $f''(x^*) > 0$.

Note que, o método de Newton encontra o ótimo global em apenas uma iteração para as funções quádricas e converge rapidamente para o mínimo global da função.

5.3 MÉTODO DE NEWTON NA OTIMIZAÇÃO DE FUNÇÕES DE MÚLTIPLAS VARIÁVEIS

Suponha que pretende-se resolver o seguinte problema de minimização:

$$\begin{aligned} \min f(x) \\ x \in \mathbb{R}^n \end{aligned} \quad (11)$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função de classe C^2 e os pontos estacionários deste problema são caracterizados pela solução da equação $\nabla f(x) = 0$. Assim aplicando a relação (5) para $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ obtêm-se:

$$F(x) = \nabla f(x).$$

Tomando $x = x_k$ como uma estimativa, pode-se aproximar $f(x)$ por um polinômio de Taylor de segunda ordem, da seguinte forma:

$$f(x) \approx h(x) := f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2} \nabla^2 f(x_k) (x_{k+1} - x_k), \quad (12)$$

onde $\nabla f(x)$ é o gradiente de $f(x)$ e $\nabla^2 f(x)$ é a matriz hessiana de $f(x)$.

Observe que $h(x)$ é uma função quádrática, nesse caso se $\nabla h(x) = 0$, então:

$$\begin{aligned} \nabla h(x) &= \nabla f(x_k) + \nabla^2 f(x_k) (x - x_k) \Rightarrow \\ \nabla f(x_k) + \nabla^2 f(x_k) (x - x_k) &= 0 \Rightarrow \\ x &= x_k - \nabla^2 f(x_k)^{-1} \nabla f(x_k). \end{aligned} \quad (13)$$

O método de Newton requer o ponto inicial x_0 e um procedimento que calcula o gradiente e a inversa da hessiana da função a ser minimizada. Assim, o algoritmo pode ser descrito como um processo iterativo em que novos pontos x_{k+1} são gerados a partir do ponto atual x_k .

O algoritmo do método de Newton implementado foi formulado como:

$$x_{k+1} = x_k + \alpha_k d_k \quad (14)$$

onde $d_k = -\nabla^2 f(x_k)^{-1} \nabla f(x_k)$, é a direção de busca linear, α_k representa o tamanho do passo a ser dado na direção de busca e k é o número de iterações.

O algoritmo, então, é dividido em duas etapas, a determinação da direção de busca d_k e a avaliação do parâmetro α_k . Para isso, é necessário utilizar o método da Seção Áurea, usada como estratégia na escolha do passo ótimo (IZMAILOV; SOLODOV, 2012).

5.4 MÉTODO DA SEÇÃO ÁUREA

O método da Seção Áurea é frequentemente aplicado em problemas de otimização, pois, qualquer que seja a dimensão inicial do problema, quando o método numérico já calculou a direção de busca, d_k , então, é necessário obter o passo ótimo, α_k .

A escolha de α_k é feita por uma minimização local de $f(x_k + \alpha_k d_k)$, com o uso da função auxiliar $h_k(\alpha_k) = f(x_k + \alpha_k d_k)$. Observe que $h_k(\alpha_k)$ é uma função de uma única variável. Então, pode-se fazer uso do método da seção Áurea para encontrar o seu mínimo. Ao aplicar o método da seção Áurea na função $h_k(\alpha_k)$ deve-se tomar o cuidado de restringir α_k , de tal forma que o vetor $x_k + \alpha_k d_k$ permaneça em D . Assim, ao se minimizar a função $h_k(\alpha_k)$, encontra-se os valores ótimos de α_k para cada valor de x_k em cada iteração (IZMAILOV; SOLODOV, 2012).

O Método da Seção Áurea é dividido em duas fases, a primeira fase constitui-se em determinar um intervalo $[a, b]$ que contém um minimizador de f . Na segunda fase o intervalo $[a, b]$ é reduzido, descartando-se o subintervalo deste que não contém o minimizador, até que reste um intervalo de comprimento ε , sendo ε a tolerância pré-fixada.

Suponha que o minimizador de f pertence ao intervalo $[a, b]$, então:

- i) Considere $a < x < y < b$ em $[0, \infty)$.
- ii) Se $f(x) < f(y)$ então o trecho $(y, b]$ não pode conter o minimizador e pode ser descartado.
- iii) Se $f(x) \geq f(y)$ então o trecho $[a, y)$ pode ser descartado.
- iv) Particione o intervalo obtido e repita o processo.

Outra estratégia é escolher x e y que dividem o segmento $[a, b]$ na razão áurea, de acordo com a definição abaixo.

Definição 32. Um ponto c divide o segmento $[a, b]$ na razão áurea quando a razão entre o maior segmento e o segmento todo é igual à razão entre o menor e o maior dos segmentos. Tal razão é conhecida como o número de ouro e vale $\frac{\sqrt{5}-1}{2} \approx 0.618$.

O Algoritmo 2 apresenta as fases para a implementação do método da seção Áurea. As entradas necessárias são a função f , o intervalo da seção Áurea escolhido, o valor inicial x_0 , a direção de busca d_k e o erro ε . A saída é o passo ótimo correspondente ao método de Newton.

Algoritmo 2: Método da Seção Áurea.

Entrada: $f(x)$, a , b , x_0 , d_k , ε

Saída: Intervalo I

- 1 Inicialize $k = 1$;
 - 2 REPITA enquanto $|(b - a)| > \varepsilon$;
 - 3 $x_a = b - 0.618(b - a)$;
 - 4 $x_b = a + 0.618(b - a)$;
 - 5 FAÇA
 - 6 se $f(x_k + x_a d_k) < f(x_k + x_b d_k)$;
 - 7 $b = x_b$;
 - 8 senão $a = x_a$;
 - 9 fim se;
 - 10 fim enquanto;
 - 11 $I = \frac{(x_a + x_b)}{2}$;
 - 12 retorna;
 - 13 Intervalo = I
-

Teorema 19. *O Algoritmo do método da Seção Áurea, na primeira fase, encontra um intervalo $[a, b]$ que contém pelo menos um minimizador da função em um número finito de iterações.*

Detalhes do Método da Seção Áurea podem ser encontrados em Izmailov e Solodov (2012).

O Algoritmo 3 descreve as fases para a implementação do método de Newton. As entradas do método são a função $f(x)$, o gradiente da função $\nabla f(x)$, a hessiana $\nabla^2 f(x)$, o erro $\varepsilon > 0$, o ponto inicial x_0 e a saída é o ponto ótimo x^* . A determinação do passo α_k é obtida pela estratégia da Seção Áurea.

Algoritmo 3: Método de Newton para Múltiplas Variáveis.

Entrada: $f(x), \nabla f(x), \nabla^2 f(x), \varepsilon > 0, x_0$

Saída: Solução ótima x^*

- 1 Inicialize $k = 0$;
 - 2 REPITA enquanto $\|\nabla f(x_k)\| \geq \varepsilon$;
 - 3 DEFINA
 - 4 $d_k = -\nabla^2 f(x_k)^{-1} \nabla f(x_k)$;
 - 5 $\alpha_k =$ mínimo da seção áurea tal que $f(x_k + \alpha_k d_k) < f(x_k)$;
 - 6 $x_{k+1} = x_k + \alpha_k d_k$;
 - 7 $k = k + 1$;
 - 8 **retorna;**
 - 9 Solução ótima = x_k^*
-

5.5 CONVERGÊNCIA DO MÉTODO DE NEWTON

O Método de Newton baseia-se na busca de um ponto ótimo onde o gradiente de $f(x)$ é nulo. No caso de funções não convexas, não existe garantia de convergência para o método de Newton. Porém, quando trata-se de problemas que envolvem funções convexas obtêm-se resultados satisfatórios.

O método de Newton apresenta convergência quadrática uma vez $\nabla f(x^*)$ tem posto completo e se x_0 é suficientemente próximo do ponto ótimo x^* .

Teorema 20. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe C^2 e que a Hessiana $\nabla^2 f(x)$ seja Lipschitz contínua em uma vizinhança da solução x^* , na qual as condições suficientes de segunda ordem são satisfeitas. Considere a iteração da forma $x_{k+1} = x_k + d_k$ com $d_k = -\nabla^2 f(x_k)^{-1} \nabla f(x_k)$. Então,*

- se o ponto inicial x_0 está suficientemente próximo de x^* , a sequência de iterandos converge a x^* ;
- a ordem de convergência de $\{x_k\}$ é quadrática;
- e a sequência da norma dos gradientes $\{\|\nabla f_k\|\}$ converge quadraticamente para 0.

O Teorema 20 garante que o método de Newton gera uma sequência convergente uma vez que as condições apresentadas são garantidas.

6 SIMULAÇÕES NUMÉRICAS

Neste Capítulo serão apresentados os resultados de simulações numéricas usando como estratégia o Método de Newton. Serão utilizadas funções da literatura para a realização dos testes computacionais para verificar a convergência do método e o tempo de processamento. Para validar os resultados utiliza-se a rotina de minimização *fminunc* presente no *Optimization Toolbox* do MATLAB. Todas as simulações foram implementadas utilizando a linguagem de programação presente no MATLAB e processadas em um computador Intel CORE i5.

6.1 ANÁLISE DOS RESULTADOS

Para a implementação do Método de Newton para minimizar funções $f(x)$ considera-se o problema de programação não linear irrestrito (3). O objetivo inicial é verificar o tempo de convergência do método de Newton utilizando a busca linear na direção d_k com escolha do passo ótimo α_k em cada iteração k , onde $0 < \alpha_k < 1$ é obtido pelo algoritmo da Seção Áurea. O critério de parada adotado é $\|\nabla f(x)\| < \varepsilon$ ou atingir um número máximo de iterações, nesse caso, fixou-se 100 iterações e $\varepsilon = 0.0001$.

Na Tabela 1 são descritos os resultados das simulações numéricas para problemas de minimização de funções convexas de única variável. A Tabela 1 apresenta o valor inicial x_0 , a quantidade de iterações k , o valor da função no ponto x^* (*fval*) e o tempo de processamento em segundos.

Tabela 1: Resultados obtidos com a implementação do Algoritmo 1.

Funções Convexas	x_0	k	fval	x^*	Tempo
$f(x) = 4x^2 - 2x - 10$	1	2	-10.2500	0.2500	0.0348
$t(x) = 2(x - 3)^2 + (x + 2)^2$	2	2	16.6667	1.3333	0.0473

No caso de problemas de PNL irrestrito para múltiplas variáveis serão exposto os resultados baseados em funções convexas presente na literatura encontradas em Moré et al. (1981). Nesse momento, o objetivo é verificar a eficiência da convergência e o tempo de processamento

do método de Newton. Os resultados foram obtidos através da implementação do Algoritmo 3 e as funções utilizadas estão descritas na Tabela 2.

Tabela 2: Funções utilizadas para avaliação do Algoritmo 3.

Funções	
Booth Function	$h(x) = ((x + (2y) - 7)^2) + ((2x + y - 5)^2)$
Parabolic Function	$p(x) = (x + 1)^2 + (y + 2)^2$
Rosenbrock Function	$r(x) = 100(y - x^2)^2 + (1 - x)^2$
Matyas Function	$m(x) = 0.26(x^2 + y^2) - 0.48xy$
Freudenstein Function	$f(x) = (-13 + x + ((5 - y)y - 2)y)^2 + (-29 + x + ((y + 1)y - 14)y)^2$
Beales Function	$l(x) = (1.5 - x(1 - y))^2 + (2.25 - x(1 - y^2))^2 + (2.625 - x(1 - y^3))^2$

Tabela 3: Resultados obtidos com a implementação do Algoritmo 3.

Funções	x_0	x^*	k	fval	tempo(s)
Booth Function	[1 1]	[1 3]	3	1.3889e-08	0.0126
Parabolic Function	[-3 3]	[-1 -2]	3	0.0422e-7	0.0078
Rosenbrock Function	[0.5 0.5]	[0.9714 0.9435]	10	0.0008	0.0167
Matyas Function	[0.5 0.5]	[0 0]	2	3.6875e-13	0.0062
Freudenstein Function	[0 -2]	[11.4093 -0.8971]	7	48.9843	0.0300
Beales Funcion	[1 0]	[2.9977 0.4994]	8	8.8371e-07	0.0070

Na Tabela 3 são apresentados os resultados obtidos com a implementação do método de Newton descrito no Algoritmo 3. Os resultados obtidos foram comparados com a rotina *fminunc* presente no Toolbox do MATLAB. O *Optimization Toolbox* do MATLAB, consiste em uma coleção de funções que estendem a capacidade do MATLAB em resolver problemas de otimização já pré-definidas. Para isso, optou-se em utilizar a rotina de minimização *fminunc*, a qual é usada para resolver problemas de otimização irrestrito. A Tabela 4 ilustra os resultados obtidos com a rotina *fminunc* presente no MATLAB.

Tabela 4: Resultados obtidos com a rotina *fminunc* presente no MATLAB.

Funções	x_0	x^*	k	fval	tempo(s)
Booth Function	[1 1]	[1 3]	9	8.3401e-12	0.0958
Parabolic Function	[-3 3]	[-1 -2]	2	4.4691e-14	2.3493
Rosenbrock Function	[0.5 0.5]	[1 1]	20	2.0062e-11	0.5570
Matyas Function	[0.5 0.5]	[0 0]	2	2.6867e-16	0.377
Freudenstein Function	[0 -2]	[11.4128 -0.8968]	19	48.9843	0.0451
Beales Function	[1 0]	[3 0.5]	11	4.3262e-14	5.5938

Ao analisar os resultados obtidos nas Tabelas 3 e 4, observa-se que o método do Newton solucionou todos os problemas de forma satisfatória, ou seja, todas as funções convergiram, uma vez que a sequência gerada x_k esta se aproximando do mínimo da função. Entretanto, a rotina *fminunc* presente no *Toolbox* do MATLAB mostrou-se menos eficiente, uma vez que

apresentou um número maior de iterações e o tempo de processamento foi superior. A Tabela 5 sintetiza dos resultados obtidos com o uso da implementação do método de Newton e da rotina *fminunc*. Observa-se que k_i , $i = 1, 2$ representa o número de iterações, respectivamente e $tempo_i$, $i = 1, 2$ o tempo de processamento, respectivamente. Note que ambas as estratégias resolveram os problemas de forma satisfatória, porém, apresentaram resultados diferentes em termos de número de iteração e tempo. Esta diferença pode ser justificado pelo fato da rotina *fminunc* utilizar outros métodos de solução, como o método Quase-Newton que consiste em realizar aproximação da inversa da matriz hessiana. Neste caso, obtêm-se uma aproximação para a matriz hessiana $H(x_k)$ e, em seguida, é gerado uma sequência convergente x_k (MARTINEZ; SANTOS, 1995).

Tabela 5: Comparação dos resultados

Function	k_1	k_2	$tempo_1(s)$	$tempo_2(s)$
Booth Function	3	9	0.0126	0.0958
Parabolic Function	3	2	0.0078	2.3493
Rosenbrock Function	10	20	0.0167	0.5570
Matyas Function	2	2	0.0062	0.377
Freudenstein Function	7	19	0.0300	0.0451
Beales Funcion	8	11	0.0070	5.5938

A seguir, apresentam-se a análise gráfica dos problemas apresentados, a fim de melhor ilustrar os resultados obtidos.

A Figura 12 ilustra o comportamento da função Booth's dada por $h(x) = (x + 2y - 7)^2 + (2x + y - 5)^2$. A superfície correspondente a função de Booth's é ilustrada na Figura 13a. As curvas de nível são descritas na Figura 13b. Note que a norma do vetor gradiente, critério de parada do algoritmo, tende a zero em apenas 3 iterações como é apresentada na Figura 13c.

O comportamento da função de Parabolic dada por $p(x) = (x + 1)^2 + (y + 2)^2$ é ilustrado na Figura 13. A superfície correspondente a função de Parabolic é ilustrada na Figura 14a. As curvas de nível são descritas na Figura 14b. Note que a norma do vetor gradiente, critério de parada do algoritmo, tende a zero em apenas 3 iterações como é apresentada na Figura 14c.

A Figura 14 ilustra o comportamento da função de Rosenbrock dada por $r(x) = 100(y - x^2)^2 + (1 - x)^2$. A superfície correspondente a função de Rosenbrock é ilustrada na Figura 15a. As curvas de nível são descritas na Figura 15b. Note que a norma do vetor gradiente, critério de parada do algoritmo, tende a zero em 10 iterações como é apresentada na Figura 15c.

O comportamento da função Matyas dado por $m(x) = 0.26(x^2 + y^2) - 0.48xy$ é ilustrado na Figura 15. A superfície correspondente a função de Matyas é ilustrada na Figura 16a. As curvas de nível são descritas na Figura 16b. Note que a norma do vetor gradiente, critério

de parada do algoritmo, tende a zero em apenas 2 iterações como é apresentada na Figura 16c.

O comportamento da função Freudenstein dada por $f(x) = (-13 + x + ((5 - y)y - 2)y)^2 + (-29 + x + ((y + 1)y - 14)y)^2$ é ilustrado na Figura 16. A superfície correspondente a função de Freudenstein é ilustrada na Figura 17a. As curvas de nível são descritas na Figura 17b. Note que a norma do vetor gradiente, critério de parada do algoritmo, tende a zero em 7 iterações como é apresentada na Figura 17c.

A Figura 17 ilustra o comportamento da função de Beales dada por $l(x) = (1.5 - x(1 - y))^2 + (2.25 - x(1 - y^2))^2 + (2.625 - x(1 - y^3))^2$. A superfície correspondente a função de Beales é ilustrada na Figura 18a. As curvas de nível são descritas na Figura 18b. Note que a norma do vetor gradiente, critério de parada do algoritmo, tende a zero em 8 iterações como é apresentada na Figura 18c.

Note que a norma do gradiente, em todas as funções analisadas, tende a zero satisfazendo uma das condições de convergência do método de Newton.

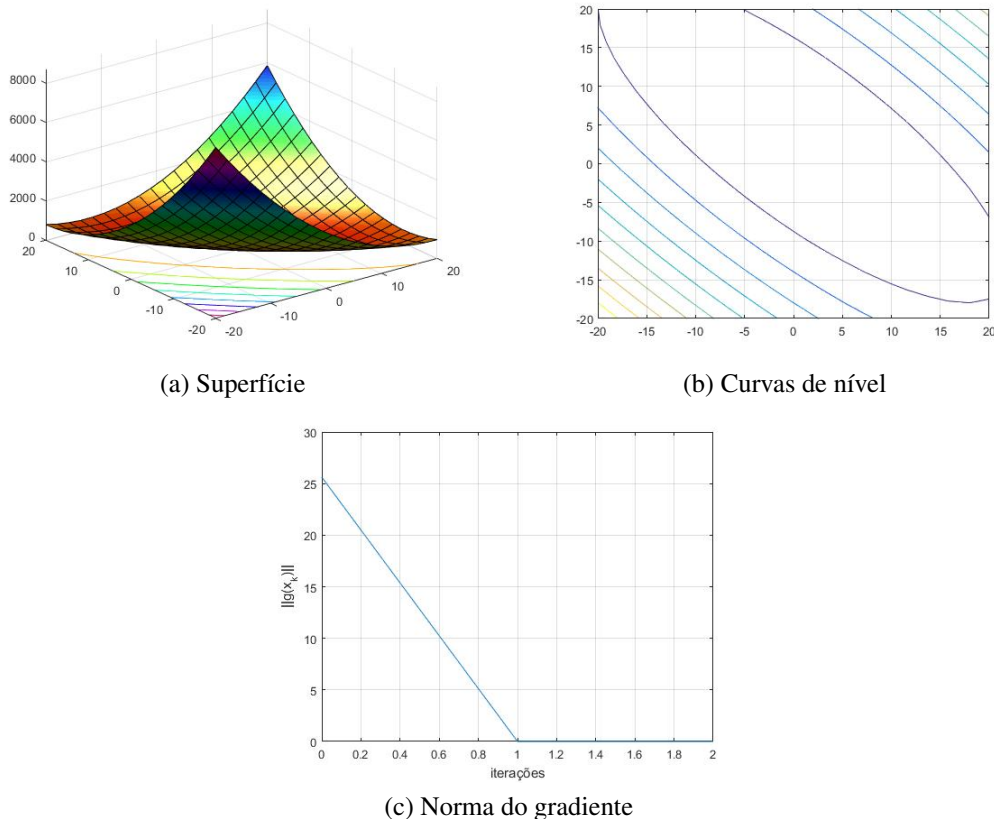


Figura 12: Análise gráfica da função Booth's.

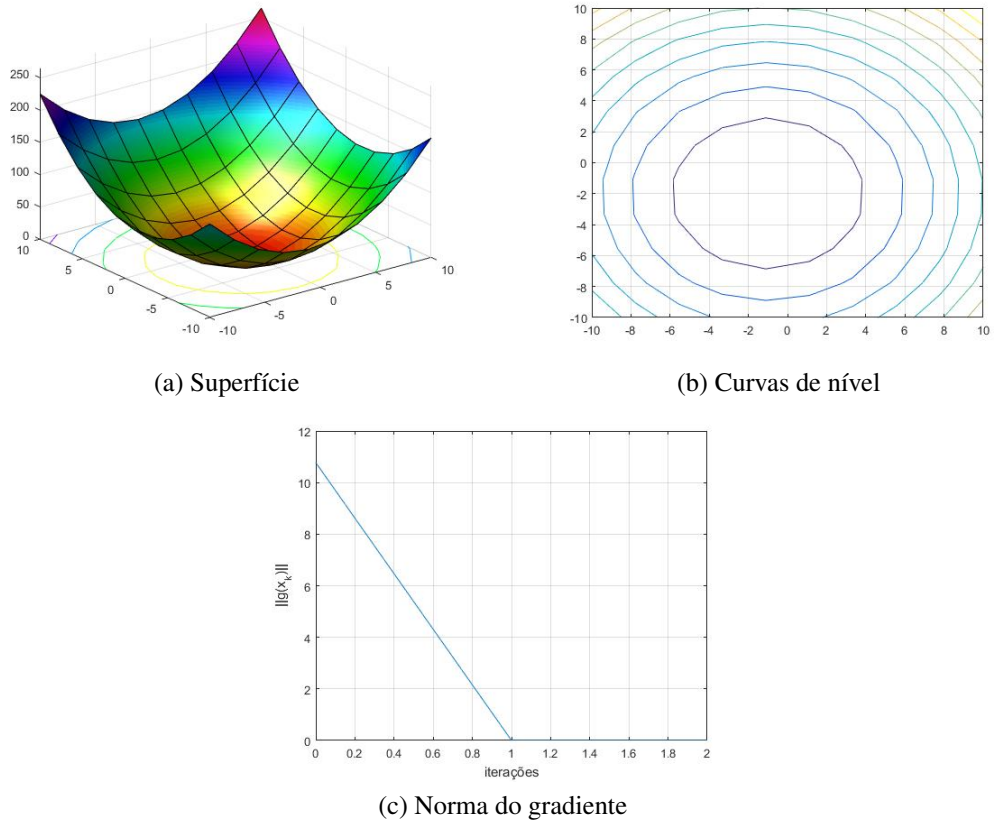


Figura 13: Análise gráfica da função Parabolic.

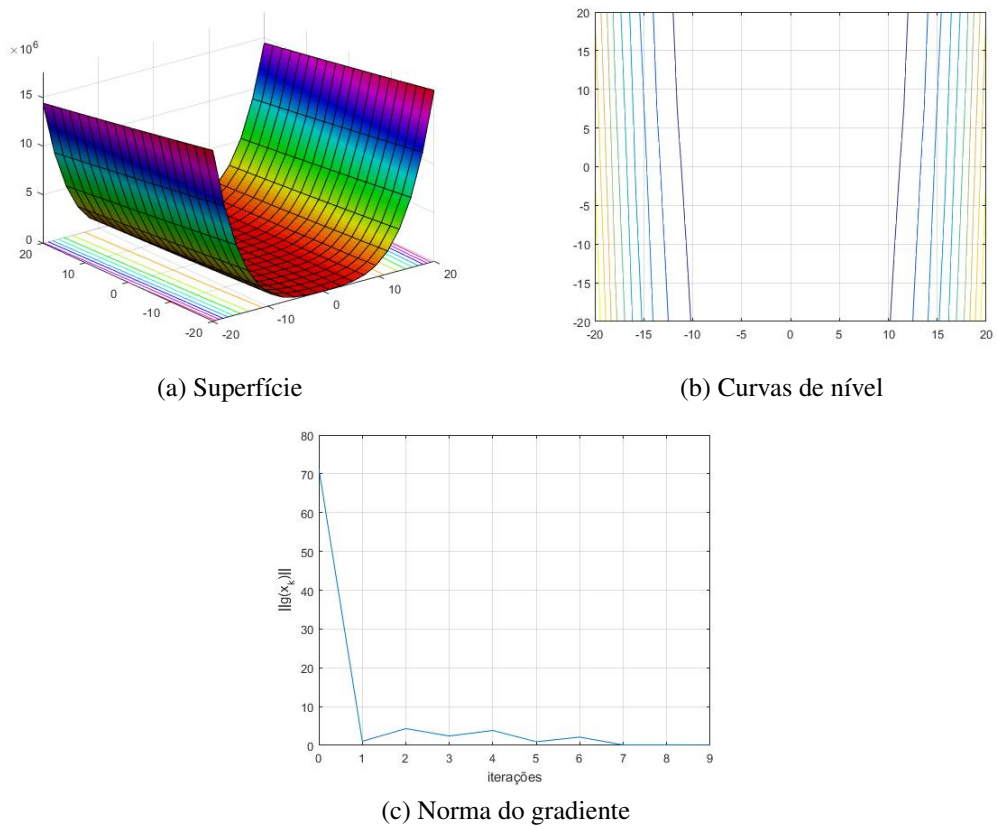


Figura 14: Análise gráfica da função Rosenbrock.

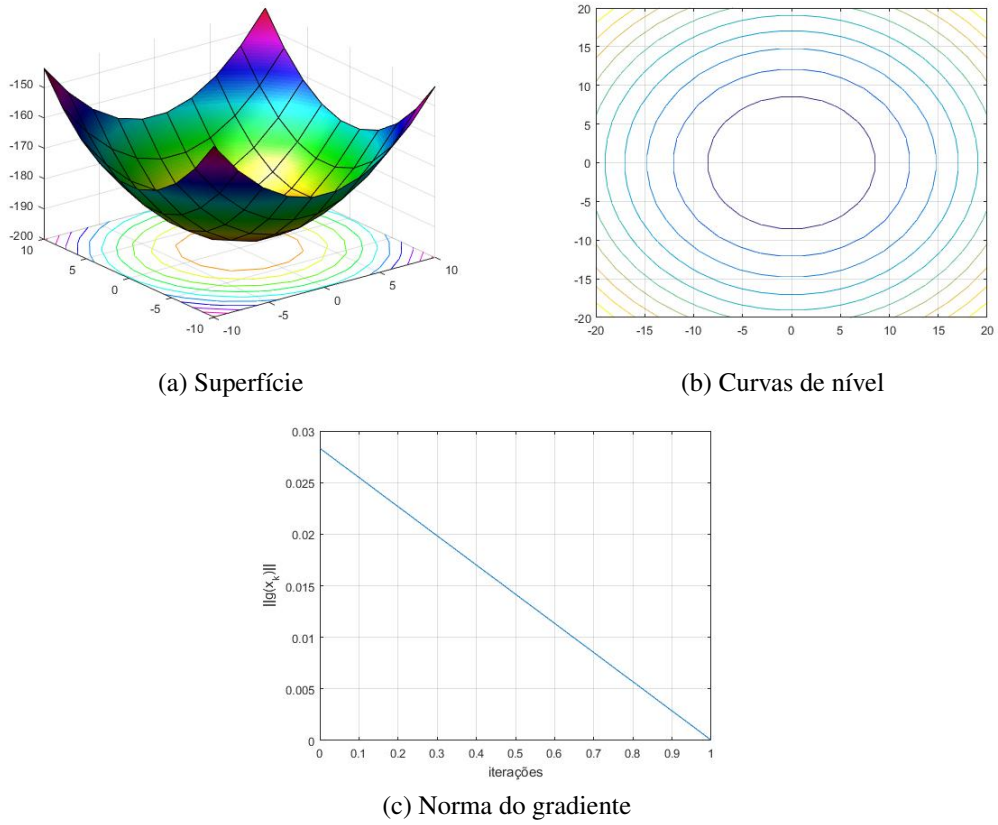


Figura 15: Análise gráfica da função Matyas.

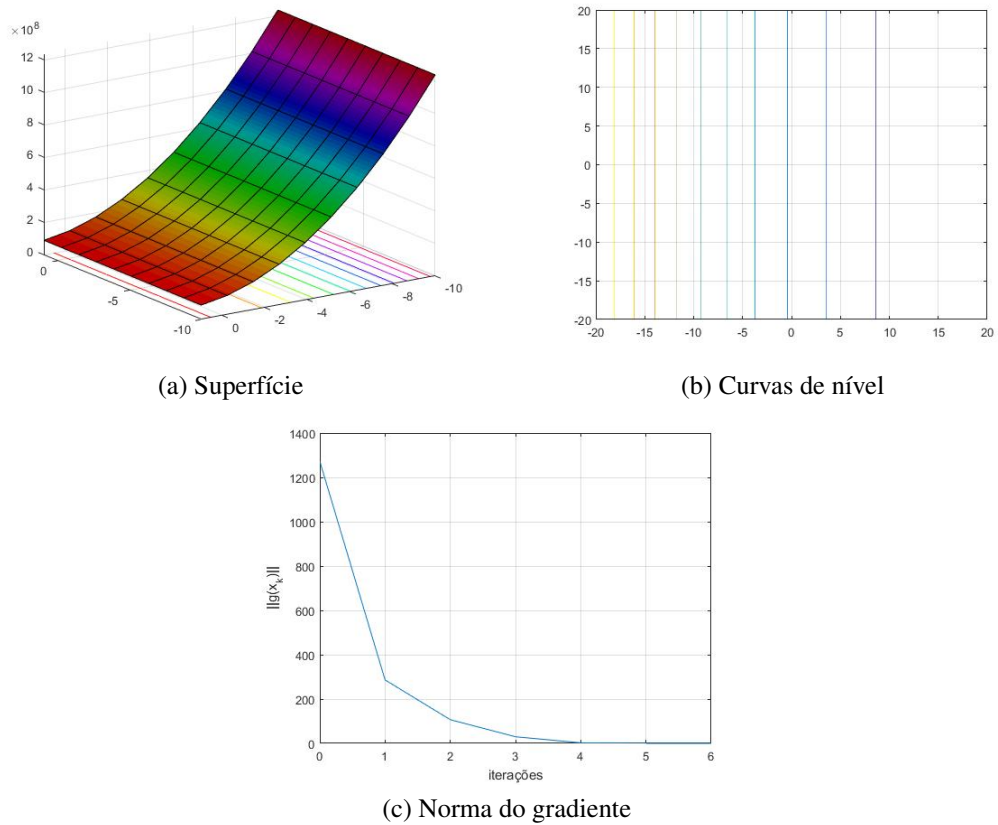


Figura 16: Análise gráfica da Freudenstein Function

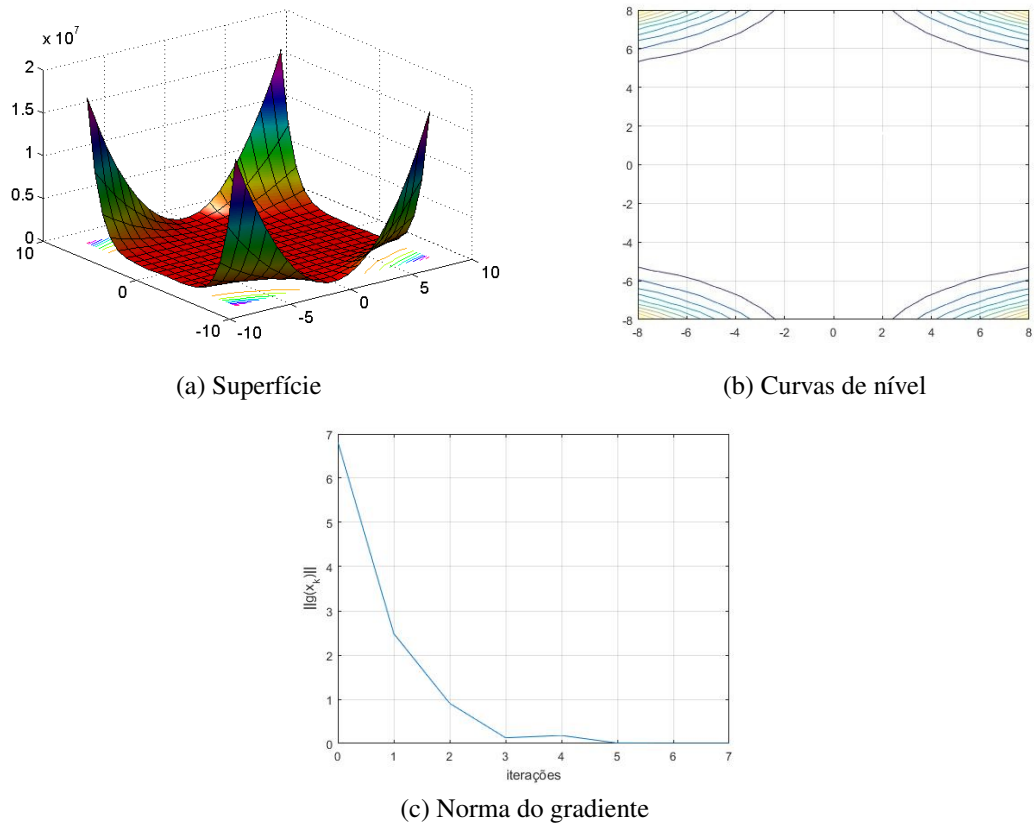


Figura 17: Análise gráfica da função Beales.

7 CONCLUSÃO

No presente trabalho foram apresentadas ferramentas matemáticas utilizadas para a solução de problemas de otimização não linear irrestrita. E ainda, foi apresentada a fundamentação teórica para problemas de minimização convexa a qual garante a existência de um minimizador global para problemas de otimização irrestrito. Destaca-se, também, a existência de soluções e as condições de otimalidade para problemas convexos.

Os resultados numéricos foram obtidos através do estudo do método de Newton para minimização de funções de uma única variável e para funções de múltiplas variáveis. Observou-se que para ambos os problemas abordados o método converge rapidamente para a solução local, o que corresponde ao ótimo global no caso de funções convexas.

Vale ressaltar que o método de Newton implementado apresenta uma velocidade computacional satisfatória uma vez que o tempo de processamento foi superior comparado à rotina presente no MATLAB. Outra observação importante do método de Newton é que se a matriz Hessiana for singular o passo α_k pode não estar bem definido, podendo não apresentar uma direção de descida. Caso contrário, se a matriz Hessiana é definida positiva, $\nabla^2 f > 0$ e o passo α_k for bem definido, o método de Newton apresenta uma direção de descida e, portanto, se o problema envolver funções convexas tem-se a garantia de convergência global do problemas de forma eficiente.

A escolha do passo α_k foi baseado no método da Seção Áurea, pois com isso, pode-se obter uma melhora na direção de busca, uma vez que está se otimizando o passo α_k . O método da Seção Áurea auxilia no melhor desempenho do método de Newton, uma vez que a cada iteração está se buscando o melhor passo.

A partir da implementação do método de Newton comparada com a rotina de minimização *fminunc* presente no *Optimization Toolbox* do MATLAB, verificou-se que os resultados apresentados são satisfatórios uma vez que a velocidade de convergência foi superior e o número de iterações foi menor.

Sendo assim, a partir dos estudos realizados, observa-se que o uso de ferramentas

computacionais para a resolução dos problemas de otimização são fundamentais quando se trata de problemas não lineares. Com isso, auxiliando na tomada de decisão dos processos de forma rápida e eficiente.

REFERÊNCIAS

- ARENALES, M. et al. **Pesquisa Operacional**. Rio de Janeiro: Elsevier, 2007.
- BOLDRINI, J. L. et al. **Álgebra linear**. [S.l.]: Harper & Row do Brasil, 1980.
- COOPPE. **Pesquisa Operacional**. 2015. Disponível em: <<http://www.po.coppe.ufrj.br>>. Acesso em: 13 de abril de 2016.
- FRIEDLANDER, A. **Elementos de programação não-linear**. [S.l.]: Editora da UNICAMP, 1994.
- GASS, S. I. **Great Moments in History**. 2002. Disponível em: <<http://www.orms-today.org/orms-10-02/frhistorysb1.html>>. Acesso em: 20 de fevereiro de 2016.
- GUIDORIZZI, H. L. **Um curso de cálculo Volume 1**. Rio de Janeiro: LTC, 2001.
- IZMAILOV, A.; SOLODOV, M. **Otimização Volume 1. Condições de otimalidade, elementos de análise convexa e dualidade**. Rio de Janeiro: IMPA, 2009.
- IZMAILOV, A.; SOLODOV, M. **Otimização Volume 2. Métodos Computacionais**. Rio de Janeiro: IMPA, 2012.
- LACHTERMACHER, G. **Pesquisa operacional na tomada de decisões: modelagem em Excel**. [S.l.]: Elsevier, 2007.
- LIMA, E. L. **Curso de Análise Volume 1**. São Paulo: SBM, 1999.
- LIMA, E. L. **Análise Real Vol. 2. Funções de n Variáveis**. [S.l.]: IMPA, Brazil, 2006.
- MACHADO, I. de A.; ALVES, R. R. **MÉTODO DE NEWTON**. 2013. 30 p.
- MARTINEZ, J. M.; SANTOS, S. A. **Métodos computacionais de otimização**. [S.l.]: Colóquio Brasileiro de Matemática, Apostilas, 1995.
- MORÉ, J. J.; GARBOW, B. S.; HILLSTROM, K. E. Testing unconstrained optimization software. **ACM Transactions on Mathematical Software (TOMS)**, ACM, v. 7, n. 1, p. 17–41, 1981.
- NOCEDAL, J.; WRIGHT, S. **Numerical optimization**. [S.l.]: Springer Science & Business Media, 2006.
- RIBEIRO, A. A.; KARAS, E. W. **Otimização Contínua: Aspectos Teóricos e Computacionais**. São Paulo: Cengage Learning, 2013.
- STIEGELMEIER, E. W. **Modelo de otimização para o controle de plantas daninhas usando programação não linear inteira mista**. Tese (Doutorado) — Universidade de São Paulo, 2012.

VALENTINO, M. C. **Controle ótimo de doenças fúngicas no feijoeiro**. Dissertação (Mestrado) — Centro de Matemática, Computação e Cognição – Universidade Federal do ABC, Santo André/São Paulo, 2009.