

Lucas Marchand

# **Multiplicadores de Lagrange: uma aplicação em problemas de otimização global restrita**

Rio Grande, Rio Grande do Sul, Brasil

Novembro, 2016

Lucas Marchand

## **Multiplicadores de Lagrange: uma aplicação em problemas de otimização global restrita**

Trabalho de Conclusão de Curso, Matemática Aplicada Bacharelado, submetido por Lucas Marchand junto ao Instituto de Matemática, Estatística e Física da Universidade Federal do Rio Grande.

Universidade Federal do Rio Grande - FURG

Instituto de Matemática, Estatística e Física - IMEF

Curso de Matemática Aplicada Bacharelado

Orientador: Dra. Bárbara Denicol do Amaral Rodriguez

Coorientador: Dr. Leonardo Ramos Emmendorfer

Rio Grande, Rio Grande do Sul, Brasil

Novembro, 2016

*Este trabalho é dedicado aos meus pais Lucio Floro Machado de Sousa e Carmen Rosane Saraiva Marchand, aos meus Irmãos Guilherme de Sousa e Fernanda de Sousa e aos meus professores do IMEF, por terem me proporcionado condições de enfrentar este desafio e concluí-lo.*

# Agradecimentos

Agradeço pela grande ajuda dos professores do IMEF, em especial à orientadora Dra. Bárbara Denicol do Amaral Rodriguez, ao coorientador Dr. Leonardo Ramos Emendorfer e ao Dr. Mario Rocha Retamoso pelo incentivo para a conclusão do curso. Aos amigos que percorreram, ao meu lado, o caminho do meio acadêmico, àqueles que fazem parte do meu dia a dia e aos amigos professores da Escola de Ensino Médio Roberto Basto Tellechea.

*“ Acredite em si próprio e chegará um dia em que os outros não terão outra escolha  
senão acreditar com você”  
(Cynthia Kersey)*

# Resumo

A teoria dos Multiplicadores de Lagrange é aplicada para determinar pontos de máximos e mínimos de funções sujeitas a restrições. Neste trabalho é avaliada a eficiência deste método quando aplicado a problemas de otimização com restrições de igualdade. Os resultados obtidos são visualizados no software de geometria dinâmica GeoGebra e comparados com dados disponíveis na literatura.

**Palavras-chaves:** Multiplicadores de Lagrange, otimização, restrições, GeoGebra, Hessiano Orlado.

# Abstract

The Lagrange Multipliers theory is applied to determine maximum and minimum points of constrained functions. This work evaluates the efficiency of this method when applied to optimization equality constraints problems. The obtained results are visualized in the dynamic geometry software GeoGebra and compared with data available in the literature.

**Key-words:** Lagrange Multiplier, optimization, restrictions, GeoGebra, Hessiano Orlado.

# Lista de ilustrações

|                                                                                                |    |
|------------------------------------------------------------------------------------------------|----|
| Figura 1 – Representação Gráfica da Superfície $S$ . . . . .                                   | 14 |
| Figura 2 – Curvas de Níveis de $f(x, y)$ . . . . .                                             | 14 |
| Figura 3 – Representação do Vetor Unitário $\mathbf{u}$ . . . . .                              | 16 |
| Figura 4 – $D_{\mathbf{u}}f$ no $P$ na Direção Unitário $u$ . . . . .                          | 16 |
| Figura 5 – Máximos e Mínimos da Intersecção $g_1$ e $g_2$ . . . . .                            | 26 |
| Figura 6 – Fluxograma PSO. . . . .                                                             | 32 |
| Figura 7 – Representação do ponto $(0, 0, 1)$ para o 1º caso, $x = 0$ , da função G11. . . . . | 35 |
| Figura 8 – Representação do ponto Crítico para o 2º caso da função G11. . . . .                | 36 |
| Figura 9 – Pontos de máximos e mínimos para a função G11. . . . .                              | 39 |
| Figura 10 – Pontos de máximo e mínimos para G3. . . . .                                        | 41 |



# Lista de tabelas

|                                                |    |
|------------------------------------------------|----|
| Tabela 1 – Comparação dos Resultados . . . . . | 39 |
|------------------------------------------------|----|

# Sumário

|            |                                                                     |           |
|------------|---------------------------------------------------------------------|-----------|
|            | <b>Introdução</b>                                                   | <b>11</b> |
| <b>1</b>   | <b>FUNDAMENTAÇÃO MATEMÁTICA</b>                                     | <b>13</b> |
| <b>1.1</b> | <b>Funções Reais de Várias Variáveis</b>                            | <b>13</b> |
| 1.1.1      | Funções Reais de Duas Variáveis                                     | 13        |
| 1.1.2      | Funções Reais de Três Variáveis.                                    | 14        |
| 1.1.2.1    | Superfícies de Nível                                                | 15        |
| <b>1.2</b> | <b>Derivadas Parciais</b>                                           | <b>15</b> |
| <b>1.3</b> | <b>Derivadas Direcionais</b>                                        | <b>15</b> |
| 1.3.1      | Vetor Gradiente                                                     | 17        |
| <b>1.4</b> | <b>Máximos e Mínimos de Funções de Várias Variáveis</b>             | <b>17</b> |
| 1.4.1      | Máximos e Mínimos de Funções de Duas Variáveis                      | 17        |
| 1.4.2      | Ponto Crítico de uma Função de Duas Variáveis                       | 18        |
| 1.4.2.1    | Valores Máximo e Mínimo Absolutos.                                  | 18        |
| <b>2</b>   | <b>OTIMIZAÇÃO</b>                                                   | <b>20</b> |
| <b>2.1</b> | <b>Definições e Alguns Fatos Básicos</b>                            | <b>20</b> |
| <b>2.2</b> | <b>Existência de Soluções Globais</b>                               | <b>21</b> |
| <b>2.3</b> | <b>Condições de Otimalidade</b>                                     | <b>22</b> |
| <b>2.4</b> | <b>Multiplicadores de Lagrange, Máximos e Mínimos Condicionados</b> | <b>24</b> |
| 2.4.1      | Problemas Envolvendo Funções a Três Variáveis e Duas Restrições     | 26        |
| <b>2.5</b> | <b>O Hessiano Aumentado (Orlado)</b>                                | <b>27</b> |
| 2.5.1      | O Caso de $n$ -variáveis (CHIANG, 1982).                            | 27        |
| 2.5.2      | Caso de Múltiplas Restrições                                        | 29        |
| <b>3</b>   | <b>ALGORITMOS DE OTIMIZAÇÃO</b>                                     | <b>31</b> |
| <b>3.1</b> | <b>Algoritmos Evolutivos</b>                                        | <b>31</b> |
| 3.1.1      | <i>Particle Swarm Optimization</i> - PSO (Enxame de Partículas)     | 31        |
| <b>4</b>   | <b>RESULTADOS NUMÉRICOS</b>                                         | <b>33</b> |
| <b>4.1</b> | <b>Grupo de Funções G</b>                                           | <b>33</b> |
| <b>4.2</b> | <b>Descrição do Problema 1 - Função G11</b>                         | <b>33</b> |
| 4.2.1      | Desenvolvimento do Problema                                         | 34        |
| <b>4.3</b> | <b>Descrição do Problema 2 - Função G3</b>                          | <b>40</b> |
| 4.3.1      | Desenvolvimento do Problema                                         | 40        |
| <b>5</b>   | <b>CONCLUSÕES</b>                                                   | <b>45</b> |

|                       |    |
|-----------------------|----|
| REFERÊNCIAS . . . . . | 46 |
|-----------------------|----|

|        |    |
|--------|----|
| ANEXOS | 48 |
|--------|----|

|     |                                                      |    |
|-----|------------------------------------------------------|----|
|     | ANEXO A – MÉTODOS DE OTIMIZAÇÃO IRRESTRITA . . . . . | 49 |
| A.1 | O Método do Gradiente . . . . .                      | 49 |
| A.2 | O Método de Newton . . . . .                         | 49 |
|     | ANEXO B – MÉTODOS DE OTIMIZAÇÃO RESTRITA . . . . .   | 50 |
| B.1 | O Método de Direções Viáveis . . . . .               | 50 |
| B.2 | O Método das Penalidades . . . . .                   | 50 |

# Introdução

No mundo moderno, cada vez mais as pessoas estão interessadas em encurtar distâncias para ganhar tempo ou economizar recursos financeiros. Agindo desse modo, elas estão buscando maneiras de otimizar a aplicação de seus recursos. Essa busca constante pelo melhor desempenho, teve início no final da década de 30 e início da década de 40 do século XX, com a situação de guerra e competição em que viviam e o forte desenvolvimento industrial. Nessa época, a gestão dos recursos impunha que utilizassem técnicas sofisticadas de otimização que acabaram por, de uma forma ou de outra, cair no âmbito da matemática (SOARES, 2004).

Nesse sentido, a essência da otimização é melhorar algo em um conjunto de alternativas disponíveis. Este algo tem uma representação matemática que recebe o nome de função objetivo ou índice de performance (ID). Mais especificamente, resolver um problema de otimização, consiste em determinar uma solução ou um conjunto de soluções ótimas para uma determinada função ou conjunto de funções. Trata-se de uma ferramenta de grande aplicabilidade que se constitui em uma vasta e atraente área do conhecimento, sendo de grande importância não apenas para as engenharias e ciências exatas, como também para as mais diversas áreas de biologia e tecnologia (OLIVEIRA, 2012).

Vários problemas clássicos de otimização com restrições nas áreas de engenharia e matemática aplicada vêm sendo estudados utilizando-se diversas técnicas encontradas na literatura em que um dos objetivos é minimizar o custo computacional necessário para a localização da solução ótima. Uma das formas mais comuns de tratar os problemas com restrições consiste em transformá-los em problemas sem restrições.

À medida que o número de funções e o número de variáveis aumenta, a dificuldade de se determinar o conjunto de soluções ótimas também aumenta. É neste contexto que surge a necessidade de desenvolver técnicas matemáticas e computacionais que refinem o processo de otimização, dado que este é amplamente utilizado para resolver problemas de engenharia (SARAMAGO; JÚNIOR, 2008).

Os métodos de otimização podem ser classificados em dois grandes grupos:

Os métodos randômicos ou aleatórios (*Random Strategies*), ditos Não - Determinísticos. Pertencem a esse grupo os Métodos Evolutivos, como por exemplo, os Algoritmos Genéticos (SARAMAGO; JÚNIOR, 2008). Esse último é baseado nas teorias de Darwin sobre a seleção natural, reprodução genética e evolução das espécies. Esse método não garante uma solução ótima, mas pode atingir um resultado sub-ótimo com o custo computacional razoável (SILVA; LEMONGE; LIMA, 2014).

Os métodos baseados no cálculo (*Deterministic Optimization*) são chamados de Métodos Determinísticos. Dentre estes, está o método dos Multiplicadores de Lagrange. Essa metodologia consiste em resolver problemas de otimização restrita e permite analisar situações mais gerais. A partir de sua aplicação, um problema de otimização restrita com  $n$  variáveis e  $m$  variáveis de igualdade é transformado em um problema de otimização sem restrições com  $(m + n)$  variáveis (GONÇALVES; FLEMMING, 2007).

O objetivo do trabalho é aplicar a teoria dos Máximos e Mínimos Condicionados, com o emprego do Método dos Multiplicadores de Lagrange na determinação do ponto ótimo global para dois problemas sujeitos a uma restrição de igualdade. Tais problemas, compõem um grupo de funções chamadas funções G. Além disso, pretende-se identificar quais são os pontos que satisfazem os problemas através de uma representação gráfica com software GeoGebra e verificar, pelo emprego do método de Hessiano Orlado, se tais pontos são ótimos.

Para verificar a metodologia proposta, os resultados obtidos são comparados com os trabalhos de Pedrosa e Afonso (PEDROSA; AFONSO, 2011), que utiliza o método de Enxame de Partículas (PSO), e de Koziel, que emprega Algoritmos Evolutivos Modificados (EAs) (KOZIEL; MICHALEWICZ, 1999).

A fim de atingir os objetivos propostos, o trabalho está organizado da seguinte forma: no Capítulo 1, abordam-se definições e teoremas relacionados à teoria matemática necessária para a elaboração deste trabalho.

No Capítulo 2, descrevem-se algoritmos de otimização evolutivos, particularmente o Enxame de Partículas.

No Capítulo 3 faz-se a descrição dos problemas abordados neste trabalho, bem como aplica-se a metodologia proposta e apresentam-se os resultados obtidos.

Por fim, as conclusões e as propostas para trabalhos futuros estão presentes no Capítulo 4.

# 1 Fundamentação Matemática

Neste capítulo apresenta-se a fundamentação teórico-matemática, base para a realização deste trabalho. São abordados os conceitos a cerca de Funções Reais de Várias Variáveis, Derivadas Parciais, Derivadas Direcionais, Máximos e Mínimos de Funções de Várias Variáveis, Otimização, Multiplicadores de Lagrange e Hessiano Orlado.

## 1.1 Funções Reais de Várias Variáveis

Nesta seção, estuda-se uma das noções centrais da Matemática, mais especificamente na área do cálculo diferencial, o conceito de função de várias variáveis.

Uma função de várias variáveis reais é uma regra que descreve como uma quantidade é determinada por outras de maneira única. Através destas funções pode-se modelar uma grande quantidade de fenômenos dos mais diversos ramos da Ciência.

Nas subseções seguintes descreve-se as definições de funções reais de duas e três variáveis.

### 1.1.1 Funções Reais de Duas Variáveis

**Definição 1.1.1. Funções Reais de Duas Variáveis.** Uma função real  $f$  de duas variáveis reais é uma relação que transforma em um único número real  $z$  cada par ordenado  $(x, y)$  de um números reais de um certo conjunto  $D$ , chamado de domínio da função. Se a relação  $f$  transforma no número real  $z$  o par ordenado  $(x, y)$  em  $D$ , então escrevemos  $z = f(x, y)$  (MUNEM; FOULIS, 1982).

Na equação  $z = f(x, y)$ , chamamos  $z$  de variável **dependente** e nos referimos a  $x$  e a  $y$  como variáveis **independentes**. O conjunto de todos os valores possíveis de  $z$ , que pode ser obtido aplicando a relação  $f$  aos pares ordenados  $(x, y)$  em  $D$ , é denominado **imagem** da função  $f$  (MUNEM; FOULIS, 1982).

**Definição 1.1.2. Representação Gráfica.** Se  $f$  é uma função de duas variáveis com domínio  $D$ , então o **gráfico** de  $f$  é o conjunto de todos os pontos  $(x, y, z)$  em  $\mathbb{R}^3$  tal que  $z = f(x, y)$  e  $(x, y) \in D$  (STEWART, 2006).

Assim como o gráfico de uma função  $f$  de uma única variável é uma curva  $C$  com equação  $y = f(x)$ , o gráfico de uma função com duas variáveis é uma superfície  $S$  de  $f$  com o domínio  $D$  sobre o plano  $xy$  (Figura 1) (STEWART, 2006).

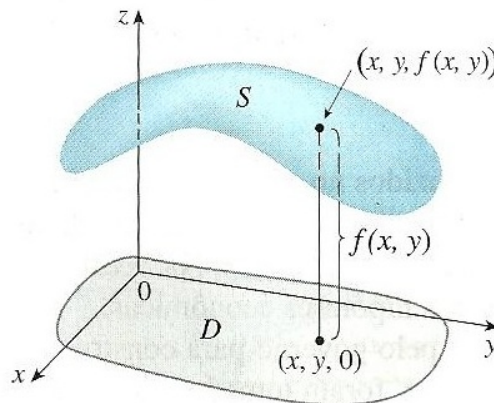


Figura 1 – Representação Gráfica da Superfície  $S$ .

Fonte: (STEWART, 2006)

**Definição 1.1.3. Curvas de Nível.** As **curvas de nível** de uma função  $f$  de duas variáveis são aquelas com equação  $f(x, y) = k$  onde  $k$  é uma constante (na imagem de  $f$ ). Em outras palavras, ela mostra onde o gráfico de  $f$  tem altura  $k$  (STEWART, 2006).

As curvas de nível  $f(x, y) = k$  são apenas traços do gráfico de  $f$  no plano horizontal  $z = k$  projetado sobre o plano  $xy$  (Figura 2).

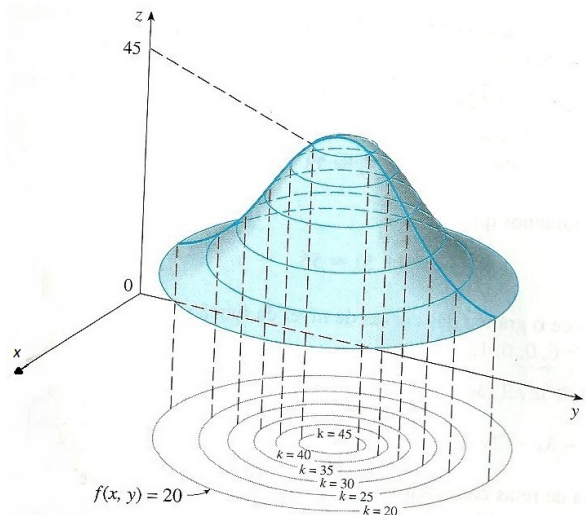


Figura 2 – Curvas de Níveis de  $f(x, y)$ .

Fonte: (STEWART, 2006)

### 1.1.2 Funções Reais de Três Variáveis.

**Definição 1.1.4. Funções Reais de Três Variáveis.** Uma função com três variáveis,  $f$ , é uma regra que associa a cada tripla ordenada  $(x, y, z)$  em um domínio  $D \subset \mathbb{R}^3$  um

único número real denotado por  $f(x, y, z)$  (STEWART, 2006).

### 1.1.2.1 Superfícies de Nível

Pelo conhecimento das **curvas de nível**, pode-se definir as **superfícies de nível** da função  $f$ .

**Definição 1.1.5. Superfícies de Nível.** As **superfícies de nível** de uma função  $f$  de três variáveis são aquelas com equação  $f(x, y, z) = k$  onde  $k$  é uma constante.

Se um ponto  $(x, y, z)$  move-se ao longo de uma superfície de nível, o valor de  $f(x, y, z)$  permanece fixo.

## 1.2 Derivadas Parciais

**Definição 1.2.1. Derivadas Parciais de Funções a Duas Variáveis.** Se  $f$  é uma função de duas variáveis e  $(x, y)$  é um ponto no domínio de  $f$ , então as derivadas parciais  $\frac{\partial f(x, y)}{\partial x}$  e  $\frac{\partial f(x, y)}{\partial y}$  de  $f$  em  $(x, y)$  em relação à primeira e à segunda variável são definidas pelas seguintes derivadas parciais

$$\begin{aligned}\frac{\partial f(x, y)}{\partial x} &= \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} \\ \frac{\partial f(x, y)}{\partial y} &= \lim_{\Delta y \rightarrow 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y},\end{aligned}$$

contanto que os limites existam. O procedimento para calcular as derivadas parciais é denominado **diferenciação parcial** (MUNEM; FOULIS, 1982).

**Definição 1.2.2. Derivadas Parciais de Funções a  $n$  Variáveis.** Seja  $f$  é uma função a  $n$  variáveis e suponha que  $(x_1, x_2, \dots, x_k, \dots, x_n)$  pertença ao domínio de  $f$ . Se  $1 \leq k \leq n$ . Então a derivada parcial de  $f$  em relação à  $k$ -ésima variável  $x_k$  é denotada por  $f_k$  e definida por

$$\frac{\partial f(x_1, \dots, x_n)}{\partial x_k} = \lim_{\Delta x_k \rightarrow 0} \frac{f(x_1, x_2, \dots, x_k + \Delta x_k, \dots, x_n) - f(x_1, x_2, \dots, x_k, \dots, x_n)}{\Delta x_k},$$

deste que o limite exista (MUNEM; FOULIS, 1982).

## 1.3 Derivadas Direcionais

**Definição 1.3.1. Derivada Direcional.**(STEWART, 2006) A **Derivada Direcional** de  $f$  em  $(x_0, y_0)$  na direção e sentido do vetor unitário  $\mathbf{u} = (a, b)$  é

$$D_{\mathbf{u}}f(x_0, y_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h},$$



se esse limite existir.

Suponha que se quer determinar a taxa de variação de  $z$  no ponto  $(x_0, y_0)$  na direção e sentido de um vetor unitário arbitrário  $\mathbf{u} = (a, b)$  (Figura 3).

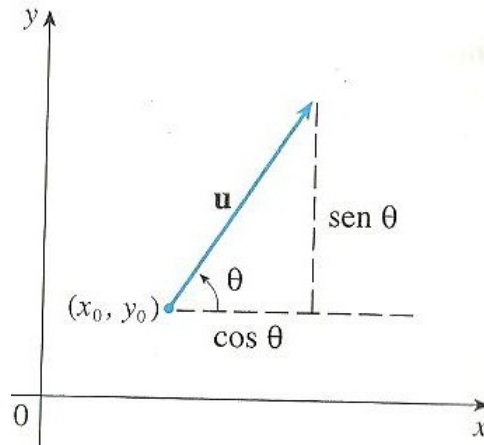


Figura 3 – Representação do Vetor Unitário  $\mathbf{u}$ .

Fonte: (STEWART, 2006)

Para fazê-lo, deve-se considerar a superfície  $S$  com a equação  $z = f(x, y)$  e tomar  $z_0 = f(x_0, y_0)$ . O ponto  $P(x_0, y_0, z_0)$  pertence a  $S$ . O plano vertical que passa por  $P$  na direção de  $\mathbf{u}$  intercepta  $S$  em uma curva  $C$  (Figura 4). A inclinação da reta tangente  $T$  a  $C$  em  $P$  é a taxa de variação de  $z$  na direção e sentido de  $\mathbf{u}$  (STEWART, 2006).

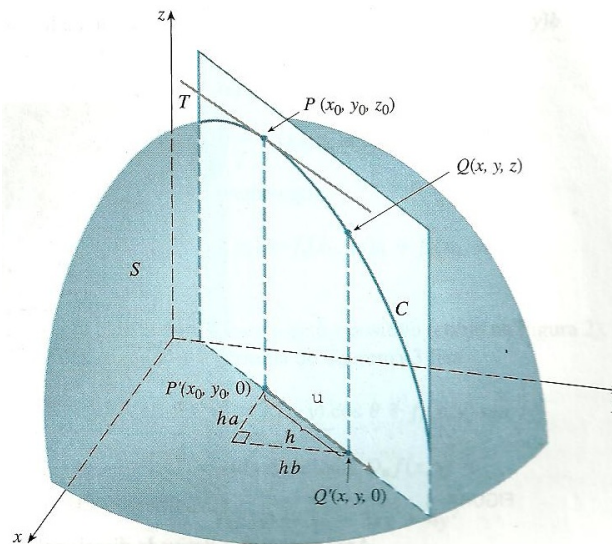


Figura 4 –  $D_{\mathbf{u}}f$  no  $P$  na Direção Unitário  $\mathbf{u}$ .

Fonte: (STEWART, 2006)

Se  $Q(x, y, z)$  é outro ponto sobre  $C$  e  $P', Q'$  são as projeções de  $P, Q$  sobre o plano  $xy$ , então o vetor  $\vec{P'Q'}$  é paralelo a  $\mathbf{u}$ , e portanto

$$\vec{P'Q'} = h\mathbf{u} = (ha, hb),$$

para algum valor do escalar  $h$ .

Dessa forma,  $x - x_0 = ha, y - y_0 = hb$ , logo  $x = x_0 + ha, y = y_0 + hb$ , e

$$\frac{\Delta z}{h} = \frac{z - z_0}{h} = \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}.$$

Tomando-se o limite quando  $h \rightarrow 0$ , obtém-se a taxa de variação de  $z$  (em relação à distância) na direção e sentido de  $\mathbf{u}$ , que é chamada derivada direcional de  $f$  na direção e sentido de  $\mathbf{u}$ .

**Teorema 1.3.1.** (STEWART, 2006) Se  $f$  é uma função diferenciável em  $x$  e  $y$ , então  $f$  tem derivada direcional na direção e sentido de qualquer versor  $\mathbf{u} = (a, b)$  e

$$D_{\mathbf{u}}f(x, y) = f_x(x, y)a + f_y(x, y)b.$$

### 1.3.1 Vetor Gradiente

Note que no Teorema 1.3.1 a derivada direcional pode ser escrita como o produto escalar entre dois vetores:

$$D_{\mathbf{u}}f(x, y) = f_x(x, y)a + f_y(x, y)b = \langle f_x(x, y), f_y(x, y) \rangle \cdot \langle a, b \rangle = \langle f_x(x, y), f_y(x, y) \rangle \cdot \mathbf{u}. \quad (1.1)$$

O primeiro vetor no produto escalar na equação (1.1) ocorre não somente no cálculo da derivada direcional, mas também em muitas outras situações. Logo chamaremos de **o gradiente de  $f$** , em uma notação matemática **grad** $f$  ou  $\nabla f$  e lê-se (del  $f$ ).

**Definição 1.3.2.** (STEWART, 2006) Se  $f$  é uma função de duas variáveis  $x$  e  $y$ , o **gradiente** de  $f$  é a função vetorial  $\nabla f$  definida por

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j}.$$

## 1.4 Máximos e Mínimos de Funções de Várias Variáveis

### 1.4.1 Máximos e Mínimos de Funções de Duas Variáveis

**Definição 1.4.1.** (GONÇALVES; FLEMMING, 2007) Seja  $z = f(x, y)$  uma função de duas variáveis. Diz que  $(x_0, y_0) \in D(f)$  é ponto de máximo absoluto ou global de  $f$  se,  $\forall (x, y) \in D(f), f(x, y) \leq f(x_0, y_0)$ .

Diz-se ainda que  $f(x_0, y_0)$  é o valor máximo de  $f$ .

**Definição 1.4.2.** (GONÇALVES; FLEMMING, 2007) Seja  $z = f(x, y)$  uma função de duas variáveis. Diz-se que  $(x_0, y_0) \in D(f)$  é ponto de mínimo absoluto ou global de  $f$  se,  $\forall (x, y) \in D(f)$ ,  $f(x, y) \geq f(x_0, y_0)$ .

Diz-se ainda que  $f(x_0, y_0)$  é o valor mínimo de  $f$ .

**Definição 1.4.3.** (GONÇALVES; FLEMMING, 2007) Seja  $z = f(x, y)$  uma função de duas variáveis. Diz-se que:

- $(x_0, y_0) \in D(f)$  é ponto de máximo relativo ou local de  $f$  se existir um ponto aberto  $B((x_0, y_0) : r) \mid f(x, y) \leq f(x_0, y_0), \forall (x, y) \in B \cap D(f)$ .
- $(x_0, y_0) \in D(f)$  é ponto de mínimo relativo ou local de  $f$  se existir um ponto aberto  $B((x_0, y_0) : r) \mid f(x, y) \geq f(x_0, y_0), \forall (x, y) \in B \cap D(f)$ .

### 1.4.2 Ponto Crítico de uma Função de Duas Variáveis

Se as inequações da definição 1.4.3 valem para todos os pontos  $(x, y)$  do domínio de  $f$ , então  $f$  tem um **máximo absoluto** (ou **mínimo absoluto**) em  $(x_0, y_0)$  (STEWART, 2006).

**Teorema 1.4.1.** Se uma função  $f$  tem um máximo ou mínimo locais em  $(x_0, y_0)$  e as derivadas parciais de primeira ordem de  $f$  existem nesses pontos, então  $f_x(x_0, y_0) = 0$  e  $f_y(x_0, y_0) = 0$  (STEWART, 2006).

#### 1.4.2.1 Valores Máximo e Mínimo Absolutos.

**Proposição 1.4.1.** Seja  $z = f(x, y)$  uma função cujas derivadas parciais de 1ª e 2ª ordem são contínuas em um conjunto aberto que contém  $(x_0, y_0)$  e  $(x_0, y_0)$  seja um ponto crítico de  $f$ . Considera-se  $H(x, y)$  o determinante

$$H(x, y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = f_{xx}f_{yy} - (f_{xy})^2.$$

1. Se  $H(x_0, y_0) > 0$  e  $f_{xx}(x_0, y_0) > 0$ , então  $f(x_0, y_0)$  é um mínimo local.
2. Se  $H(x_0, y_0) > 0$  e  $f_{xx}(x_0, y_0) < 0$ , então  $f(x_0, y_0)$  é um máximo local.
3. Se  $H(x_0, y_0) < 0$ , então  $(x_0, y_0)$  não é extremante local. Nesse caso  $(x_0, y_0)$  é um ponto sela
4. Se  $H(x_0, y_0) = 0$ , nada se pode afirmar.

**Observação 1.4.1.** A matriz

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

é conhecida como matriz Hessiana. O seu determinante,  $H(x, y)$ , é chamado determinante Hessiano da função  $z = f(x, y)$  (GONÇALVES; FLEMMING, 2007).

Na seção seguinte descrevem-se conceitos importantes acerca de otimização.

## 2 Otimização

Otimização consiste em determinar uma solução ou um conjunto de soluções ótimas para uma determinada função ou conjunto de funções.

O conceito de solução ótima é específico do problema que se deseja otimizar. No entanto, pode-se ter uma única solução ou um conjunto de soluções ou ainda não haver solução que satisfaça todas as funções.

À medida que o número de funções e o número de variáveis aumentam, a dificuldade em se determinar o conjunto de soluções ótimas também aumentam. É neste contexto que surge a necessidade de desenvolver técnicas matemáticas e computacionais que refinem o processo de otimização, uma vez que este é amplamente utilizado para resolver problemas nas áreas de Engenharia, Economia, Negócios, Física, Química, Biologia, Arquitetura e Administração.

Os métodos para a solução de problemas de otimização dividem-se em dois grupos: métodos baseados no cálculo (*Deterministic Optimization*) e os métodos aleatórios (*Random Strategies*).

No grupo de métodos aleatórios podem ser citados os Métodos de Ordem Zero, Algoritmos Genéticos, *Simulated Annealing*, Redes Neurais e Evolução Diferencial.

Já nos métodos determinísticos têm-se, por exemplo, o método de Multiplicador de Lagrange e os métodos de Penalidade.

Quando à presença de limitantes ao problema, tem-se a otimização irrestrita e a otimização restrita.

Na otimização irrestrita tem-se, por exemplo, o Método de Newton e o Método de Gradiente, enquanto que na otimização restrita existem os métodos indiretos por exemplo métodos Sequenciais e os métodos diretos como a Programação Linear (SARAMAGO; JÚNIOR, 2008).

Neste trabalho dar-se-á uma atenção especial aos problemas restritos, que no caso o Multiplicadores de Lagrange.

### 2.1 Definições e Alguns Fatos Básicos

Sejam os conjuntos  $D \subset \mathbb{R}^n$  e  $\Omega \subset \mathbb{R}^n$  tais que  $D \subset \Omega$ , e uma função  $f : \Omega \rightarrow \mathbb{R}^n$ . O problema principal a ser considerado é o de achar um minimizador (ou maximizador) de  $f$  no conjunto  $D$ .

Esse problema pode ser escrito como

$$\min f(x) \text{ sujeito a } x \in D. \quad (2.1)$$

O conjunto  $D$  será chamado conjunto viável do problema, os pontos do conjunto  $D$  serão chamados pontos viáveis, e  $f$  será chamada função objetivo.

**Definição 2.1.1.** (IZMAILOV; SOLODOV, 2007a) Dizemos que um ponto  $\bar{x} \in D$  é

- minimizador (ou maximizador) global de (2.1), se

$$f(\bar{x}) \leq f(x), \forall x \in D. \quad (2.2)$$

- minimizador (ou maximizador) local de (2.1), se existe uma vizinhança  $U$  de  $\bar{x}$  tal que

$$f(\bar{x}) \leq f(x), \forall x \in D \cap U. \quad (2.3)$$

Pela Definição 2.1.1, todo minimizador global também é um minimizador local, mas não reciprocamente. Se  $\forall x \neq \bar{x}$ , a desigualdade (2.2) ou (2.3) é restrita,  $\bar{x}$  será chamado minimizador restrito (global ou local, respectivamente).

## 2.2 Existência de Soluções Globais

Serão definidos alguns critérios que garantem a existência de solução global.

**Teorema 2.2.1. Teorema de Weierstrass** (IZMAILOV; SOLODOV, 2007a) Seja  $D \subset \mathbb{R}^n$  um conjunto compacto não - vazio e  $f : D \rightarrow \mathbb{R}$  uma função contínua.

Então, os problemas de minimizar e de maximizar  $f$  em  $D$  têm soluções globais.

**Definição 2.2.1.** (IZMAILOV; SOLODOV, 2007a) O conjunto de nível da função  $f : D \rightarrow \mathbb{R}$  associado a  $c \in \mathbb{R}$  é o conjunto dado por

$$L_{f,D}(c) = \{x \in D \mid f(x) \leq c\}.$$

**Corolário 2.2.1.** (IZMAILOV; SOLODOV, 2007a) Seja  $D \subset \mathbb{R}^n$  e  $f : D \rightarrow \mathbb{R}$  contínua no conjunto  $D$ . Suponha que exista  $c \in \mathbb{R}$  tal que o conjunto de nível  $L_{f,D}(c)$  seja não-vazio e compacto.

Então o problema de minimizar  $f$  em  $D$  possui uma solução global.

**Definição 2.2.2.** (IZMAILOV; SOLODOV, 2007a) Dados o conjunto  $D \subset \mathbb{R}^n$  e o ponto  $y \in \mathbb{R}^n$ , uma *projeção* (ortogonal) de  $y$  sobre  $D$  é uma solução global do problema

$$\min \|x - y\| \text{ sujeito a } x \in D.$$

**Corolário 2.2.2.** (IZMAILOV; SOLODOV, 2007a) Seja  $D \subset \mathbb{R}^n$  um conjunto fechado não-vazio. Então a projeção de  $y$  sobre  $D$  existe para todo ponto  $y \in \mathbb{R}^n$ .

## 2.3 Condições de Otimalidade

Quando  $D = \mathbb{R}^n$ , diz-se que o problema (2.1) é sem restrição (ou irrestrito), e quando  $D \neq \mathbb{R}^n$  falamos de otimização com restrição (ou restrito).

Apresentam-se, a seguir, as condições de otimalidade para o problema de minimização sem restrição

$$\min f(x), x \in \mathbb{R}^n. \quad (2.4)$$

Percebe-se que estes resultados são também verdadeiros para um problema com restrição (2.1), desde que o ponto de interesse  $\bar{x}$  esteja no interior do conjunto viável, isto é,  $\bar{x} \in \text{int } D$ .

**Teorema 2.3.1. (Condição de Otimalidade Sem Restrição)**(IZMAILOV; SOLODOV, 2007b)

- Supõe-se que a função  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  seja diferenciável no ponto  $\bar{x} \in \mathbb{R}^n$  e também que  $\bar{x}$  seja um minimizador local do problema (2.4).

Então

$$f'(\bar{x}) = 0. \quad (2.5)$$

Se  $f$  é duas vezes diferenciável em  $\bar{x}$ , então além de (2.5) tem-se que a Matriz Hessiano de  $f$  no ponto  $\bar{x}$  é semi-definida positiva, isto é,

$$^1 \langle f''(\bar{x})d, d \rangle \geq 0, \forall d \in \mathbb{R}^n. \quad (2.6)$$

Onde a condição necessária e suficiente para que uma matriz simétrica  $H$  seja **semi-definida positiva**:

- $x^T H x \geq 0$  para todos os vetores  $x$  (isso define uma semi-definida positiva).
- Todos os autovalores de  $H$  satisfazem  $\lambda \geq 0$ .
- Nenhuma submatriz principal possui determinantes negativos.

<sup>1</sup>  $\langle x, y \rangle$  - produto interno Euclidiano entre  $x \in \mathbb{R}^n$  e  $y \in \mathbb{R}^n$ . (IZMAILOV; SOLODOV, 2007a)

- Nenhum pivô é negativo.

- Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  seja duas vezes diferenciável no ponto  $\bar{x} \in \mathbb{R}^n$ .

Se  $\bar{x}$  satisfaz (2.5) e se a Matriz Hessiana de  $f$  em  $\bar{x}$  é definida positiva, isto é, se existe  $\gamma > 0$  tal que

$$\langle f''(\bar{x})d, d \rangle \geq \gamma \|d\|^2, \quad \forall d \in \mathbb{R}^n \quad (2.7)$$

então  $\bar{x}$  é minimizador local estrito do problema (2.4).

A condição (2.5) chama-se condição necessária de 1ª ordem para o problema (2.4), e os pontos que a satisfazem se chamam pontos críticos deste problema.

A combinação de (2.5) com (2.6) chama-se condição necessária de 2ª ordem, e a combinação (2.5) com (2.7) é a condição suficiente de 2ª ordem para o problema (2.4).

A seguir, apresentam-se as condições de otimalidade para o caso de restrição de igualdade:

$$\min f(x) \text{ sujeito a } x \in D = \{x \in \mathbb{R}^n \mid h(x) = 0\}, \quad (2.8)$$

onde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  e  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  são funções dadas. A função Lagrange ou Lagrangiana do problema (2.8) é dada por

$$L : \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R},$$

$$L(x, \lambda) = f(x) + \langle \lambda, h(x) \rangle = f(x) + \sum_{i=1}^l \lambda_i h_i(x).$$

Observa-se que

$$L'_x(x, \lambda) = f'(x) + (h'(x))^T \lambda = f'(x) + \sum_{i=1}^l \lambda_i h'_i(x),$$

$$L'_\lambda(x, \lambda) = h(x)$$

Condições deste tipo necessitam de hipóteses de regularidade das restrições. As condições de regularidade são:

$$\{h'_i(\bar{x}), i = 1, \dots, l\} \text{ é um conjunto linearmente independentes,} \quad (2.9)$$

ou  $h$  é uma função afim, isto é,

$$h(x) = Ax - a, \text{ onde } A \in \mathbb{R}(l, n), a \in \mathbb{R}^l \quad (2.10)$$

onde  $\mathbb{R}(l, n)$  é o espaço de matrizes reais de dimensão  $l \times n$ .

Sob qualquer uma das hipótese acima, tem-se como resultado o Teorema 2.3.2.



**Teorema 2.3.2. (Condição de Otimalidade no Caso com Restrição de Igualdade)**(IZMAILOV; SOLODOV, 2007b).

- Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  seja diferenciável no ponto  $\bar{x} \in \mathbb{R}^n$  e que  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  seja diferenciável numa vizinhança deste ponto, com derivada contínua em  $\bar{x}$  seja minimizador local do problema (2.8).

Se vale uma das condições de regularidade das restrições ((2.9) ou (2.10)) então existe  $\bar{\lambda} \in \mathbb{R}^l$  tal que

$$L'_x(\bar{x}, \bar{\lambda}) = 0. \quad (2.11)$$

Sob a condição (2.9),  $\bar{\lambda} \in \mathbb{R}^l$  que satisfazendo (2.11) é único.

Se  $f$  e  $h$  são duas vezes diferenciáveis no ponto  $\bar{x}$ , então para qualquer  $\bar{\lambda} \in \mathbb{R}^l$  que satisfaz (2.11) temos o seguinte:

$$\langle L''_{xx}(\bar{x}, \bar{\lambda})d, d \rangle \geq 0, \quad \forall d \in \ker h'(\bar{x})^2.$$

O sistema de equações

$$L'_x(x, \lambda) = 0, \quad h(x) = 0$$

em relação a  $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$ , que caracteriza os pontos estacionários do problema (2.8) e os Multiplicadores de Lagrange associados, chamam-se sistemas de Lagrange. De forma equivalente, este sistema pode ser escrito como

$$L'(x, \lambda) = 0,$$

onde a derivada é em relação a todas as variáveis da Lagrangiana. O número de variáveis neste sistema é igual ao número de equações (IZMAILOV; SOLODOV, 2007b).

Na subseção 2.4 descreve-se o método do Multiplicadores de Lagrange. Metodologia empregada neste trabalho que permite solucionar problemas de otimização restrita (B).

## 2.4 Multiplicadores de Lagrange, Máximos e Mínimos Condicionados

Problemas de otimização restrita podem ser muito complexos, não havendo um método geral para determinar a solução de todas as classes de problemas. Em algumas situações simples, é possível obter um resultado explicitando uma variável em função das outras na restrição; substituindo este resultado na função objetivo e resolvendo o problema de otimização irrestrita (ou não restrita) resultante.

<sup>2</sup>  $\ker A$  - o núcleo do operador linear  $A$  (ou da matriz  $A$ ) (IZMAILOV; SOLODOV, 2007b)

Entre os métodos que permitem analisar situações mais gerais, está o método dos Multiplicadores de Lagrange <sup>3</sup>. Por meio desse método, um problema de otimização restrita com  $n$  variáveis e  $m$  variáveis de igualdade é transformado em um problema de otimização não restrita com  $(n + m)$  variáveis (GONÇALVES; FLEMMING, 2007).

Em outras palavras, este método converte problemas de otimização com restrição em problemas sem restrição, através da inserção de um novo parâmetro: o Multiplicador de Lagrange -  $\lambda$ .

Resumidamente, um problema de otimização com restrição pode ser escrito da seguinte forma: minimize (ou maximize) uma função real submetida a uma função limitadora (ou restritiva)

**Teorema 2.4.1. Método dos Multiplicadores de Lagrange** (MUNEM; FOULIS, 1982)

Suponha que  $f$  e  $g$  sejam funções definidas e tenham derivadas parciais contínuas num subconjunto  $D$  consistindo inteiramente em pontos interiores. Suponha que, em cada ponto  $(x, y, z)$  em  $D$ , pelo menos uma das três derivadas parciais  $g_1(x, y, z)$ ,  $g_2(x, y, z)$ ,  $g_3(x, y, z)$  seja diferente de zero. Então os pontos  $(x, y, z)$  em  $D$ , nos quais  $f$  tem extremos relativos, sujeito à restrição

$$g(x, y, z) = k,$$

onde  $k$  é uma constante, podem ser determinados como segue.

Seja a função  $L$  definida por

$$L(x, y, z) = f(x, y, z) + \lambda g(x, y, z)$$

para  $(x, y, z)$  em  $D$  e  $\lambda$  (chamado de Multiplicador de Lagrange) uma constante a ser determinada. Então, resolvendo o sistema de equações

$$\begin{aligned}\frac{\partial L}{\partial x} &= 0 \\ \frac{\partial L}{\partial y} &= 0 \\ \frac{\partial L}{\partial z} &= 0 \\ \frac{\partial L}{\partial \lambda} &= k\end{aligned}$$

para  $x, y, z$  e  $\lambda$ , diversas soluções podem ser obtidas. Os pontos  $(x, y, z)$  desejados, onde  $f$  tem seus extremos sujeitos à restrição, estão entre essas soluções.

<sup>3</sup> Joseph Louis Lagrange (1736-1813)

### 2.4.1 Problemas Envolvendo Funções a Três Variáveis e Duas Restrições

Suponha agora que se quer determinar os valores máximos e mínimos de  $f(x, y, z)$  sujeita a duas restrições (condições laterais) da forma  $g_1(x, y, z) = k$  e  $g_2(x, y, z) = c$ . Geometricamente, isso significa a procura pelos valores extremos de  $f$  quando  $(x, y, z)$  está restrito a pertencer à curva  $C$ , obtida pela intersecção das superfícies de níveis  $g_1(x, y, z) = k$  e  $g_2(x, y, z) = c$  (Figura 5) (STEWART, 2006).

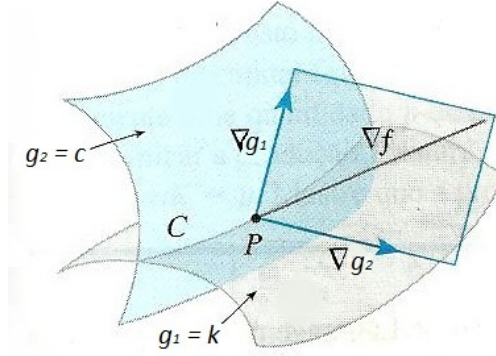


Figura 5 – Máximos e Mínimos da Intersecção  $g_1$  e  $g_2$ .

Fonte: (STEWART, 2006)

Supõe-se que  $f$  tenha valor extremo no ponto  $P(x_0, y_0, z_0)$ . Sabe-se que  $\nabla f$  é ortogonal a  $C$ . Além disso,  $\nabla g_1$  é ortogonal a  $g_1(x, y, z) = k$  e  $\nabla g_2$  é ortogonal a  $g_2(x, y, z) = c$ . Portanto,  $\nabla g_1$  e  $\nabla g_2$  são ambos ortogonais a  $C$ . Isso significa que o vetor gradiente  $\nabla f(x_0, y_0, z_0)$  pertence ao plano determinado por  $\nabla g_1(x_0, y_0, z_0)$  e  $\nabla g_2(x_0, y_0, z_0)$ .

Supõe-se também que esses vetores gradientes não são paralelos e nem nulos. Logo, existem números  $\lambda$  e  $\mu$  (chamados Multiplicadores de Lagrange), tais que

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g_1(x_0, y_0, z_0) + \mu \nabla g_2(x_0, y_0, z_0). \quad (2.12)$$

Neste caso o Método de Lagrange nos leva a procurar os valores extremos resolvendo as cinco equações nas cinco incógnitas  $x, y, z, \lambda$  e  $\mu$ . Tais equações podem ser obtidas escrevendo-se a equação (2.12) como

$$f_x = \lambda g_{1x} + \mu g_{2x}$$

$$f_y = \lambda g_{1y} + \mu g_{2y}$$

$$f_z = \lambda g_{1z} + \mu g_{2z}$$

$$g_1(x, y, z) = k$$

$$g_2(x, y, z) = c.$$

## 2.5 O Hessiano Aumentado (Orlado)

Tal como no caso do extremo livre não condicionado, é possível expressar a condição de segunda ordem na forma de um determinante para situações envolvendo extremos condicionados.

No lugar do determinante Hessiano  $|H|$ , encontrado para o caso do extremo condicionado, usa-se o chamado Hessiano Aumentado (ou Hessiano Orlado) (CHIANG, 1982).

O desenvolvimento para se chegar na matriz Hessiano Orlado a partir da matriz Hessiano, se encontra no livro Matemática para Economistas no capítulo 12 de (CHIANG, 1982), o foco deste trabalho é utilizar o método de Hessiano Orlado.

De forma simples, o discriminante consiste no Hessiano  $|H| = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix}$ .

O determinante denominado como Hessiano Aumentado ou Hessiano Orlado, denota-se por  $|\bar{H}|$ , em que a barra na parte superior simboliza o aumento (ou orla). Dado um valor crítico ou estacionário de  $z = f(x, y)$  ou de  $L = f(x, y) + \lambda[c - g(x, y)]$  em que  $L$  é a função Lagrangiana, um  $|\bar{H}|$  positivo é suficiente para estabelecer o ponto crítico como um máximo relativo da função objetiva,  $f$ . Analogamente um  $|\bar{H}|$  negativo é suficiente para estabelecer o ponto crítico como um mínimo da função objetiva. Todas as derivadas envolvidas em  $|\bar{H}|$  são calculadas nos valores críticos de  $x$  e  $y$  (CHIANG, 1982).

De forma resumida, **o ponto local ou relativo é** (SILVA; MACHADO, 2010):

- Um mínimo local ou relativo da função  $f$  se  $|\bar{H}| < 0$ .
- Um máximo local ou relativo da função  $f$  se  $|\bar{H}| > 0$ .

### 2.5.1 O Caso de $n$ -variáveis (CHIANG, 1982).

Quando a função objetivo toma a forma

$$z = f(x_1, x_2, x_3, \dots, x_n).$$

Sujeita a

$$g(x_1, x_2, x_3, \dots, x_n) = c,$$

as condições para que função lagrangiana seja positiva definida (ou negativa definida) envolvem, uma vez mais, o Hessiano Aumentado (ou Orlado). Dessa forma, é preciso estabelecer condições necessárias para serem expressas em termos dos menores principais líderes aumentados do Hessiano.

Onde a condição necessária e suficiente para que a matriz simétrica real  $H$  seja **positiva definida**:

- $x^T H x > 0$  para todos os vetores reais  $x$  não nulos.
- Todos os autovalores de  $H$  sejam positivos.
- Todas as submatrizes superiores  $H_k$  à esquerda têm determinantes positivos.
- Todos os pivôs (sem alterações de linhas) são positivas.

Dado o Hessiano Aumentado (ou Orlado)

$$|\bar{H}| = \begin{vmatrix} 0 & g_1 & g_2 & \dots & g_n \\ g_1 & L_{11} & L_{12} & \dots & L_{1n} \\ g_2 & L_{21} & L_{22} & \dots & L_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ g_n & L_{n1} & L_{n2} & \dots & L_{nn} \end{vmatrix},$$

os menores principais aumentados podem ser definidos como

$$|\bar{H}_2| = \begin{vmatrix} 0 & g_x & g_y \\ g_x & L_{11} & L_{12} \\ g_y & L_{21} & L_{22} \end{vmatrix} \quad (2.13)$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & g_1 & g_2 & g_3 \\ g_1 & L_{11} & L_{12} & L_{13} \\ g_2 & L_{21} & L_{22} & L_{23} \\ g_3 & L_{31} & L_{32} & L_{33} \end{vmatrix} \dots, \quad (2.14)$$

sendo que o último é  $|\bar{H}_n| = |\bar{H}|$ . Nesta notação a barra horizontal acima de  $H$  novamente significa o aumento, ou melhor aumentado. Por exemplo,  $|\bar{H}_2|$  envolve o segundo menor principal líder do hessiano (simples), aumentado com 0,  $g_1$  e  $g_2$ ; e analogamente para os demais (CHIANG, 1982).

As condições para definir positiva (mínimo local) ou negativa (máximo local) do Hessiano Orlado são:

**O ponto crítico (ou extremo) é:**

- Um mínimo local da função  $f$  se  $|\bar{H}_2|, |\bar{H}_3|, \dots, |\bar{H}_n| < 0$ .
- Um máximo local da função  $f$  se  $|\bar{H}_2| < 0; |\bar{H}_3| > 0; |\bar{H}_4| < 0; \dots$

- Se nenhuma dessas condições se verifica, a função deve ser examinada na vizinhança no ponto crítico. (SILVA; MACHADO, 2010)

Na Matriz (2.13), para todos os menores principais líderes aumentados (Orlados), necessitam ser negativos; e na matriz (2.14), os sinais devem se alternar (CHIANG, 1982).

### 2.5.2 Caso de Múltiplas Restrições

Quando aparece mais de uma restrição no problema, a condição de segunda ordem envolve um Hessiano Aumentado (Orlado) por mais de uma linha e coluna.

Supõe-se que há  $n$  variáveis de escolha e  $m$  restrições ( $m < n$ ) da forma  $g^j(x_1, \dots, x_n) = c_j$ . Então, a função de Lagrange será

$$L = f(x_1, \dots, x_n) + \sum_{j=1}^m \lambda_j [c_j - g^j(x_1, \dots, x_n)]$$

e o Hessiano Aumentado aparecerá como

$$|\bar{H}| = \begin{vmatrix} 0 & 0 & \dots & 0 & | & g_1^1 & g_2^1 & \dots & g_n^1 \\ 0 & 0 & \dots & 0 & | & g_1^2 & g_2^2 & \dots & g_n^2 \\ \dots & \dots & \dots & \dots & | & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & | & g_1^m & g_2^m & \dots & g_n^m \\ - & - & - & - & | & - & - & - & - \\ g_1^1 & g_1^2 & \dots & g_1^m & | & L_{11} & L_{12} & \dots & L_{1n} \\ g_2^1 & g_2^2 & \dots & g_2^m & | & L_{21} & L_{22} & \dots & L_{2n} \\ \dots & \dots & \dots & \dots & | & \dots & \dots & \dots & \dots \\ g_n^1 & g_n^2 & \dots & g_n^m & | & L_{n1} & L_{n2} & \dots & L_{nn} \end{vmatrix}$$

onde  $g_i^j \equiv \frac{\partial g^j}{\partial x_i}$ ,  $i = (1, \dots, n)$  e  $j = (1, \dots, m)$ , são as derivadas parciais das funções de restrição e o símbolo  $L$  com dois índices denota a derivada parcial de segunda ordem da função de Lagrange. Note que o Hessiano Orlado foi subdividido em quatro *blocos* para uma visualização mais clara. O bloco superior esquerdo é composto somente de zeros, e o bloco inferior direito é apenas o Hessiano simples. Os outros dois blocos, contendo as derivadas  $g_i^j$ , guardam uma relação de imagem especular entre si com referência à diagonal principal, resultando, assim, em um arranjo simétrico de elementos em todo o Hessiano Aumentado.

Vários menores principais líderes aumentados podem ser formados a partir de  $|\bar{H}|$ . O que contiver  $L_{22}$  como último elemento de sua diagonal principal será denotado por  $|\bar{H}_2|$ . Incluindo mais uma linha e mais coluna, de modo que  $L_{33}$  entra em cena, têm-se  $|\bar{H}_3|$  e assim por diante. Com essa simbologia, pode-se enunciar a condição suficiente de segunda ordem em termos dos sinais dos seguintes  $(n - m)$  menores principais líderes

aumentados:

$$|\bar{H}_{m+1}|, |\bar{H}_{m+2}|, \dots, |\bar{H}_n| \text{ (ou } |\bar{H}|).$$

Para um máximo de  $z$  uma condição suficiente é que esses menores principais líderes aumentados alternem sinais, sendo que o sinal de  $|\bar{H}_{m+1}|$  é o de  $(-1)^{m+1}$ . Para um mínimo de  $z$ , uma condição suficiente é que todos esses menores principais tenham o mesmo sinal, a saber, o de  $(-1)^m$ .

Note que ter um número par ou um número ímpar de restrições faz uma grande diferença, porque  $(-1)$  elevado a uma potência ímpar resultará no sinal oposto ao de  $(-1)$  elevado a uma potência par (CHIANG, 1982).

No capítulo seguinte é descrito o algoritmo de Otimização: Algoritmo Evolutivo e Enxame de Partículas.

## 3 Algoritmos de Otimização

Neste capítulo é descrito o algoritmo evolutivo para otimização chamado Enxame de Partículas (PEDROSA; AFONSO, 2011). A escolha deste método, deve-se ao fato de que os resultados obtidos pela metodologia proposta nesse trabalho serão comparados com dados disponíveis na literatura resultantes da implementação deste algoritmo (PEDROSA; AFONSO, 2011).

### 3.1 Algoritmos Evolutivos

Problemas de otimização modelados através de uma função matemática podem ser resolvidos através de métodos clássicos de otimização, como a programação linear e inteira, onde o objetivo é localizar o valor ótimo para a função que representa o problema. Embora esta técnica seja muito utilizada, o processo de localização do ótimo pode ser computacionalmente custoso. Como alternativa, podem ser utilizados os Algoritmos Evolutivos, que não garantem uma solução ótima, mas podem chegar a um resultado sub-ótimo satisfatório com um custo computacional aceitável.

Os Algoritmos Evolutivos têm sido exaustivamente investigados neste contexto devido à sua capacidade de resolver problemas de otimização complexos, trabalhando de forma eficiente na avaliação das variações específicas dos problemas estudados. Dentre os algoritmos evolutivos mais utilizados estão os: Algoritmos Genéticos, Algoritmos Meméticos, os Sistemas Imunológicos Artificiais e as Colônias de Formigas e, por fim, o Enxame de Partículas (PSO) (SILVA; LEMONGE; LIMA, 2014).

#### 3.1.1 *Particle Swarm Optimization* - PSO (Enxame de Partículas)

PSO é uma técnica que se baseia no movimento coletivo de um grupo de partículas, o enxame de partícula. Cada membro desse enxame é movimentado através do espaço de busca do problema por duas forças. Uma força os atrai com uma magnitude aleatória para a melhor localização já encontrada por ele próprio (*pbest*) e a outra força os atrai para a melhor localização encontrada entre alguns ou todos os membros do enxame (*gbest*). A posição e a velocidade de cada partícula são atualizadas a cada iteração até todo o enxame convergir.

O algoritmo (Figura 6) inicia cada partícula com valores aleatórios de posição e velocidade. Durante a sua execução cada partícula avaliará sua posição atual (ou solução atual) em relação à melhor posição já encontrada por ela mesma, fazendo com que o valor de *pbest* seja atualizado. Cada partícula também avaliará a qualidade da melhor



solução encontrada na sua vizinhança, permitindo atualização do valor  $gbest$ . A avaliação da qualidade de uma posição (ou solução) é realizada por meio da função de *fitness* ou aptidão. Após atualizar o valor de velocidade com os novos valores de  $pbest$  e  $gbest$ , cada partícula irá se deslocar para uma nova posição (NASCIMENTO et al., 2012).

O algoritmo PSO não trabalha diretamente com restrições. Uma estratégia para que esse algoritmo manipule restrições, é utilizar as funções de penalidade. Assim, a função objetivo penalidade  $f_p(x)$  é obtida através da modificação da função  $f(x)$  da seguinte forma:

$$f_p(x) = f(x) - vp.VT,$$

em que:

$f_p(x)$  = valor da função-objetivo penalidade para solução  $x$ ;

$f(x)$  = valor da função-objetivo do problema;

$vp$  = penalização para cada unidade de produção violada;

$VT$  = violação total das restrições de produção (mínima e máxima);

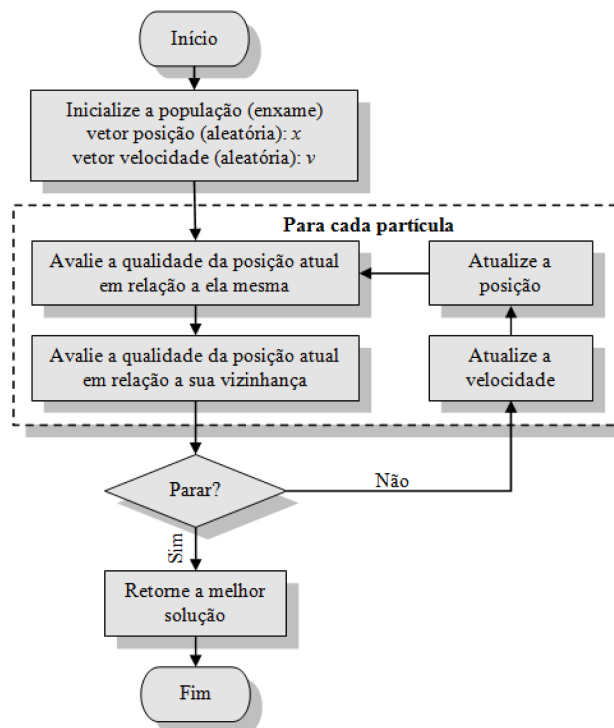


Figura 6 – Fluxograma PSO.

Fonte: (NASCIMENTO et al., 2012)

No capítulo seguinte são apresentados os resultados e as discussões pertinentes.

## 4 Resultados Numéricos

Neste capítulo, são apresentados os resultados obtidos após a implementação da metodologia proposta, tendo como premissa as formulações descritas nas seções 1.4 e 2.4.

A teoria dos Máximos e Mínimos condicionados com o emprego do Método dos Multiplicadores de Lagrange é aplicada na determinação do ponto ótimo global para problemas sujeitos a uma restrição de igualdade. Inicialmente, realizaram-se testes para a função do tipo G11. Em um segundo momento, fez-se um estudo utilizando os mesmos procedimentos para identificar os pontos ótimos para a função chamada de G3. Os pontos obtido, que satisfazem os problemas, foram representados graficamente no software GeoGebra e, finalmente verificou-se, pelo emprego do método de Hessiano Orlado, se tais pontos são, de fato, os pontos ótimos.

### 4.1 Grupo de Funções G

O grupo de funções G, do inglês *G-suite*, citados (KOZIEL; MICHALEWICZ, 1999), tem sido utilizado de forma a validar e comparar o desempenho de algoritmos de otimização. Optou-se por utilizar este grupo para que fosse possível comparar com outros métodos de otimização de problemas com restrições.

O grupo de funções G é composto por 24 funções. Neste trabalho optou-se por utilizar as funções denominada G11 e G3 para validação da metodologia proposta. Os dados para as funções G11 e G3 são comparados com os resultados obtido por (PEDROSA; AFONSO, 2011), onde foi utilizado o algoritmo Enxame de partículas (*Particle Swarm Optimization* – PSO) e por (KOZIEL; MICHALEWICZ, 1999), onde foi utilizado um Algoritmo Evolutivo modificado, respectivamente.

### 4.2 Descrição do Problema 1 - Função G11

Nesta seção são apresentados os resultados numéricos para a função G11. A função G11 é representada por,

$$f(x, y) = x^2 + (y - 1)^2 \quad (4.1)$$

e sujeita à restrição

$$y - x^2 = 0.$$

### 4.2.1 Desenvolvimento do Problema

Empregando-se o método dos Multiplicadores de Lagrange obtém-se,

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y)$$

$$f(x, y) = x^2 + (y - 1)^2$$

$$g(x, y) = y - x^2$$

$$L(x, y, \lambda) = x^2 + (y - 1)^2 - \lambda(y - x^2).$$

Aplicando as derivadas de 1ª ordem tem-se que,

$$\frac{\partial L}{\partial x} = 2x + 2\lambda x = 0 \quad (4.2)$$

$$\frac{\partial L}{\partial y} = 2y - 2 - \lambda = 0 \quad (4.3)$$

$$\frac{\partial L}{\partial \lambda} = y - x^2 = 0. \quad (4.4)$$

Colocando-se em evidência  $2x$  na equação (4.2), tem-se que:

$$2x(1 + \lambda) = 0.$$

Logo, obtêm-se duas condições:  $x = 0$  e  $\lambda = -1$ .

1º **caso**,  $x = 0$ :

Substituindo-se o valor encontrado para  $x$  na equação (4.4), tem-se

$$y - (0)^2 = 0 \rightarrow y = 0.$$

Uma vez obtido o valor de  $y$ , determina-se o valor de  $\lambda$ . Logo, substituindo-se  $y = 0$  na equação (4.3), tem-se que

$$2(0) - 2 - \lambda = 0 \rightarrow \lambda = -2.$$

Neste caso, o ponto crítico encontrado para o 1º caso,  $x = 0$ , é o ponto  $(0, 0, 1)$  e o Multiplicador de Lagrange,  $\lambda = -2$ .

Com o valor do par de pontos  $(x, y)$  pode-se determinar a coordenada  $z$  a partir da função  $f(x, y)$ , isto é, a função objetivo  $f$  sujeita a restrição  $g$  do problema G11. Desta forma, com o resultado do ponto  $(x, y, z)$ , e utilizando-se o software GeoGebra, é possível obter a representação deste no gráfico (Figura 7).

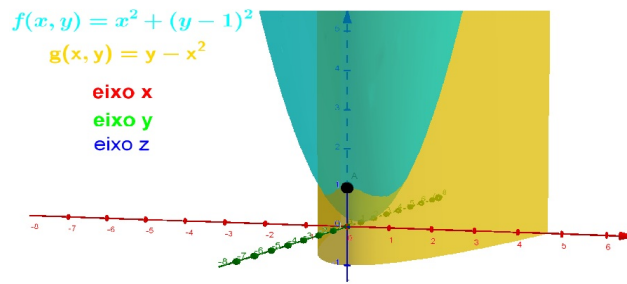


Figura 7 – Representação do ponto  $(0, 0, 1)$  para o 1º caso,  $x = 0$ , da função G11.

Observa-se pelo gráfico (Figura 7) que o ponto  $(0, 0, 1)$  não pode ser um ponto de máximo ao decorrer da intersecção das duas funções. Isto será analisado, posteriormente, com o emprego do método do Hessiano Orlado.

2º caso:  $\lambda = -1$ .

Substituindo-se  $\lambda = -1$  na equação (4.3), obtém-se,

$$2y - 2 - (-1) = 0 \rightarrow y = \frac{1}{2}.$$

Uma vez obtido o novo valor de  $y$ , substitui-se na equação (4.4). Logo,

$$\left(\frac{1}{2}\right) - x^2 = 0 \rightarrow x = \pm\sqrt{\frac{1}{2}}.$$

Neste caso, o ponto crítico encontrado para o 2º caso são, na verdade, dois pontos  $\left(\sqrt{\frac{1}{2}}, \frac{1}{2}, \frac{3}{4}\right)$  e  $\left(-\sqrt{\frac{1}{2}}, \frac{1}{2}, \frac{3}{4}\right)$ , o multiplicador de Lagrange  $\lambda = -1$ .

Com os novos pares de pontos de  $(x, y)$  pode-se determinar a coordenada  $z$  a partir da função  $f(x, y)$ .

É importante ressaltar que deve-se analisar o caso para  $x > 0$  e  $x < 0$ . A partir desta análise, obtém-se o gráfico pelo *software* GeoGebra. Uma vez obtido o resultado para as coordenadas  $(x, y, z)$ , representam-se os pontos no gráfico (Figura 8).

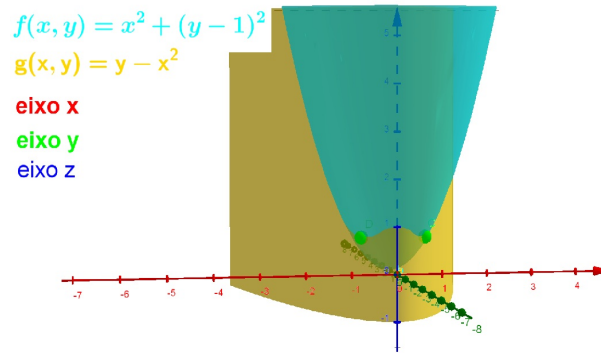


Figura 8 – Representação do ponto Crítico para o 2º caso da função G11.

Observa-se pelo gráfico (Figura 8), ao decorrer da intersecção das duas funções, que os pontos determinados são pontos de mínimo.

Para verificar se os pontos obtidos para os dois casos, são de fato pontos de máximo ou de mínimo ou nenhum dos dois, aplica-se o método do Hessiano Orlado (Aumentado).

Relembrando que o ponto  $(x, y, z)$  será:

- Um mínimo local da função  $f$  se  $|\bar{H}_2|, |\bar{H}_3|, \dots, |\bar{H}_n| < 0$ .
- Um máximo local da função  $f$  se  $|\bar{H}_2| < 0; |\bar{H}_3| > 0; |\bar{H}_4| < 0; \dots$
- Se nenhuma dessas condições se verifica, a função deve ser examinada na vizinhança do ponto crítico.

Esse método será aplicado, como segue, nos 1º e 2º casos decorrentes da aplicação dos Multiplicadores de Lagrange na resolução do problema G11.

Para a aplicação do método, são necessárias as derivadas de 1ª ordem da função  $g$ , e também das derivadas de 2ª ordem da função  $L$ .

Calculando-se as derivadas necessárias, tem-se:

$$\frac{\partial g}{\partial x} = -2x;$$

$$\frac{\partial g}{\partial y} = 1;$$

$$\frac{\partial^2 L}{\partial x^2} = 2 - 2\lambda;$$

$$\frac{\partial^2 L}{\partial y^2} = 2;$$

$$\frac{\partial^2 L}{\partial x \partial y} = 0;$$

$$\frac{\partial^2 L}{\partial y \partial x} = 0.$$

Substituindo-se os valores encontrados para os pontos críticos do 1º caso,  $(0, 0, 1)$  e o valor do Multiplicador de Lagrange ( $\lambda = -2$ ), tem-se:

$$g_x = 0; g_y = 1; L_{xx} = 6; L_{yy} = 2; L_{xy} = 0; L_{yx} = 0.$$

Sendo assim tem-se:

$$|\bar{H}_2| = \begin{vmatrix} 0 & g_x \\ g_x & L_{xx} \end{vmatrix}$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{yx} & L_{yy} \end{vmatrix}.$$

Substituindo-se os valores nos determinantes do Hessiano Orlado, obtêm-se;

$$|\bar{H}_2| = \begin{vmatrix} 0 & 0 \\ 0 & 6 \end{vmatrix} = 0 - 0 = 0$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 6 & 0 \\ 1 & 0 & 2 \end{vmatrix} = 0 + 0 + 0 - 6 - 0 - 0 = -6.$$

Como, o valor obtido, não corresponde aos requisitos de ponto máximo ou de ponto mínimo, segundo a definição do método do Hessiano Orlado, com  $|\bar{H}_2| = 0$  e  $|\bar{H}_3| < 0$ , nada pode-se concluir sobre o 1º caso.

De fato, pelo gráfico da Figura 7, o ponto  $(0, 0, 1)$ , não pode ser um ponto máximo extremo, como já se foi observado.

Para o 2º caso, aplica-se novamente o processo.

Os valores encontrados para o 2º caso são:

$$\left( \sqrt{\frac{1}{2}}, \frac{1}{2}, \frac{3}{4} \right), \left( -\sqrt{\frac{1}{2}}, \frac{1}{2}, \frac{3}{4} \right) \text{ e } \lambda = -1.$$

Primeiramente, avalia-se o caso em que  $x = +\sqrt{\frac{1}{2}}$  e obtém-se:

$$g_x = -2x; g_y = 1; L_{xx} = 2 - 2\lambda; L_{yy} = 2; L_{xy} = 0; L_{yx} = 0.$$

Substituindo os valores tem-se

$$g_x = -2\sqrt{\frac{1}{2}}; g_y = 1; L_{xx} = 4; L_{yy} = 2; L_{xy} = 0; L_{yx} = 0.$$

Aplicando-se o método do Hessiano Orlado chega-se a

$$|\bar{H}_2| = \begin{vmatrix} 0 & -2\sqrt{\frac{1}{2}} \\ -2\sqrt{\frac{1}{2}} & 4 \end{vmatrix} = 0 - \left(-2\sqrt{\frac{1}{2}}\right) \left(-2\sqrt{\frac{1}{2}}\right) = -2$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & -2\sqrt{\frac{1}{2}} & 1 \\ -2\sqrt{\frac{1}{2}} & 4 & 0 \\ 1 & 0 & 2 \end{vmatrix} = -4 - 0 - 2 \left(-2\sqrt{\frac{1}{2}}\right) \left(-2\sqrt{\frac{1}{2}}\right) = -8.$$

Logo, tem-se que  $|\bar{H}_2| < 0$  e  $|\bar{H}_3| < 0$ , isto é, o ponto  $\left(\sqrt{\frac{1}{2}}, \frac{1}{2}, \frac{3}{4}\right)$  do 2º caso, para situação de  $x$  positivo, é um ponto extremo mínimo.

Considera-se agora a situação para  $x = -\sqrt{\frac{1}{2}}$ . Logo,

$$g_x = -2x; g_y = 1; L_{xx} = 2 - 2\lambda; L_{yy} = 2; L_{xy} = 0; L_{yx} = 0.$$

Substituindo-se os valores tem-se

$$g_x = -2(-\sqrt{\frac{1}{2}}); g_y = 1; L_{xx} = 4; L_{yy} = 2; L_{xy} = 0; L_{yx} = 0$$

Simplificando-se os resultados chega-se à

$$|\bar{H}_2| = \begin{vmatrix} 0 & 2\sqrt{\frac{1}{2}} \\ 2\sqrt{\frac{1}{2}} & 4 \end{vmatrix} = 0 - \left(2\sqrt{\frac{1}{2}}\right) \left(2\sqrt{\frac{1}{2}}\right) = -2$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & 2\sqrt{\frac{1}{2}} & 1 \\ 2\sqrt{\frac{1}{2}} & 4 & 0 \\ 1 & 0 & 2 \end{vmatrix} = -4 - 0 - 4 = -8.$$

Logo, tem-se que  $|\bar{H}_2| < 0$  e  $|\bar{H}_3| < 0$ , isto é, o ponto do 2º caso, de fato, é um ponto de mínimo local.

Com a aplicação do método do Hessiano Orlado, é possível comprovar que o ponto encontrado é de fato um ponto extremo de máximo (ou mínimo). Para situações de funções com mais de 2 variáveis, não é possível a visualização gráfica, entretanto esta verificação ainda pode ser feita pelo emprego deste método.

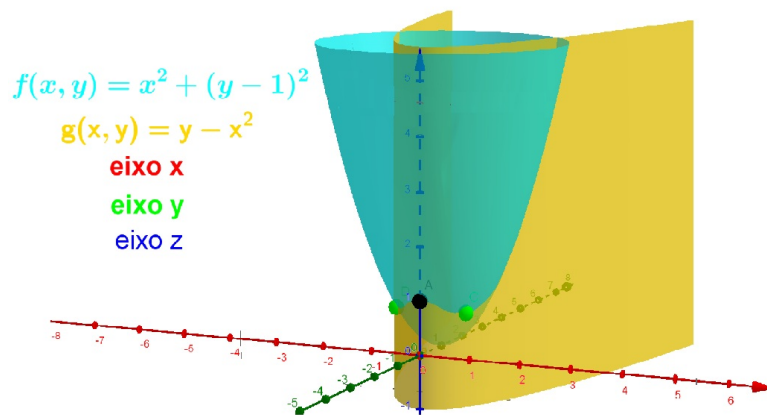


Figura 9 – Pontos de máximos e mínimos para a função G11.

A partir da aplicação do método dos Multiplicadores de Lagrange foi possível definir os pontos que solucionam o problema. A representação gráfica, obtida com o software GeoGebra, auxilia na identificação dos pontos candidatos a máximo e/ou mínimo (pontos críticos) e com o emprego do método do **Hessiano Orlado** verifica-se e classifica-se quem são, de fato, os pontos de máximos e/ou mínimos extremos.

Na Tabela 1 são apresentados os resultados obtidos a partir da aplicação do Método dos Multiplicadores de Lagrange, proposto neste trabalho, os dados gerados por Pedrosa e Afonso (PEDROSA; AFONSO, 2011) e por Koziel (KOZIEL; MICHALEWICZ, 1999). Observa-se uma excelente concordância com os resultados apresentados por Koziel e Michalewicz (KOZIEL; MICHALEWICZ, 1999).

Tabela 1 – Comparação dos Resultados

|                                            | Teste 1 | Teste 2 | Teste 3 |
|--------------------------------------------|---------|---------|---------|
| <b>Koziel e Michalewicz</b>                | 0.75    | 0.75    | 0.75    |
| <b>Pedrosa e Afonso</b>                    | 0.7487  | 0.7355  | 0.7525  |
| <b>Método do Multiplicador de Lagrange</b> | 0.75    | -       | -       |

Ao confrontar os resultados, pode-se notar que a metodologia proposta neste trabalho possibilitou uma melhor aproximação que o método proposto no trabalho de Pedrosa e Afonso (PEDROSA; AFONSO, 2011). Além disso, por se tratar de um método determinístico, não exige um grande esforço computacional.



### 4.3 Descrição do Problema 2 - Função G3

O segundo problema abordado neste trabalho considera - se a função G3 representada por

$$f(x_1, \dots, x_n) = (\sqrt{n})^n \prod_{i=1}^n x_i \quad (4.5)$$

e sujeita a restrição

$$\sum_{i=1}^n x_i^2 = 1.$$

onde  $0 \leq x_i \leq 1$ ;  $i = 1, 2, \dots, n$

Neste problema foi restringido para  $i = 1, 2$ , logo o problema reescrito é

$$f(x_1, x_2) = (\sqrt{2})^2 x_1 x_2 \quad (4.6)$$

e sujeita a restrição

$$g(x_1, x_2) = x_1^2 + x_2^2 = 1.$$

O problema G3 foi resolvido de maneira análoga à sub-seção anterior. No entanto, os resultados foram apresentados de maneira simplificado.

#### 4.3.1 Desenvolvimento do Problema

Empregando-se o método dos Multiplicadores de Lagrange obtêm-se

$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda g(x_1, x_2)$$

$$f(x_1, x_2) = (\sqrt{2})^2 x_1 x_2$$

$$g(x_1, x_2) = x_1^2 + x_2^2 - 1$$

$$L(x_1, x_2, \lambda) = (\sqrt{2})^2 x_1 x_2 - \lambda(x_1^2 + x_2^2 - 1).$$

Aplicando as derivadas de 1ª ordem tem-se

$$\frac{\partial L}{\partial x_1} = 2x_2 - 2x_1\lambda = 0 \quad (4.7)$$

$$\frac{\partial L}{\partial x_2} = 2x_1 - 2x_2\lambda = 0 \quad (4.8)$$

$$\frac{\partial L}{\partial \lambda} = x_1^2 + x_2^2 - 1 = 0. \quad (4.9)$$

Isolando  $x_1$  na equação 4.9 tem-se

$$x_1 = \pm \sqrt{1 - \frac{1}{\lambda^2 + 1}}.$$

Substituindo-se o valor encontrado para  $x_1$  na equação 4.8 e isola-se  $x_2$ , chega-se a

$$x_2 = \pm \sqrt{\frac{1}{\lambda^2 + 1}}.$$

Substituindo-se o valor encontrado para  $x_1$  e  $x_2$  na equação 4.7, isola-se  $\lambda$ , obtendo-se

$$\lambda = \pm 1.$$

Os valores encontrados para  $\lambda = 1$  e  $\lambda = -1$ , onde nas duas situações de  $\lambda$ , substituindo-se encontra-se o mesmos valores para  $x_1$  e  $x_2$ , que são  $x_1 = \pm \sqrt{\frac{1}{2}}$  e  $x_2 = \pm \sqrt{\frac{1}{2}}$ .

Neste caso, os pontos críticos obtidos são  $(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}}, 1)$ ,  $(-\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}}, 1)$ ,  $(-\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}}, -1)$  e  $(\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}}, -1)$ .

Com os pontos de  $(x_1, x_2)$  pode-se encontrar a coordenada  $z$  a partir da função  $f(x_1, x_2)$ , isto é, a função objetivo  $f$  sujeita à restrição  $g$  do problema G3, desta forma obtém-se o gráfico pelo software **GeoGebra**, com o resultado das coordenadas  $(x_1, x_2, z)$  e se obtém os 4 pontos representados em azul na (Figura 10).

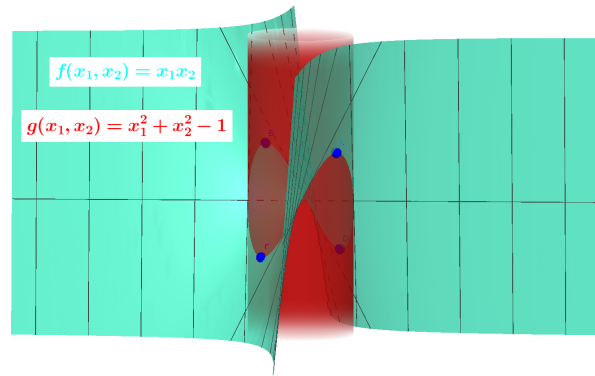


Figura 10 – Pontos de máximo e mínimos para G3.

Para verificar se os quatros pontos encontrados são de fato pontos de máximo ou de mínimo ou não nenhum dos dois, aplica-se o método do Hessiano Orlado (Aumentado).

Sendo assim, optou-se por dividir esse problema em 2 casos, 1º para  $\lambda = 1$  com 4 sub-casos para cada ponto obtido e 2º caso  $\lambda = -1$  com mais 4 sub-casos para cada ponto obtido.

Para se aplicar o método, é necessária a obtenção do valor das derivadas de 1ª ordem da função  $g$ , e também das derivadas de 2ª ordem da função  $L$ .

Calculando-se as derivadas necessárias, tem-se:

$$\begin{aligned}
\frac{\partial g}{\partial x_1} &= 2x_1; \\
\frac{\partial g}{\partial x_2} &= 2x_2; \\
\frac{\partial^2 L}{\partial x_1 x_1} &= -2\lambda; \\
\frac{\partial^2 L}{\partial x_2 x_2} &= -2\lambda; \\
\frac{\partial^2 L}{\partial x_1 x_2} &= 2; \\
\frac{\partial^2 L}{\partial x_2 x_1} &= 2.
\end{aligned}$$

Analogamente, substituem-se todos os pontos nas suas derivadas e se aplica o método de Hessiano Orlado.

Sabe-se que as matrizes Hessiano Orlado são:

$$|\bar{H}_2| = \begin{vmatrix} 0 & g_{x_1} \\ g_{x_1} & L_{x_1 x_1} \end{vmatrix}$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & g_{x_1} & g_{x_2} \\ g_{x_1} & L_{x_1 x_1} & L_{x_1 x_2} \\ g_{x_2} & L_{x_2 x_1} & L_{x_2 x_2} \end{vmatrix}.$$

Substituindo-se os valores nos determinantes do Hessiano Orlado para;

1º caso  $\lambda = 1$ .

Sub-caso 1, ponto dado  $(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$ .

$$|\bar{H}_2| = \begin{vmatrix} 0 & 2\sqrt{\frac{1}{2}} \\ 2\sqrt{\frac{1}{2}} & -2 \end{vmatrix} = -2$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & 2\sqrt{\frac{1}{2}} & 2\sqrt{\frac{1}{2}} \\ 2\sqrt{\frac{1}{2}} & -2 & 2 \\ 2\sqrt{\frac{1}{2}} & 2 & -2 \end{vmatrix} = 16.$$

Logo, tem-se que  $|\bar{H}_2| < 0$  e  $|\bar{H}_3| > 0$ , isto é, o ponto  $(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$  é o máximo.

Sub-caso 2, ponto dado  $(-\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$ .

$$|\bar{H}_2| = \begin{vmatrix} 0 & -2\sqrt{\frac{1}{2}} \\ -2\sqrt{\frac{1}{2}} & -2 \end{vmatrix} = -2$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & -2\sqrt{\frac{1}{2}} & -2\sqrt{\frac{1}{2}} \\ -2\sqrt{\frac{1}{2}} & -2 & 2 \\ -2\sqrt{\frac{1}{2}} & 2 & -2 \end{vmatrix} = 16.$$

Logo, tem-se que  $|\bar{H}_2| < 0$  e  $|\bar{H}_3| > 0$ , isto é, o ponto  $(-\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$  é o máximo.

Sub-caso 3, ponto dado  $(-\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$ .

$$|\bar{H}_2| = \begin{vmatrix} 0 & -2\sqrt{\frac{1}{2}} \\ -2\sqrt{\frac{1}{2}} & -2 \end{vmatrix} = -2$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & -2\sqrt{\frac{1}{2}} & 2\sqrt{\frac{1}{2}} \\ -2\sqrt{\frac{1}{2}} & -2 & 2 \\ 2\sqrt{\frac{1}{2}} & 2 & -2 \end{vmatrix} = 0.$$

Logo, tem-se que  $|\bar{H}_2| < 0$  e  $|\bar{H}_3| = 0$ , isto é, para o ponto  $(-\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$  não se pode confirmar nada.

Sub-caso 4, ponto dado  $(\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$ .

$$|\bar{H}_2| = \begin{vmatrix} 0 & 2\sqrt{\frac{1}{2}} \\ 2\sqrt{\frac{1}{2}} & -2 \end{vmatrix} = -2$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & 2\sqrt{\frac{1}{2}} & -2\sqrt{\frac{1}{2}} \\ 2\sqrt{\frac{1}{2}} & -2 & 2 \\ -2\sqrt{\frac{1}{2}} & 2 & -2 \end{vmatrix} = 0.$$

Logo, tem-se que  $|\bar{H}_2| < 0$  e  $|\bar{H}_3| = 0$ , isto é, para o ponto  $(\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$  não se pode confirmar nada.

2º caso  $\lambda = -1$ .

Sub-caso 1, ponto dado  $(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$ .

$$|\bar{H}_2| = \begin{vmatrix} 0 & 2\sqrt{\frac{1}{2}} \\ 2\sqrt{\frac{1}{2}} & 2 \end{vmatrix} = -2$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & 2\sqrt{\frac{1}{2}} & 2\sqrt{\frac{1}{2}} \\ 2\sqrt{\frac{1}{2}} & 2 & 2 \\ 2\sqrt{\frac{1}{2}} & 2 & 2 \end{vmatrix} = 0.$$

Logo, tem-se que  $|\bar{H}_2| < 0$  e  $|\bar{H}_3| = 0$ , isto é, para o ponto  $(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$  não se pode confirmar nada.

Sub-caso 2, ponto dado  $(-\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$ .

$$|\bar{H}_2| = \begin{vmatrix} 0 & -2\sqrt{\frac{1}{2}} \\ -2\sqrt{\frac{1}{2}} & 2 \end{vmatrix} = -2$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & -2\sqrt{\frac{1}{2}} & -2\sqrt{\frac{1}{2}} \\ -2\sqrt{\frac{1}{2}} & 2 & 2 \\ -2\sqrt{\frac{1}{2}} & 2 & 2 \end{vmatrix} = 0.$$

Logo, tem-se que  $|\bar{H}_2| < 0$  e  $|\bar{H}_3| = 0$ , isto é, para o ponto  $(-\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$  não se pode confirmar nada.

Sub-caso 3, ponto dado  $(-\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$ .

$$|\bar{H}_2| = \begin{vmatrix} 0 & -2\sqrt{\frac{1}{2}} \\ -2\sqrt{\frac{1}{2}} & 2 \end{vmatrix} = -2$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & -2\sqrt{\frac{1}{2}} & 2\sqrt{\frac{1}{2}} \\ -2\sqrt{\frac{1}{2}} & 2 & 2 \\ 2\sqrt{\frac{1}{2}} & 2 & 2 \end{vmatrix} = -16.$$

Logo, tem-se que  $|\bar{H}_2| < 0$  e  $|\bar{H}_3| < 0$ , isto é, o ponto do  $(-\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$  é mínimo.

Sub-caso 4, ponto dado  $(\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$ .

$$|\bar{H}_2| = \begin{vmatrix} 0 & 2\sqrt{\frac{1}{2}} \\ 2\sqrt{\frac{1}{2}} & 2 \end{vmatrix} = -2$$

e

$$|\bar{H}_3| = \begin{vmatrix} 0 & 2\sqrt{\frac{1}{2}} & -2\sqrt{\frac{1}{2}} \\ 2\sqrt{\frac{1}{2}} & 2 & 2 \\ -2\sqrt{\frac{1}{2}} & 2 & 2 \end{vmatrix} = -16$$

Logo, tem-se que  $|\bar{H}_2| < 0$  e  $|\bar{H}_3| < 0$ , isto é, o ponto do  $(\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$  é o mínimo.

Portanto, os pontos  $(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}}, 1)$  e  $(-\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}}, 1)$  para  $\lambda = 1$  são os pontos de máximo e  $(-\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}}, -1)$  e  $(\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}}, -1)$  para  $\lambda = -1$  são os pontos de mínimo, de fato é visto na (Figura 10).

De acordo com Koziel (KOZIEL; MICHALEWICZ, 1999), o resultado exato depende a quantidade de variáveis escolhidas, genericamente  $f(x_1, \dots, x_n) = f(\frac{1}{n}, \dots, \frac{1}{n})$  onde o valor do ponto é 1.

A aplicação da metodologia proposta no trabalho foi feita manualmente. Exceto o gráfico, gerado no software GeoGebra, levando cerca de dez segundos para obter os resultados desejados. No trabalho de (PEDROSA; AFONSO, 2011) e (KOZIEL; MICHALEWICZ, 1999) não é citado o tempo computacional.

## 5 Conclusões

A fim de mostrar a aplicação da teoria de Máximo e Mínimo condicionados, com o emprego do Método dos Multiplicadores de Lagrange na determinação do ponto ótimo global para um problema sujeito a uma restrição de igualdade, apresentou-se neste trabalho uma fundamentação teórico-matemática sobre funções reais de várias variáveis e extremos condicionados, bem como conceitos básicos sobre otimização e métodos para determinar pontos ótimos locais ou globais.

Os algoritmos, baseados em métodos aleatório, utilizados na determinação dos pontos que minimizam (ou maximizam) uma função objetivo, na sua maioria, dispendem um certo tempo computacional. Entretanto, os métodos determinísticos, como os Multiplicadores de Lagrange, não apresentam tal dificuldade.

Diante desses aspectos, com aplicação da metodologia proposta neste trabalho foi possível determinar os pontos que satisfazem a função objetivo e a restrição no conjunto viável. Os pontos encontrados foram representados através de um gráfico obtidos com o software GeoGebra, sendo possível a sua visualização. Além disso, pode-se verificar, pelo emprego do método Hessiano Orlado que tais pontos são de fato pontos extremos e não apenas pontos críticos.

Para a efetivação do objetivo deste trabalho, desenvolveu-se a solução analítica, pelo emprego de Multiplicadores de Lagrange, de duas funções globais restritas do grupo de funções  $G$ , determinadas  $G11$  e  $G3$ . A representação gráfica dos problemas foi obtida com o emprego do software de geometria dinâmica GeoGebra. Em um segundo momento, os resultados obtidos foram comparados com dados disponíveis na literatura. A partir dessa comparação, verificou-se uma excelente concordância com os valores determinados com o emprego de algoritmos Evolutivos.

Diante do exposto, considera-se que os objetivos deste trabalho foram atingidos, tendo em vista que, através dos comparativos estabelecidos, pôde-se verificar que a formulação determinística, Multiplicadores de Lagrange, pode propiciar ótimos resultados quando comparada com métodos aleatórios.

Em trabalhos futuros, pretende-se empregar a metodologia de Multiplicadores de Lagrange em problemas com restrição envolvendo desigualdade.

# Referências

- CHIANG, A. C. *Matemática para Economistas*. São Paulo: Universidade de São Paulo, 1982. Citado 5 vezes nas páginas 9, 27, 28, 29 e 30.
- FRITZSCHE, H. *Programação Não Linear; Análise e Métodos*. São Paulo: Universidade de São Paulo, 1978. Citado 3 vezes nas páginas 49, 50 e 51.
- GONÇALVES, M. B.; FLEMMING, D. M. *Cálculo B*. 2. ed. São Paulo: Pearson, Prentice Hall, 2007. Citado 5 vezes nas páginas 12, 17, 18, 19 e 25.
- IZMAILOV, A.; SOLODOV, M. *Otimização; Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade*: volume 1. 2. ed. Rio de Janeiro: Impa, 2007. Citado 2 vezes nas páginas 21 e 22.
- IZMAILOV, A.; SOLODOV, M. *Otimização; Métodos Computacionais*: volume 2. 2. ed. Rio de Janeiro: Impa, 2007. Citado 2 vezes nas páginas 22 e 24.
- KOZIEL, S.; MICHALEWICZ, Z. Evolutionary algorithms, homomorphous mappings, and constrained parameter optimization. *Evolutionary Computation*, v. 5, n. 1, p. 19–44, 1999. Citado 4 vezes nas páginas 12, 33, 39 e 44.
- MUNEM, M. A.; FOULIS, D. J. *Cálculo, Volume 2*: volume 2. Rio de Janeiro: Guanabara, 1982. Citado 3 vezes nas páginas 13, 15 e 25.
- NASCIMENTO, F. A. F. do et al. Uso da meta-heurística otimização por exame de partículas no planejamento florestal. *Scientia Forestalis*, Piracicaba, v. 40, n. 96, p. 557–565, Dezembro 2012. Disponível em: <[http://www.ufjf.br/simmec-emmcomp-2014/files/2014/05/simmec-emmcomp-2014\\_submission\\_19.pdf](http://www.ufjf.br/simmec-emmcomp-2014/files/2014/05/simmec-emmcomp-2014_submission_19.pdf)>. Acesso em: 27.10.2016. Citado na página 32.
- OLIVEIRA, F. R. de. *Estudo de Alguns Métodos Clássicos de Otimização Restrita Não Linear*. Dissertação (Mestrado) — Programa de Pós Graduação em Matemática da Universidade Federal de Uberlândia, Uberlândia, fevereiro 2012. Disponível em: <<http://www.posgrad.famat.ufu.br/node/87>>. Acesso em: 21.09.2016. Citado na página 11.
- PEDROSA, G. T. M. C.; AFONSO, S. M. B. Desenvolvimento de ferramenta de otimização considerando o uso de procedimentos de aproximação com algoritmos evolucionários. *Anais XIX Congresso de Iniciação Científica da UFPE (XIX CONIC)*, Recife, 2011. Citado 5 vezes nas páginas 12, 31, 33, 39 e 44.
- SARAMAGO, S. P.; JÚNIOR, V. S. Introdução às técnicas de otimização em engenharia. *Horizonte Científico*, v. 2, n. 2, 2008. Disponível em: <<http://www.seer.ufu.br/index.php/horizontecientifico/article/viewFile/4050/3015>>. Acesso em: 20.09.2016. Citado 2 vezes nas páginas 11 e 20.
- SILVA, A. F.; LEMONGE, A. C. C.; LIMA, B. S. L. P. Algoritmo de otimização com enxame de partículas auxílio por metamodelos. *XI Simpósio de Mecânica Computacional, II Encontro Mineiro de Modelagem Computacional*, Mato Grosso, 2014. Disponível em:

<[http://www.ufjf.br/simmec-emmcomp-2014/files/2014/05/simmec-emmcomp-2014\\_submission\\_19.pdf](http://www.ufjf.br/simmec-emmcomp-2014/files/2014/05/simmec-emmcomp-2014_submission_19.pdf)>. Acesso em: 26.10.2016. Citado 2 vezes nas páginas 11 e 31.

SILVA, L. M. O. da; MACHADO, M. A. S. *Matemática Aplicada à Administração, Economia e Contabilidade; Funções de uma e mais variáveis*. São Paulo: Cengage Learning, 2010. Citado 2 vezes nas páginas 27 e 29.

SOARES, J. L. Optimização matemática. *Anais Optimization'2004*, Lisboa, 2004. Disponível em: <<http://www.mat.uc.pt/~jsoares/>> Acesso em: 21.09.2016. Citado na página 11.

STEWART, J. *Cálculo*: volume 2. 5. ed. São Paulo: Thomson, 2006. Citado 7 vezes nas páginas 13, 14, 15, 16, 17, 18 e 26.



# Anexos

# ANEXO A – Métodos de Otimização Irrestrita

A seguir apresentam-se dois métodos clássicos de maximização de funções irrestritas, o método do Gradiente e o método de Newton.

## A.1 O Método do Gradiente

Um dos métodos mais antigos e melhor conhecido de maximização de uma função de várias variáveis é o método do Gradiente. Esse método é muito importante do ponto de vista teórico, uma vez que é um dos métodos mais simples que apresenta uma análise satisfatória.

### O Método

Seja  $f$  uma função contínua, com derivadas parciais contínuas de 1ª ordem e matriz identidade de dimensão  $n \times n$ .

O método do Gradiente é definido pelo algoritmo iterativo

$$x_{k+1} = x_k + \tau_k \cdot g_k,$$

onde  $\tau_k$  é um escalar não - negativo que maximiza  $f(x_k + \tau \cdot g_k)$ , isto é, pesquisa-se a partir do ponto  $x_k$  ao longo da direção do Gradiente, até um ponto máximo ser encontrado nessa linha. Neste caso, o máximo é  $x_{k+1}$  (FRITZSCHE, 1978).

## A.2 O Método de Newton

O método do Gradiente baseia-se em aproximações lineares, ou de 1ª ordem, de Taylor a  $f$ .

Considerando uma tal aproximação  $f_L$  no ponto  $x_k$ .

$$f_L(x) = f(x_k) + \nabla f(x_k)(x - x_k).$$

O método de Newton estende a idéia do método do gradiente aproveitando aproximações quadráticas a  $f$ .

Aproximações quadráticas não são somente melhores que aproximações lineares, mas ganham importância à medida que se aproximam do ponto de solução  $x^*$ . Também no procedimento de Newton, um ponto de solução é definido como um ponto  $x^*$ , onde  $\nabla f(x^*) = 0$  (FRITZSCHE, 1978).

# ANEXO B – Métodos de Otimização Restrita

Há problemas em que se procura o máximo de uma função de  $n$  variáveis, sendo que essas variáveis estão sujeitas a restrições lineares e/ou não lineares.

Neste trabalho apresentam-se dois métodos: um dos métodos mais conhecidos, chamados o **método de direções viáveis** (restrições não lineares) e o **método das penalidades** (FRITZSCHE, 1978).

## B.1 O Método de Direções Viáveis

Considere o problema a ser encontrado a solução do seguinte problema não linear

$$\max f(x_1, \dots, x_n) \text{ sujeita à } g_p(x_1, \dots, x_n) \geq 0; h_m(x_1, \dots, x_n) = 0 \quad (\text{B.1})$$

onde  $x$  tem dimensão  $n$ , enquanto  $f$ ,  $g$  e  $h$  tem dimensões 1,  $p$  e  $m$ , respectivamente.

Suponha que todas as funções tenham derivadas parciais contínuas e de ordem 2. Muitas vezes, o método de direções viáveis é também chamado método primal. Outra possibilidade é aproximar o problema (B.1) por uma sequência de problemas aproximados que somente no limite convergirão à solução.

Em resumo, o método de Direções Viáveis é de importância central entre os algoritmos de programação não - linear, devido a sua aplicabilidade geral, simplicidade e boas propriedades de convergência, sobretudo no caso de restrição lineares.

A idéia básica consiste em iterar, através do conjunto  $F$  para que seja viável, da seguinte maneira:

$$x_{k+1} = x_k + \tau_k \cdot d_k,$$

onde  $d_k$  é um vetor de direção e  $\tau_k$  é um escalar não negativo. O escalar  $\tau_k$  é selecionado, tal que  $f$  seja máxima na direção  $d_k$  e, além disso, todo segmento de linha entre  $x_k$  e  $x_{k+1}$  esteja dentro do conjunto  $F$  viável (FRITZSCHE, 1978).

## B.2 O Método das Penalidades

O método de funções de penalidade é um procedimento que visa aproximar problemas de otimização com restrições, por problemas de otimização sem restrições.

Considera-se o seguinte problema restrito

$$\max f(x) \text{ sujeita a } g_i(x) \geq 0; i = 1, 2, \dots, m,$$

onde  $f$  é uma função contínua definida em  $\mathbb{R}^n$  com valores em  $\mathbb{R}$ , e as restrições  $g_i(x) \geq 0$ , determinam um conjunto viável  $F$ , sendo as funções  $g_i(x)$  contínuas.

Substituindo por um problema irrestrito, ou seja,  $\max f(x) + \mu P(x)$ , onde  $\mu$  é uma constante positiva e  $P$  é uma função do  $\mathbb{R}^n$  em  $\mathbb{R}$  satisfazendo as três condições:

- $P(x)$  é contínua no  $\mathbb{R}^n$ ,
- $P(x) \leq 0 \forall x \in \mathbb{R}^n$ ,
- $P(x) = 0 \Leftrightarrow x \in F$ .

$P$  é definida como uma *função de penalidade*. (FRITZSCHE, 1978)