

MILENA ALMEIDA LEITE BRANDÃO

**ESTUDO DE ALGUNS MÉTODOS  
DETERMINÍSTICOS DE OTIMIZAÇÃO  
IRRESTRITA**



UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
FACULDADE DE MATEMÁTICA  
2010

MILENA ALMEIDA LEITE BRANDÃO

# ESTUDO DE ALGUNS MÉTODOS DETERMINÍSTICOS DE OTIMIZAÇÃO IRRESTRITA

**Dissertação** apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal de Uberlândia, como parte dos requisitos para obtenção do título de **MESTRE EM MATEMÁTICA**.

**Área de Concentração:** Matemática.

**Linha de Pesquisa:** Análise Numérica e Otimização.

**Orientador(a):** Prof(a). Dr(a). Sezimária de Fátima Pereira Saramago.

UBERLÂNDIA - MG  
2010

## Dados Internacionais de Catalogação na Publicação (CIP)

B817e	<p>Brandão, Milena Almeida Leite, 1985- Estudo de alguns métodos determinísticos de otimização irrestrita [manuscrito] / Milena Almeida Leite Brandão. - 2010. 82 f.</p> <p>Orientadora: Sezimária de Fátima Pereira Saramago.</p> <p>Dissertação (mestrado) – Universidade Federal de Uberlândia, Programa de Pós-Graduação em Matemática. Inclui bibliografia.</p> <p>1. Otimização matemática - Teses. 2. Análise numérica - Teses. I. Saramago, Sezimária de Fátima Pereira. II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Matemática. III. Título.</p> <p style="text-align: right;">CDU: 519.863</p>
-------	---

Elaborada pelo Sistema de Bibliotecas da UFU / Setor de Catalogação e Classificação



**UNIVERSIDADE FEDERAL DE UBERLÂNDIA**  
**FACULDADE DE MATEMÁTICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA**  
 Av. João Naves de Ávila, 2121, Bloco 1F, Sala 1F 152  
 Campus Santa Mônica, Uberlândia - MG, CEP 38400-902

**ALUNO(A):** Milena Almeida Leite Brandão.

**NÚMERO DE MATRÍCULA:** 93817.

**ÁREA DE CONCENTRAÇÃO:** Matemática.

**LINHA DE PESQUISA:** Análise Numérica e Otimização.

**PÓS-GRADUAÇÃO EM MATEMÁTICA:** Nível Mestrado.

**TÍTULO DA DISSERTAÇÃO:** ESTUDO DE ALGUNS MÉTODOS DETERMINÍSTICOS DE OTIMIZAÇÃO IRRESTRITA.

**ORIENTADOR(A):** Prof(a). Dr(a). Sezimária de Fátima Pereira Saramago.

Esta dissertação foi **APROVADA** em reunião pública realizada na Sala Multiuso da Faculdade de Matemática, Bloco 1F, Campus Santa Mônica, em 12 de março de 2010, às 14:00 horas, pela seguinte Banca Examinadora:

**NOME**

**ASSINATURA**

Prof(a). Dr(a). Sezimária de Fátima Pereira Saramago  
 UFU - Universidade Federal de Uberlândia

Prof. Dr. Rogério Rodrigues dos Santos  
 VSE - Vale Soluções em Energia - S. José dos Campos

Prof. Dr. Valeriano Antunes de Oliveira  
 UFU - Universidade Federal de Uberlândia

Uberlândia-MG, 12 de março de 2010.

# Dedicatória

Dedico este trabalho a Jeová, meus pais Anderson e Neide e ao meu marido Jordaneo.

## Agradecimentos

Agradeço:

- Ao meu Deus Jeová por me abençoar e me guiar.
- A minha mãe Neide pelo grande apoio emocional.
- Ao meu pai Anderson por custear minhas despesas.
- Ao meu marido Jordaneio pela paciência.
- A minha irmã Núbia.
- A minha orientadora Sezimária pela ajuda na escolha do tema desse trabalho, pelos ensinamentos, pela paciência e confiança.
- Aos professores Rogério e Valeriano por terem aceito o convite para participarem da banca examinadora.
- Aos funcionários da FAMAT pela prontidão em ajudar.
- Aos amigos do mestrado, principalmente Carolina, Danilo, Marcelo e Marta, pelo companheirismo.
- A CAPES por financiar meus estudos.

Brandão, M. A. L. *Estudo de alguns métodos determinísticos de otimização irrestrita* 2010. 101 p. Dissertação de Mestrado, Universidade Federal de Uberlândia, Uberlândia-MG.

## Resumo

Neste trabalho são estudados alguns métodos clássicos de busca linear para otimização irrestrita. São apresentadas as principais formulações matemáticas de um problema de otimização e são consideradas duas estratégias, busca linear e região de confiança, para o algoritmo passar de uma iteração para a outra. Além disso, são expostas as principais considerações sobre a escolha do comprimento do passo ao longo da direção de busca para que os métodos sejam convergentes. Dentre os métodos estudados, alguns métodos minimizam uma função de várias variáveis sem o uso de derivadas (Coordenadas Cíclicas, Hooke e Jeeves com busca linear e com passos discretos e Rosenbrock). Os outros métodos necessitam, para o cálculo da direção de busca, das derivadas da função objetivo (Descida Máxima, Newton, Davidon-Fletcher-Powell, Broyden-Flotcher-Goldfarb-Shanno e Gradientes Conjugados). Para ilustração, dois problemas de otimização, um teórico e outro prático, são resolvidos usando cada um dos métodos estudados. Os resultados obtidos são analisados e são apresentadas comparações entre os métodos estudados.

*Palavras-chave:* Otimização irrestrita, otimização determinística, métodos clássicos de otimização, métodos de busca linear.

Brandão, M. A. L. *Study of some deterministic methods for unconstrained optimization* 2010. 101 p. M. Sc. Dissertation, Federal University of Uberlândia, Uberlândia-MG.

## Abstract

In this work some classical methods of linear search for unconstrained optimization are studied. The main mathematical formulations for the optimization problem are presented. Two strategies, linear search and trust region, for the algorithm to move from one iteration to another are discussed. Furthermore, the main considerations about the choice of step length along the search direction to ensure the convergence of the methods are exposed. Among the methods studied, some methods minimize a function of several variables without using derivatives (Cyclic Coordinates, Hooke and Jeeves with line searches and discrete steps and Rosenbrock). In the other methods the search direction is obtained by using derivatives of the objective function (Steepest Descent, Newton, Davidon-Fletcher-Powell, Broyden-Flotcher-Goldfarb-Shanno and Conjugate Gradient). For illustration, two optimization problems, one theoretical and one practical, are solved using the methods mentioned. The results are analyzed and comparisons between the studied methods are presented.

*Keywords:* Unconstrained optimization, deterministic optimization, classical methods of optimization, line search methods.

# Lista de Figuras

1.1	Representação da região viável de $f(x) = (x_1 - 2)^2 + (x_2 - 1)^2$ . . . . .	7
1.2	Representação de mínimos locais e global . . . . .	8
1.3	Conjunto convexo . . . . .	9
1.4	Conjunto não-convexo . . . . .	9
1.5	Bola unitária em $\mathbb{R}^2$ . . . . .	9
1.6	Poliedro convexo . . . . .	10
2.1	Algoritmo de busca linear . . . . .	15
2.2	Gráfico e curvas de nível da função $f(x) = 10(x_2 - x_1^2)^2 + (1 - x_1)^2$ . . . . .	16
2.3	Região de confiança para $f(x) = 10(x_2 - x_1^2)^2 + (1 - x_1)^2$ . . . . .	16
3.1	Comprimento de passo ideal dado pelo mínimo global . . . . .	19
3.2	Ilustração de que o decréscimo de uma função não garante convergência rápida . . . . .	19
3.3	Condição de decréscimo suficiente . . . . .	20
3.4	Condição de curvatura . . . . .	21
3.5	Condições de Wolfe . . . . .	21
3.6	Condições forte de Wolfe . . . . .	22
3.7	<i>Caso 1</i> do método da Seção Áurea: $\phi(c_0) > \phi(d_0) \Rightarrow \alpha_k \in [c_0, b_0]$ . . . . .	23
3.8	<i>Caso 2</i> do método da Seção Áurea: $\phi(d_0) > \phi(c_0) \Rightarrow \alpha_k \in [a_0, d_0]$ . . . . .	23
3.9	<i>Caso 3</i> do método da Seção Áurea: $\phi(c_0) = \phi(d_0) \Rightarrow \alpha_k \in [c_0, d_0]$ . . . . .	24
3.10	Se $\phi(c_s) \geq \phi(d_s)$ , então $[a_{s+1}, b_{s+1}] = [c_s, b_s]$ . . . . .	25
3.11	Se $\phi(c_s) < \phi(d_s)$ , então $[a_{s+1}, b_{s+1}] = [a_s, d_s]$ . . . . .	25
3.12	Gráfico e curvas de nível da função $f(x) = (2x_1 - 3x_2)^2 + (x_2 - 1)^4$ do Exemplo 1 . . . . .	27
3.13	Primeira iteração do Método da Seção Áurea aplicado ao Exemplo 1 . . . . .	28
3.14	Compressor de multi-estágios . . . . .	28
3.15	Gráfico e curvas de nível da função $f(p) = \left( \left( \frac{p_1}{1} \right)^\beta + \left( \frac{p_2}{p_1} \right)^\beta + \left( \frac{64}{p_2} \right)^\beta - 3 \right)$ . . . . .	29
3.16	Curvas de nível da função $f(p)$ para pontos fora da região viável . . . . .	29
4.1	Primeiras duas iterações do Método das Coordenadas Cíclicas . . . . .	31
4.2	Caso onde o método das coordenadas cíclicas foi finalizado prematuramente (Bazaraa <i>et al</i> , 1979) . . . . .	32
4.3	Progresso do Método das Coordenadas Cíclicas aplicado ao Exemplo 1 . . . . .	34
4.4	Progresso do Método das Coordenadas Cíclicas aplicado ao Exemplo 2 . . . . .	34
4.5	Primeiras duas iterações do Método de Hooke e Jeeves, (Bazaraa <i>et al</i> , 1979) . . . . .	35
4.6	Progresso de Hooke e Jeeves utilizando buscas lineares aplicado ao Exemplo 1 . . . . .	37
4.7	Progresso de Hooke e Jeeves utilizando buscas lineares aplicado ao Exemplo 2 . . . . .	37
4.8	Progresso de Hooke e Jeeves com passos discretos aplicado ao Exemplo 1 . . . . .	40
4.9	Progresso de Hooke e Jeeves com passos discretos aplicado ao Exemplo 2 . . . . .	40
4.10	Progresso do Método Rosenbrok aplicado ao Exemplo 1 . . . . .	45
4.11	Progresso do Método Rosenbrok aplicado ao Exemplo 2 . . . . .	45

5.1	Direção de descida máxima . . . . .	48
5.2	Uma direção de descida $p_k$ . . . . .	48
5.3	Passos do Método da Descida Máxima . . . . .	51
5.4	Progresso do Método da Descida Máxima aplicado ao Exemplo 1 . . . . .	55
5.5	Progresso do Método da Descida Máxima aplicado ao Exemplo 2 . . . . .	55
6.1	Progresso do Método de Newton aplicado ao Exemplo 1 . . . . .	60
7.1	Progresso do Método Quase-Newton (DFP) aplicado ao Exemplo 1 . . . . .	71
7.2	Progresso do Método Quase-Newton (BFGS) aplicado ao Exemplo 1 . . . . .	71
7.3	Progresso do Método Quase-Newton (DFP) aplicado ao Exemplo 2 . . . . .	72
7.4	Progresso do Método Quase-Newton (BFGS) aplicado ao Exemplo 2 . . . . .	73
8.1	Progresso do Método dos Gradientes Conjugados aplicado ao Exemplo 1 . . . . .	79
8.2	Progresso do Método dos Gradientes Conjugados aplicado ao Exemplo 2 . . . . .	82

# Lista de Tabelas

3.1	Método da Seção Áurea aplicado ao Exemplo 1 . . . . .	27
3.2	Método da Seção Áurea aplicado ao Exemplo 2 . . . . .	30
4.1	Método das Coordenadas Cíclicas aplicado ao Exemplo 1 . . . . .	33
4.2	Método das Coordenadas Cíclicas aplicado ao Exemplo 2 . . . . .	34
4.3	Método de Hooke e Jeeves utilizando buscas lineares aplicado ao Exemplo 1 . . . . .	36
4.4	Método de Hooke e Jeeves utilizando buscas lineares aplicado ao Exemplo 2 . . . . .	38
4.5	Método Hooke e Jeeves com passos discretos aplicado ao Exemplo 1 . . . . .	40
4.6	Método Hooke e Jeeves com passos discretos aplicado ao Exemplo 2 . . . . .	41
4.7	Método Rosenbrock aplicado ao Exemplo 1 . . . . .	44
4.8	Método Rosenbrock aplicado ao Exemplo 2 . . . . .	46
5.1	Método da Descida Máxima aplicado ao Exemplo 1 . . . . .	54
5.2	Método da Descida Máxima aplicado ao Exemplo 2 . . . . .	55
6.1	Método de Newton aplicado ao Exemplo 1 . . . . .	60
6.2	Método de Newton aplicado ao Exemplo 2 . . . . .	61
7.1	Método Quase-Newton (DFP) aplicado ao Exemplo 1 . . . . .	70
7.2	Método Quase-Newton (BFGS) aplicado ao Exemplo 1 . . . . .	71
7.3	Método Quase-Newton (DFP) aplicado ao Exemplo 2 . . . . .	72
7.4	Método Quase-Newton (BFGS) aplicado ao Exemplo 2 . . . . .	72
8.1	Método dos Gradientes Conjugados aplicado ao Exemplo 1 . . . . .	80
8.2	Método dos Gradientes Conjugados aplicado ao Exemplo 2 . . . . .	81
9.1	Resultados ótimos encontrados para o Exemplo 1, aplicando as técnicas estudadas	84
9.2	Resultados ótimos encontrados para o Exemplo 2, aplicando as técnicas estudadas	85

# Lista de Símbolos

$B^k$	Matriz Hessiana ou uma aproximação desta
$c_i(x)$	Função de restrição
$D$	Conjunto de índices das restrições de desigualdade
$d_1, \dots, d_n$	Direções de busca conjugadas
$f(x)$	Função objetivo
$G_k$	Matriz Hessiana média
$H^k$	Aproximação para a inversa da Hessiana
$I$	Conjunto de índices das restrições de igualdade
$I_b = [a_0, b_0]$	Intervalo de incerteza de busca
$k(Q)$	Fator de condição
$m_k$	Função modelo na $k$ -ésima iteração
$p_k$	Direção de busca na $k$ -ésima iteração
$Q$	Matriz simétrica positiva definida
$s_k$	Varição do vetor variáveis de projeto entre duas iterações consecutivas
$x$	Vetor variáveis de projeto
$x_i$	$i$ -ésima coordenada de $x$
$x^*$	Mínimo global de uma função
$x^0$	Ponto inicial do vetor das variáveis de projeto
$\{x^k\}$	Sequência de aproximações do vetor das variáveis de projeto
$y_k$	Varição do gradiente aplicado as variáveis de projeto entre duas iterações consecutivas
$\alpha_k$	Comprimento do passo na direção de busca $p_k$
$\Delta$	Comprimento de passo inicial do Método de Hooke e Jeeves
$\delta$	Raio da região de confiança
$\epsilon$	Precisão adotada em algum critério de parada do algoritmo
$\phi(\alpha)$	Função uni-dimensional do comprimento do passo $\alpha$
$\lambda$	Fator de aceleração do método de Hooke e Jeeves
$\lambda_i$	Autovalores
$\theta$	Ângulo entre $p_k$ e $\nabla f_k$
$\tau$	Taxa de redução do intervalo de incerteza do Método da Seção Áurea
$\nabla f(x)$	Vetor gradiente da função $f(x)$
$\nabla^2 f(x)$	Matriz Hessiana da função $f(x)$
$\ \cdot\ _2$	Norma Euclidiana
$\ \cdot\ $	Norma qualquer
$\ \cdot\ _F$	Norma peso de Frobenius

# Sumário

Resumo	vii
Abstract	viii
Lista de Símbolos	xii
Introdução	1
<b>1 Definição do Problema de Otimização</b>	<b>5</b>
1.1 Formulação matemática do problema de otimização . . . . .	6
1.2 Classificação dos métodos de otimização . . . . .	6
1.3 Revisão: algumas definições importantes . . . . .	9
<b>2 Fundamentos da Otimização Irrestrita</b>	<b>12</b>
2.1 Algoritmo para a direção de busca . . . . .	14
<b>3 Escolha do passo nos Métodos de Busca Linear</b>	<b>18</b>
3.1 Considerações sobre a escolha do comprimento do passo . . . . .	18
3.2 Condições de Wolfe . . . . .	20
3.3 Busca unidimensional usando o Método da Seção Áurea . . . . .	22
<b>4 Métodos de Otimização Determinística Independentes de Derivadas</b>	<b>31</b>
4.1 O Método das Coordenadas Cíclicas . . . . .	31
4.2 O Método de Hooke e Jeeves com busca linear . . . . .	35
4.3 O Método de Hooke e Jeeves com passos discretos . . . . .	39
4.4 O Método de Rosenbrock . . . . .	42
<b>5 Método da Descida Máxima</b>	<b>47</b>
5.1 Direção de descida máxima (steepest descent direction) . . . . .	47
5.2 Convergência dos Métodos de Descida Máxima . . . . .	48
5.3 Taxas de convergência do Método de Descida Máxima . . . . .	50
<b>6 Método de Newton</b>	<b>56</b>
6.1 Convergência do Método de Newton . . . . .	57
6.2 Taxa de convergência do Método de Newton . . . . .	57
<b>7 Métodos Quase-Newton</b>	<b>62</b>
7.1 Direção de busca dos Métodos de Quase-Newton . . . . .	62
7.2 Convergência dos Métodos Quase-Newton . . . . .	63
7.3 Taxas de convergência dos Métodos Quase-Newton . . . . .	64
7.4 Os métodos BFGS e DFP . . . . .	65

<b>8</b>	<b>Método dos Gradientes Conjugados</b>	<b>74</b>
	78	
8.2	Recomendações para a reinicialização do método dos gradientes conjugados . . .	82
<b>9</b>	<b>Discussões e Conclusões</b>	<b>84</b>
9.1	Conclusões . . . . .	85

# Introdução

Segundo Pirsig (1974) a “Otimização matemática é o título formal dada ao ramo da ciência computacional, que procura responder à pergunta Qual é o melhor? para os problemas em que a qualidade de qualquer resposta pode ser expressa como um valor numérico. Tais problemas surgem em todas as áreas da matemática, das ciências físicas, químicas e biológicas, engenharia, arquitetura, economia e gestão, e a gama de técnicas disponíveis para resolvê-los é quase tão grande quanto”.

Mas não é só nas ciências que se encontram problemas de otimização. Otimização é uma realidade que se aplica todos os dias, em quase todos os assuntos. Até mesmo os animais de uma forma instintiva aplicam a otimização no seu dia-a-dia. Um exemplo clássico disto são as abelhas. Esses insetos usam cera para construir os alvéolos das colméias, os quais são usados como depósitos para o mel que fabricam. Os alvéolos que compõem o favo de mel das abelhas européias possuem um formato poliédrico. Mas por que não circular, já que o círculo é a forma geométrica que tem maior volume com menor área? Por que as abelhas constroem os alvéolos procurando a forma mais econômica, ou seja, que apresente maior volume para a menor porção de material gasto. Como os alvéolos não poderiam ser, por exemplo, cilíndricos, pois a falta de paredes comuns entre eles deixaria uma grande quantidade de espaços inaproveitáveis, a solução encontrada pelas abelhas é que os alvéolos devem ter a forma de um prisma, pois assim, as paredes de um alvéolo servem também aos alvéolos vizinhos, desta forma conseguem guardar um maior volume de mel com a mínima quantidade de material gasto. Este é apenas um dos vários exemplos encontrados na natureza.

Devido a combinação elegante entre prática e teoria, a otimização hoje “está na moda”. Uma amostra desta combinação pode ser vista na tese de mestrado de Carrera (2006). No seu estudo é apresentado o desenvolvimento de um modelo dinâmico de veículo de corrida considerando as características suficientes para análise da trajetória, influenciada por parâmetros geométricos e físicos pertinentes. Em seguida é definido o problema de obtenção da trajetória empregando procedimentos de otimização, de modo a determinar como um veículo irá percorrer um traçado, considerando como função objetivo o tempo de percurso, que deverá ser mínimo, e tendo como restrições as condições dinâmicas do veículo e geométricas da pista, implementando rotinas que são usadas em conjunto com os algoritmos comerciais existentes.

Um trabalho interessante na área da economia é o de Bueno (2008). Na sua dissertação de mestrado é feita uma exposição sobre alguns modelos matemáticos com aplicações em economia. Dentre os modelos estudados é destacado a versão discreta das populares medidas de risco VaR (Value at Risk) e C-VaR (Conditional Value at Risk). Argumenta-se algumas propriedades de tais medidas, e, principalmente, apresenta-se algumas idéias para otimizá-las sob uma formulação do tipo OVO (Order Value Optimization) e propõe-se uma nova formulação para o problema de minimizar a VaR.

É por causa de exemplos como estes que a otimização se tornou atraente. A primeira técnica de otimização, a qual é conhecida como descida mais inclinada, retorna ao tempo de Gauss. Historicamente, o primeiro termo a ser introduzido foi programação linear, inventado por George Dantzig na década de 1940. Hoje o termo otimização, ou programação matemática, refere-se ao estudo de problemas em que se busca minimizar ou maximizar uma função através

da escolha dos valores de variáveis dentro de um conjunto viável. Mas na maioria das vezes o valor ótimo é difícil de ser encontrado num único cálculo. Por isso, é necessário um processo sistemático de busca, um procedimento iterativo de cálculos que, passo a passo, permita ir melhorando a seleção até que o ótimo seja encontrado, ou até que seja satisfeito algum critério.

Em muitos problemas é possível encontrar o ótimo ou pelo menos aproximar-se deste pois, grande parte dos algoritmos convergem para o ótimo quando algumas hipóteses são obedecidas. No trabalho de Shi (2000) é proposto uma nova estratégia global para resolver problemas irrestritos baseado na idéia de combinar a direção de Newton e a direção de descida máxima dentro de cada iteração. A convergência global é garantida para qualquer ponto inicial, visto que em cada iteração a direção de busca é escolhida de forma a estar tão próximo da direção de Newton quanto possível. Uma proposta de combinar a direção de Quase-Newton com direção de descida máxima também é considerada e são apresentados simulações numéricas.

Um outro trabalho interessante é de Liu *et al.* (1999) no qual é introduzido uma classe de métodos de gradiente conjugado não-monótono, que incluem os métodos de Polak-Ribière e Hestenes-Stiefel como casos especiais. Prova-se que esta classe de métodos de gradiente conjugado não-monótono é globalmente convergente quando aplicada para resolver problemas de otimização irrestrita com funções objetivo convexas. Experimentos numéricos mostram que os métodos de Polak-Ribière e Hestenes-Stiefel não-monótonos nesta classe de métodos de gradiente conjugado não-monótono são competitivos com os métodos monótonos homólogos.

Varela (2006) em sua dissertação de mestrado apresentou a formulação e solução de problemas de programação não-linear em projeto estruturado de amplificadores de retro-alimentação negativa. O projeto eletrônico estruturado surgiu como uma nova metodologia de projeto nos anos 90 e se baseia na modificação passo-a-passo da solução ideal até chegar a uma solução real que obedeça as especificações originais determinadas pelo projetista. A otimização dos amplificadores é resolvida utilizando uma metodologia híbrida onde a técnica heurística Evolução Diferencial é aplicada em conjunto com o método determinístico clássico de Hooke-Jeeves.

Técnicas clássicas de otimização determinística foram aplicadas por Bergamaschi *et al.* (2006) para a obtenção do projeto ótimo de manipuladores com três juntas de revolução, considerando a maximização do seu volume de trabalho.

Santos *et al.* (2008) descrevem um estudo sobre o planejamento ótimo de trajetórias para um robô serial, considerando um espaço de trabalho restrito, usando técnicas de controle ótimo. Trata-se de um problema multi-objetivo e entre os métodos testados pelos autores os melhores resultados foram obtidos usando técnicas sequenciais de programação.

Existem vários outros artigos interessantes em otimização e a importância desta área do conhecimento pode ser avaliada por algumas importantes revistas especializadas, tais como: SIAM Journal on Optimization (SIOPT), Journal of Global Optimization, Structural and Multidisciplinary Optimization, Mathematical Programming Society, Optimization, entre outras.

A praticidade da otimização não depende somente da eficiência e robustez dos algoritmos, mas também de boas técnicas de modelagem, interpretação cuidadosa dos resultados e um bom programa computacional. Neste trabalho foram mostrados vários aspectos do processo de otimização, contudo, mais informações podem ser encontradas nos seguintes livros: Bazaraa *et al.* (1979), Bertsekas (1999), Bonnans *et al.* (2002), Davidon (1991), Friedlander (1994), Gill *et al.* (1981), Himmelblau (1972), Lasdon (1970).

O objetivo deste trabalho é estudar alguns métodos clássicos de busca linear para otimização irrestrita e comprovar a eficiência destes aplicando-os em dois problemas, um teórico e outro prático, para em seguida analisar os resultados obtidos. Para este fim, são apresentadas as principais formulações matemáticas de um problema de otimização e duas estratégias, busca linear e região de confiança, para o algoritmo se mover de uma iteração para a outra. Além disso, são discutidas considerações sobre a escolha do comprimento do passo ao longo da direção de busca para que os métodos sejam convergentes partindo de um ponto inicial qualquer.

A dissertação está dividida do seguinte modo:

Capítulo 1: Formulação matemática do problema de otimização e classificação dos métodos de otimização. Em seguida, apresenta-se uma pequena revisão de definições importantes.

Capítulo 2: Definições de mínimos locais e globais e considerações das condições necessárias para a existência do ponto ótimo. Com a existência do ótimo estabelecida, são apresentadas duas estratégias de busca para se encontrar o ponto ótimo, a saber, busca linear e região de confiança.

Capítulo 3: Consideração sobre como escolher o comprimento do passo ao longo da direção de busca para garantir que o método seja convergente. Mostra-se que para a convergência ocorrer é necessário que o passo satisfaça as condições de Wolfe. Também é estudado o método de busca unidimensional Seção Áurea.

Nos próximos capítulos, são estudados alguns métodos de otimização irrestrita e seus algoritmos utilizados para se minimizar duas funções não-lineares. Os resultados obtidos estão expostos em tabelas e figuras e no capítulo 9 é feita uma comparação da eficiência dos métodos para estas aplicações.

Capítulo 4: São considerados quatro métodos que minimizam uma função de várias variáveis sem o uso de derivadas: Coordenadas Cíclicas, Hooke e Jeeves com busca linear e com passos discretos e o método de Rosenbrock. Para cada uma destas técnicas é estabelecido um algoritmo computacional.

Capítulo 5: Estudo do método de busca linear Descida Máxima. Mostra-se que sua direção de busca é realmente uma direção de descida e por meio do Teorema de Zoutendijk prova-se a convergência global do método. Ainda, com o auxílio de outros teoremas, obtém-se que a taxa de convergência do método é linear.

Capítulo 6: São investigadas as principais considerações sobre a direção de busca do Método de Newton. Mostra-se que se a matriz Hessiana é definida positiva, numericamente limitada e o comprimento do passo satisfaz as condições de Wolfe, então o método é convergente, mais ainda, se a Hessiana for Lipschitz contínua, a convergência é quadrática.

Capítulo 7: São apresentados dois Métodos Quase-Newton: Davidon-Fletcher-Powell (DFP) e Broyden-Flotcher-Goldfarb-Shanno (BFGS). Mostra-se que os métodos Quase-Newton possuem convergência super-linear. Também são consideradas as fórmulas de atualização da aproximação da inversa da matriz Hessiana. Por fim, é fornecida uma expressão unificada para estes métodos.

Capítulo 8: Desenvolvimento do Método dos Gradiente Conjugados: define-se as direções conjugadas e mostra-se que estas são direções de descida. É feita uma consideração sobre o parâmetro de desvio e um algoritmo do método de Fletcher e Reeves é estabelecido. Mostra-se também que o método é convergente para funções quadráticas e não-quadráticas.

Capítulo 9: É feita uma comparação dos métodos estudados e apresentadas algumas conclusões sobre este estudo. Além disso, são propostos alguns temas para a continuidade do trabalho.

Milena Almeida Leite Brandão  
Uberlândia-MG, 12 de março de 2010.

# Capítulo 1

## Definição do Problema de Otimização

A otimização é um processo que faz parte da natureza, por exemplo, os sistemas físicos tendem ao estado de energia mínima e os feixes de luzes seguem um caminho que minimiza o tempo de percurso.

A otimização é uma ferramenta importante para tomada de decisão durante a análise e o projeto de sistemas físicos. Para utilizar esta ferramenta torna-se necessário identificar qual o objetivo, ou seja, a quantidade que irá medir a performance do sistema (ex: energia, custo, potência, tempo, temperatura, etc.). O objetivo depende das características do sistema, que são as *variáveis* ou *parâmetros*, denominados variáveis de projeto ou variáveis de decisão. Assim, a otimização visa encontrar o valor das variáveis para otimizar o objetivo. Normalmente, estas variáveis devem obedecer a restrições, que são limites impostos ao sistema estudado.

O processo de identificação dos objetivos, variáveis e restrições é denominado *modelagem* do problema. Escrever um modelo adequado é muitas vezes o passo mais importante do processo de otimização. Se o modelo é muito simples, ele pode não representar o problema real. Por outro lado, se o modelo é muito complexo, ele pode trazer muitas dificuldades para a obtenção da solução.

Após a formulação do modelo, um algoritmo de otimização é utilizado para obter a solução do problema, usualmente com a ajuda de códigos computacionais. Não existe um algoritmo universal, mas um conjunto de métodos, entre os quais alguns são mais apropriados para determinadas aplicações específicas. A escolha do método de otimização depende do usuário e pode determinar a eficácia ou falha para a solução do problema.

Após a aplicação do algoritmo o projetista deve ser capaz de analisar a solução encontrada procurando identificar se o processo de otimização obteve sucesso ou se falhou. Em uma análise teórica, existem elegantes expressões matemáticas conhecidas como condições necessárias para existência do ponto ótimo (optimality condition) para verificar se a solução atual representa a solução ótima local do problema, mas nos problemas reais elas são difíceis de serem verificadas.

O modelo pode ser aperfeiçoado usando “análise de sensibilidade” que permite avaliar quais variáveis de projeto são mais relevantes no modelo que representa o problema.

Um problema de otimização é freqüentemente chamado de programação matemática. Os algoritmos de otimização são *iterativos*. Iniciam com uma estimativa inicial  $x^0$  e geram uma sequência de aproximações até encontrar o ponto mínimo. As estratégias utilizadas para se mover de uma iteração para outra é que distingue os diversos algoritmos. Muitas estratégias (métodos clássicos) usam o valor da função objetivo e suas derivadas. Algumas estratégias armazenam informações de iterações anteriores e outras consideram a informação apenas da iteração atual. Bons algoritmos de otimização devem ter as seguintes propriedades:

*Robustez*: devem ter uma boa performance para uma grande variedade de problemas e obter resultados razoáveis para quaisquer pontos de partida.

*Eficiência*: não devem exigir tempo computacional excessivo ou grande capacidade de ar-

mazenamento de dados.

*Precisão*: devem ser capazes de identificar a solução com precisão, sendo pouco sensíveis a erros nos dados ou arredondamentos numéricos quando o algoritmo é implementado no computador.

## 1.1 Formulação matemática do problema de otimização

O problema de otimização pode ser escrito como:

$$\begin{aligned} \min f(x) \quad \text{sujeito a} \quad & c_i(x) = 0, i \in I \\ x \in \mathbb{R}^n \quad & c_i(x) \geq 0, i \in D \end{aligned} \quad (1.1)$$

sendo que,

- $x$  é o vetor das variáveis de projeto (ou variáveis de decisão, ou parâmetros do projeto);
- $f$  é a função objetivo (ou índice de performance, ou função custo);
- $c_i$  são funções de restrição (funções escalares de  $x$  que definem certas equações ou inequações que as variáveis devem obedecer);
- $I$  e  $D$  representam os conjuntos de índices das restrições de igualdade e desigualdade, respectivamente.

Por exemplo, considere, o problema de otimização

$$\begin{aligned} \min (x_1 - 2)^2 + (x_2 - 1)^2 \\ \text{sujeito a} \quad & x_1^2 - x_2 \leq 0 \\ & x_1 + x_2 \leq 2 \end{aligned} \quad (1.2)$$

Reescrevendo o problema 1.2 na forma definida em 1.1, vem:

$$f(x) = (x_1 - 2)^2 + (x_2 - 1)^2, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad c(x) = \begin{bmatrix} c_1(x) \\ c_2(x) \end{bmatrix} = \begin{bmatrix} -x_1^2 + x_2 \\ -x_1 - x_2 + 2 \end{bmatrix},$$

$D = \{1, 2\}$ ,  $I = \emptyset$ .

A Figura 1.1 mostra as *curvas de nível*, que são o conjunto de pontos para os quais  $f(x)$  possui valor constante. Além disso, apresenta a *região viável*, que são o conjunto de pontos que satisfazem a todas as restrições. O ponto  $x^*$  representa a solução ótima do problema, na presença de restrições.

## 1.2 Classificação dos métodos de otimização

Existem diversos métodos de otimização e desta forma vários autores já apresentaram diferentes propostas para classificá-los, não sendo possível escolher uma proposta universalmente aceita. Os métodos dependem do tipo do problema em estudo (programação linear ou não-linear, problemas restritos ou irrestritos), das variáveis de projeto (contínuas ou discretas), do tipo de resposta desejada (otimização local ou global; projeto ótimo ou controle ótimo), das técnicas empregadas (determinísticas, estocástica ou híbrida). Além disso, os problemas podem ser classificados segundo o número de variáveis (pequenos ou grandes) ou segundo a suavidade das funções (diferenciáveis ou não-diferenciáveis). A seguir serão apresentadas as denominações mais comuns nesta área:

- *Problemas de programação linear ou não-lineares*: quando a função objetivo e todas as restrições são funções lineares de  $x$  (este tipo de problema é bastante comum em aplicações econômicas e nas áreas de transporte) denominam-se os problemas de programação linear.

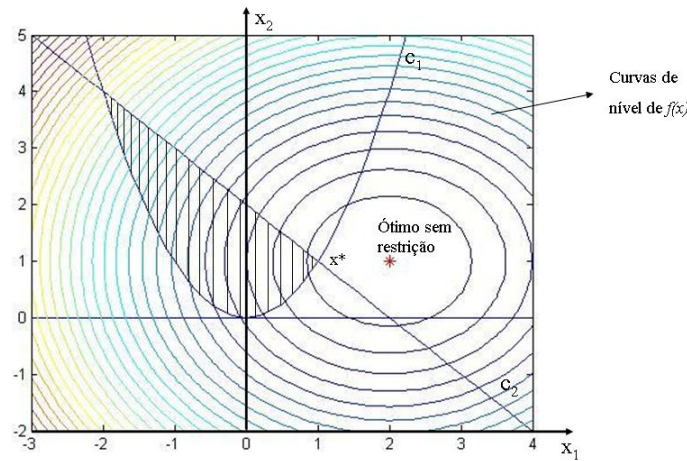


Figura 1.1: Representação da região viável de  $f(x) = (x_1 - 2)^2 + (x_2 - 1)^2$

Os problemas de programação não-lineares são aqueles onde ao menos uma das restrições ou a função objetivo são não-lineares, eles aparecem naturalmente nas ciências físicas e nas engenharias.

- *Otimização irrestrita ou restrita:* problemas de otimização irrestritos são aqueles onde  $I = D = \emptyset$ . Este caso pode acontecer quando: os problemas práticos não dependem de restrições; as restrições foram desprezadas porque não afetam diretamente a solução ou o problema com restrição foi reescrito de forma que as restrições foram substituídas por funções de penalidade.

Os problemas de otimização restritos, que aparecem devido à imposições do projeto, são aqueles onde  $I \neq \emptyset$  e  $D \neq \emptyset$ . As restrições podem ser lineares (ex:  $\sum x_i \leq 1$ ), laterais (ex:  $0 \leq x_i \leq 20$ ), não-lineares (ex:  $\sum x_i^2 - \text{sen}x_i \leq 0,5$ ), ou, ainda, combinações destas.

- *Otimização discreta ou contínua:* em alguns problemas de otimização as variáveis de projeto só têm sentido se forem considerados seus valores inteiros (ex: a variável  $x_i$  representa o número de caminhões em uma frota de transporte de cargas). Nestes tipos de problemas, deve-se incluir as “restrições de integralidade”, ou seja  $x_i \in \mathbb{Z}$  ( $\mathbb{Z}$  é o conjunto dos números inteiros). Pode-se também incluir “restrições binárias”, neste caso,  $x_i \in \{0,1\}$ . Este tipo de problemas é denominado problemas de otimização inteira ou *problemas de otimização discreta*. Neste caso, os problemas podem considerar além das variáveis inteiras ou binárias, variáveis abstratas, as quais são representadas por conjuntos finitos (que podem ser bastante grandes). Na *otimização contínua* as variáveis  $x_i \in \mathbb{R}$ .
- *Otimização global ou local:* a solução *global* é aquele ponto onde a função objetivo atinge o *menor* valor entre todos os pontos da região viável. A solução ou *ótimo local* é o ponto cuja função objetivo é menor que todos os pontos vizinhos da região viável, conforme representado na Figura 1.2. Em muitas aplicações a solução global é desejada, mas difícil de ser reconhecida. Para problemas *convexos* e na *programação linear* pode-se garantir que uma solução local é também solução global.
- *Projeto ótimo ou controle ótimo:* segundo Arora (1987) o projeto ótimo e o controle ótimo de sistemas são duas atividades distintas. Existem numerosas aplicações onde a obtenção de projeto ótimo é útil para a concepção dos sistemas. Existem outras aplicações onde os conceitos de controle ótimo são necessários. Além disso, existem algumas aplicações

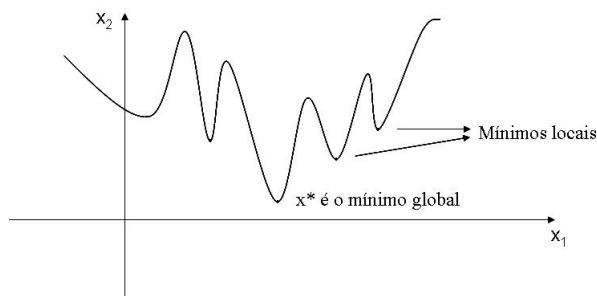


Figura 1.2: Representação de mínimos locais e global

onde tanto os conceitos de projeto ótimo quanto os de controle ótimo são usados. Uma amostra de onde isto ocorre inclui a robótica e estruturas aeroespaciais. Acontece que problemas de controle ótimo podem ser transformados em problemas de projeto ótimo e tratados pelos métodos de otimização. Um problema de controle ótimo visa encontrar uma entrada controlada para o sistema produzir a saída desejada. Neste caso, o sistema tem componentes ativos que percebem as flutuações na saída. Os sistemas de controles são automaticamente ajustados para corrigir a saída e otimizar alguma medida de desempenho. Normalmente, o controle é aplicado em problemas de natureza dinâmica. Por outro lado, no projeto ótimo escreve-se o sistema e os componentes para otimizar a função objetivo. O sistema então permanece fixo durante toda a aplicação. Esta é a maior diferença entre as duas aplicações. Como um exemplo, considere o mecanismo de controle de viagem de um carro de passageiros. A idéia da resposta do sistema é controlar a injeção de combustível para manter a velocidade do carro constante. Assim, a saída do sistema é conhecida, ou seja, a velocidade da viagem. O trabalho do mecanismo de controle é perceber as flutuações na velocidade e ajustar adequadamente a injeção de combustível. A quantidade de injeção de combustível depende das condições da estrada. Quando o carro está numa subida, a injeção de combustível é maior do que em uma descida.

- *Otimização determinística ou estocástica:* nos problemas de *otimização determinística* o modelo é completamente conhecido. Os *algoritmos clássicos* de otimização, que dependem do conhecimento das derivadas da função objetivo, são exemplos de algoritmos determinísticos. Os melhores resultados destes métodos são para funções contínuas, convexas e semi-modais. Os métodos determinísticos possuem uma grande vantagem que é o baixo número de avaliações da função objetivo, fazendo com que tenham convergência rápida e baixo custo computacional.

Na *otimização estocástica* pode-se trabalhar com “incertezas” de algumas variáveis e produzir soluções que otimizam a “performance esperada” do modelo. Pode-se garantir que as variáveis  $x$  satisfaçam as restrições para alguma probabilidade especificada anteriormente (modelos probabilísticos). Estes métodos também são conhecidos como métodos heurísticos. O termo heurístico tem sua origem na palavra grega “heuriskein” que significa “encontrar/descobrir”. Estes métodos utilizam apenas informações da função a ser otimizada, que pode ser de difícil representação, não-diferenciável, descontínua, não-linear, multimodal. Alguns exemplos dos métodos heurísticos são os métodos naturais, que tentam simular os processos usados na natureza para resolver problemas complexos. Entre estes métodos pode-se citar: Busca Tabu, Simulated Annealing, métodos baseados em população (Algoritmos Genéticos, Evolução Diferencial, Otimização por Colônia de Formigas, Otimização por Bando de Pássaros, etc). Estas técnicas trabalham com um conjunto de pontos, necessitam de muitas avaliações da função objetivo e sua maior desvantagem é o alto custo computacional.

- *A otimização híbrida:* consiste na utilização conjunta de métodos determinísticos e heurísticos. Tem sido largamente utilizada nos dias atuais, visa unir as boas características de convergência dos métodos determinísticos às vantagens dos métodos heurísticos, por exemplo, a independência do tipo de função e o fato de não ficarem “presos” pelos mínimos locais.

### 1.3 Revisão: algumas definições importantes

O conceito de convexidade é fundamental em otimização, sendo que os problemas que possuem esta propriedade são mais fáceis de serem solucionados. Considere as definições abaixo:

**Definição 1.1** *Um conjunto  $S \subset \mathbb{R}^n$  é um conjunto convexo se um segmento de reta unindo dois pontos de  $S$  permanece completamente em  $S$ . Ou seja, para quaisquer dois pontos  $x \in S$  e  $y \in S$ , então*

$$ax + (1 - a)y \in S \text{ para todo } a \in [0, 1]. \quad (1.3)$$

A Figura 1.3 representa um conjunto convexo e a Figura 1.4 um conjunto não-convexo.

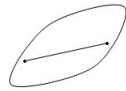


Figura 1.3: Conjunto convexo

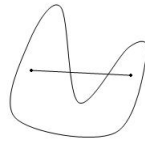


Figura 1.4: Conjunto não-convexo

**Definição 1.2** *Uma função  $f$  é convexa se o domínio  $S$  é um conjunto convexo e se para quaisquer dois pontos  $x$  e  $y$  em  $S$ , a seguinte propriedade é satisfeita:*

$$f(ax + (1 - a)y) \leq af(x) + (1 - a)f(y), \text{ para todo } a \in [0, 1]. \quad (1.4)$$

Deve-se observar que a função  $f$  é estritamente convexa se a desigualdade (1.4) é estrita quando  $x \neq y$  e  $a$  está no intervalo aberto  $(0, 1)$ . A função  $f$  é dita côncava se  $-f$  é convexa.

**Definição 1.3** *A bola unitária definida por  $S = \{y \in \mathbb{R}^n; \|y\|_2 \leq 1\}$ , é um conjunto convexo, conforme Figura 1.5.*

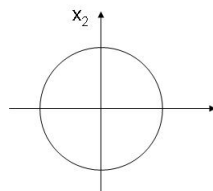


Figura 1.5: Bola unitária em  $\mathbb{R}^2$

**Definição 1.4** *Uma matriz  $B$  é chamada positiva definida se  $p^T B p > 0$ , para qualquer  $p \in \mathbb{R}^n$  e  $p \neq 0$ . A matriz  $B$  é positiva semi-definida se  $p^T B p \geq 0$ , para qualquer  $p \in \mathbb{R}^n$ .*

**Definição 1.5** Qualquer poliedro definido pelo conjunto de equações e inequações lineares  $S = \{x \in \mathbb{R}^n; Ax = b, Cx \leq d\}$  onde  $A$  e  $C$  são matrizes,  $b$  e  $d$  vetores (com dimensões apropriadas), é um conjunto convexo. Por exemplo, a região viável mostrada na Figura 1.6 é um conjunto convexo.

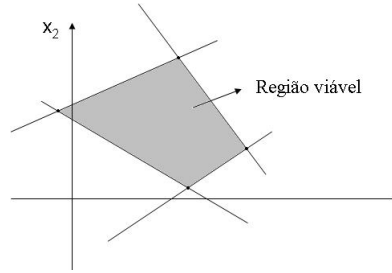


Figura 1.6: Poliedro convexo

Como exemplos de funções convexas podem ser citadas:

(a) Funções lineares:  $f(x) = c^T x + a$ , para todo vetor  $c \in \mathbb{R}^n$  e  $a \in \mathbb{R}$  (escalar).

(b) Funções quadráticas convexas:  $f(x) = x^T Q x$ , onde  $Q$  é uma matriz simétrica positiva definida.

Para o estudo dos métodos determinísticos de otimização outro conceito importante é a aproximação de funções continuamente diferenciáveis por série de Taylor:

**Teorema 1.1** (Teorema de Taylor): Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é continuamente diferenciável e que  $p \in \mathbb{R}^n$ . Sejam  $\nabla f(x)$  e  $\nabla^2 f(x)$  o gradiente e a matriz Hessiana desta função, respectivamente. Então

$$f(x + p) = f(x) + \nabla f(x + tp)^T p, \quad (1.5)$$

para algum  $t \in (0, 1)$ . Entretanto se  $f$  é duas vezes diferenciável, tem-se

$$\nabla f(x + p) = \nabla f(x) + \int_0^1 \nabla^2 f(x + tp) p dt, \quad (1.6)$$

e que

$$f(x + p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x + tp) p, \quad (1.7)$$

para algum  $t \in (0, 1)$ .

A demonstração deste teorema (que é a junção do Teorema do Valor Médio com uma variação do Teorema Fundamental do Cálculo) pode ser encontrada em vários textos de literatura, por exemplo, Lima (2006).

**Definição 1.6** Seja  $H$  uma matriz simétrica  $n \times n$ . Os vetores  $d_1, \dots, d_n$  são  $H$ -conjugados ou simplesmente conjugados se eles são linearmente independentes e se  $d_i^T H d_j = 0$  para  $i \neq j$ .

O seguinte teorema fornece uma condição necessária para que um direção  $d$  seja uma direção de descida.

**Teorema 1.2** *Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  seja diferenciável em  $\bar{x}$ . Se existe um vetor  $d$  tal que  $\nabla f(\bar{x})^T d < 0$ , então existe um  $\delta > 0$  tal que  $f(\bar{x} + \alpha d) < f(\bar{x})$  para cada  $\alpha \in (0, \delta)$ , ou seja,  $d$  é uma direção de descida de  $f$  para  $\bar{x}$ .*

**Demonstração:** Pela diferenciabilidade de  $f$  em  $\bar{x}$ , tem-se

$$f(\bar{x} + \alpha d) = f(\bar{x}) + \alpha \nabla f(\bar{x})^T d + \alpha \|d\| \beta(\bar{x}, \alpha d), \quad (1.8)$$

sendo que  $\beta(\bar{x}, \alpha d) \rightarrow 0$  quando  $\alpha \rightarrow 0$ . Re-agrupando os termos e dividindo por  $\alpha$ ,  $\alpha \neq 0$ , tem-se

$$\frac{f(\bar{x} + \alpha d) - f(\bar{x})}{\alpha} = \nabla f(\bar{x})^T d + \|d\| \beta(\bar{x}, \alpha d). \quad (1.9)$$

Já que  $\nabla f(\bar{x})^T d < 0$  e  $\beta(\bar{x}, \alpha d) \rightarrow 0$  quando  $\alpha \rightarrow 0$ , existe um  $\delta > 0$  tal que  $\nabla f(\bar{x})^T d + \|d\| \beta(\bar{x}, \alpha d) < 0$  para todo  $\alpha \in (0, \delta)$ . O que conclui a demonstração.  $\square$

# Capítulo 2

## Fundamentos da Otimização Irrestrita

Seja um problema de otimização que depende de variáveis reais não sujeitas a restrição:

$$\min f(x), \quad x \in \mathbb{R}^n. \quad (2.1)$$

O otimizador *global* de  $f$  é um ponto que corresponde ao menor valor, ou seja, em uma definição formal tem-se:

**Definição 2.1** *Um ponto  $x^* \in \mathbb{R}^n$  é mínimo global se  $f(x^*) \leq f(x)$  para todo  $x \in \mathbb{R}^n$  ou no mínimo ao domínio de interesse.*

Afirmar que  $x^*$  trata-se de um mínimo global não é fácil, pois muitas vezes temos um conhecimento apenas local de  $f(x)$ . Muitos algoritmos são capazes de obter o mínimo local, que corresponde ao ponto de menor valor de  $f$  em uma dada vizinhança. Formalmente, pode-se dizer:

**Definição 2.2** *Um ponto  $x^* \in \mathbb{R}^n$  é mínimo local se existe uma vizinhança  $V$  de  $x^*$  tal que  $f(x^*) \leq f(x)$  para todo  $x \in V$ .*

Observe que uma vizinhança  $V$  de  $x^*$  é simplesmente um conjunto aberto que contém  $x^*$ .

O ponto que satisfaz esta definição é denominado *mínimo local fraco*. Um mínimo local estrito (ou forte) é definido como:

**Definição 2.3** *Um ponto  $x^* \in \mathbb{R}^n$  é mínimo local estrito (ou forte) se existe uma vizinhança  $V$  de  $x^*$  tal que  $f(x^*) < f(x)$  para todo  $x \in V$ , com  $x \neq x^*$ .*

A Figura 1.2 apresentou uma situação difícil para obter o mínimo global, pois o algoritmo pode ficar “preso” nos mínimos locais. Normalmente, problemas com funções potenciais possuem muitos mínimos locais.

As condições necessárias para existência do ponto ótimo (optimality condition) são obtidas assumindo que  $x^*$  é mínimo local e que são conhecidos o gradiente da função  $\nabla f(x^*)$  e sua matriz Hessiana  $\nabla^2 f(x^*)$ ; conforme será apresentado nos teoremas abaixo.

**Teorema 2.1** *(Condição necessária de 1ª ordem): Se  $x^*$  é um mínimo local e  $f$  é continuamente diferenciável em uma vizinhança aberta de  $x^*$ , então  $\nabla f(x^*) = 0$ .*

**Demonstração** Suponha por contradição que  $\nabla f(x^*) \neq 0$ . Defina o vetor  $p = -\nabla f(x^*)$ . Observe que  $p^T \nabla f(x^*) = -\|\nabla f(x^*)\|^2 < 0$ . Como  $\nabla f$  é contínuo próximo de  $x^*$ , então existe um escalar  $t'$  tal que  $p^T \nabla f(x^* + tp) < 0$ , para todo  $t \in (0, t')$ .

Para algum  $t'' \in (0, t')$ , pelo Teorema de Taylor tem-se que  $f(x^* + t''p) = f(x^*) + t'' p^T \nabla f(x^* + tp)$ , para algum  $t \in (0, t'')$ . Consequentemente,  $f(x^* + t''p) < f(x^*)$  para todo  $t'' \in (0, t')$ . Assim, foi encontrada uma direção ao longo da qual partindo de  $x^*$  a função  $f$  decresce. Como  $x^*$  é um mínimo local, tem-se uma contradição.  $\square$

O ponto  $x^*$  é chamado *ponto estacionário* de  $f$  se  $\nabla f(x^*) = 0$ . De acordo com o Teorema 2.1, um mínimo local tem que ser um ponto estacionário.

**Teorema 2.2** (*Condição necessária de 2ª ordem*): Se  $x^*$  é um mínimo local de  $f$  e  $\nabla^2 f$  existe e é contínua em uma vizinhança aberta de  $x^*$ , então  $\nabla f(x^*) = 0$  e  $\nabla^2 f(x^*)$  é semi positiva definida.

**Demonstração** Do Teorema 2.1 sabe-se que  $\nabla f(x^*) = 0$ . Por contradição, suponha que  $\nabla^2 f(x^*)$  é não positiva definida. Então pode-se escolher um vetor  $p$  tal que  $p^T \nabla^2 f(x^*) p < 0$ , e como  $\nabla^2 f$  é contínua próximo de  $x^*$ , existe um escalar  $t' > 0$  tal que  $p^T \nabla^2 f(x^* + tp) p < 0$  para todo  $t \in (0, t')$ .

Fazendo a expansão em série de Taylor em torno  $x^*$ , tem-se que para todo  $t \in (0, t')$  e algum  $t'' \in (0, t)$  que

$$f(x^* + tp) = f(x^*) + tp^T \nabla f(x^*) + \frac{1}{2} t^2 p^T \nabla^2 f(x^* + t''p) p < f(x^*).$$

Assim, como no Teorema 2.1, encontra-se uma direção na qual partindo de  $x^*$  a função  $f$  está decrescendo, e novamente, como o  $x^*$  é um mínimo local, absurdo. Logo,  $\nabla^2 f(x^*)$  é positiva definida.  $\square$

**Teorema 2.3** (*Condição suficiente de 2ª ordem*): Suponha que  $\nabla^2 f$  é contínua em uma vizinhança aberta de  $x^*$  e que  $\nabla f(x^*) = 0$  e  $\nabla^2 f(x^*)$  é positiva definida. Então  $x^*$  é um mínimo local estrito de  $f$ .

**Demonstração** Como a Hessiana  $\nabla^2 f$  é contínua positiva definida em  $x^*$ , pode-se escolher um raio  $r > 0$  tal que  $\nabla^2 f(x)$  permaneça positiva definida para todo  $x$  na bola aberta  $S = \{z \in \mathbb{R}^n / \|z - x^*\| < r\}$ .

Tomando algum vetor não nulo  $p$  com  $\|p\| < r$ , tem-se que  $x^* + p \in S$  e então

$$f(x^* + p) = f(x^*) + p^T \nabla f(x^*) + \frac{1}{2} p^T \nabla^2 f(z) p > f(x^*) + \frac{1}{2} p^T \nabla^2 f(z) p,$$

onde  $z = x^* + tp$  para algum  $t \in (0, 1)$ . Como  $z \in S$ , então  $p^T \nabla^2 f(z) p > 0$ , e então  $f(x^* + p) > f(x^*)$ , provando o resultado.  $\square$

Observe que esta é uma condição suficiente, mas não necessária. O seguinte exemplo ilustra este fato: tome  $f(x) = x_1^4 + x_2^2$ . Logo,  $x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  e  $\nabla f(x) = \begin{pmatrix} 4x_1^3 \\ 2x_2 \end{pmatrix}$ , donde  $\nabla f(x^*) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . Tem-se também que  $\nabla^2 f(x) = \begin{pmatrix} 12x_1 & 0 \\ 0 & 2 \end{pmatrix}$ , assim,  $\nabla^2 f(x^*) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$ . Então,  $\nabla^2 f(x^*)$  não é positiva definida, mas  $x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  é um ponto de mínimo global da função  $f(x) = x_1^4 + x_2^2$ .

Quando a função é convexa, o mínimo global e local são fáceis de serem caracterizados, conforme teorema a seguir:

**Teorema 2.4 :** Quando  $f$  é convexa, todo mínimo local  $x^*$  é um mínimo global de  $f$ . Se, além disso,  $f$  é diferenciável, então todo ponto estacionário  $x^*$  é um mínimo global de  $f$ .

**Demonstração** Suponha que  $x^*$  é um mínimo local mas não um mínimo global. Então podemos encontrar um ponto  $z \in \mathbb{R}^n$  tal que  $f(z) < f(x^*)$ . Considere um seguimento de reta que une  $x^*$  a  $z$ , isto é,

$$x = \gamma z + (1 - \gamma)x^*, \quad \gamma \in [0, 1]. \quad (2.2)$$

Pela propriedade de convexidade de  $f$ , temos

$$f(x) \leq \gamma f(z) + (1 - \gamma)f(x^*) \leq f(x^*). \quad (2.3)$$

Alguna vizinhança  $V$  de  $x^*$  contem um pedaço de segmento de reta definido em (2.2). Assim sempre existirá pontos  $x \in V$  tal que (2.3) seja satisfeita, o que contradiz o fato de  $x^*$  ser um mínimo local.

Para a segunda parte do teorema, suponha que  $x^*$  é um ponto estacionário mas não é um mínimo global e escolha  $z$  como acima. Então, da convexidade, tem-se que

$$\begin{aligned} \nabla f(x^*)^T(z - x^*) &= \frac{d}{d\gamma} f(x^* + \gamma(z - x^*))|_{\gamma=0} \\ &= \lim_{\gamma \rightarrow 0} \frac{f(x^* + \gamma(z - x^*)) - f(x^*)}{\gamma} \leq \lim_{\gamma \rightarrow 0} \frac{\gamma f(z) + (1 - \gamma)f(x^*) - f(x^*)}{\gamma} = f(z) - f(x^*) < 0. \end{aligned}$$

Portanto,  $\nabla f(x^*) \neq 0$ , o que contradiz o fato de  $x^*$  ser um ponto estacionário.  $\square$

## 2.1 Algoritmo para a direção de busca

Nos últimos 40 anos foram desenvolvidos vários algoritmos poderosos para a otimização ir-restrita de funções suaves (smooth function). Nesta seção serão apresentadas as características principais de alguns destes métodos, que serão estudados detalhadamente nos próximos capítulos.

Todos os algoritmos necessitam que o usuário forneça o *ponto inicial*,  $x^0$ . Se o usuário tiver um bom conhecimento do problema em estudo, pode escolher  $x^0$  como uma boa estimativa da solução procurada. À partir de  $x^0$ , o algoritmo de otimização gerará uma *sequência de aproximações*  $\{x^k\}_{k=0, \dots, \infty}$ . Este processo termina quando o algoritmo não consegue progredir (melhorar a solução) ou quando a aproximação  $x^k$  atende uma precisão considerada suficiente.

Para decidir como mover de uma iteração para a próxima, o algoritmo usa informações da função  $f(x^k)$  e sendo possível também usa informação da função em pontos anteriores ( $x^0, x^1, \dots, x^{k-1}$ ). O objetivo é encontrar a aproximação  $x^{k+1}$  que corresponda a um valor da função *inferior* a  $f(x^k)$ , ou seja,  $f(x^{k+1}) < f(x^k)$ . Existem duas estratégias fundamentais para o algoritmo mover de  $x^k$  para  $x^{k+1}$ . Muitos dos algoritmos atuais seguem algumas dessas estratégias:

- *Primeira estratégia - Busca linear (linear search)*: o algoritmo escolhe uma direção de busca  $p_k$  e saindo de  $x^k$  ao longo desta direção busca  $x^{k+1}$ , de forma que o valor da função diminua.

A distância que deve ser percorrida ao longo da direção  $p_k$ , representada por  $\alpha$ , pode ser calculada de forma aproximada resolvendo o seguinte problema de otimização unidimensional:

$$\min f(x^k + \alpha p_k), \quad \alpha > 0. \quad (2.4)$$

A resolução de (2.4) permite encontrar o comprimento do passo  $\alpha$ . A solução exata de (2.4) permite obter a minimização máxima na direção  $p_k$ . Mas, este processo (exato)

pode ser caro e normalmente desnecessário. Em vez disso, o algoritmo de busca linear gera um número limitado de tentativas para o comprimento  $\alpha$ , até encontrar um valor que aproxime o mínimo.

A partir do novo ponto, uma nova direção de busca e um novo comprimento são calculados e o processo se repete (vide Figura 2.1).

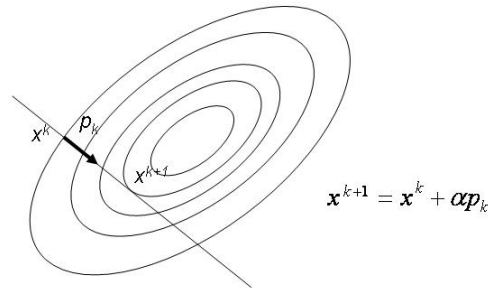


Figura 2.1: Algoritmo de busca linear

- *Segunda estratégia - Região de confiança (Trust region)*: a informação colhida acerca de  $f$  é usada para construir uma *função modelo*  $m_k$ , cujo comportamento próximo ao ponto atual  $x^k$  é similar àquele da função objetivo atual. Como o modelo  $m_k$  pode não ser uma boa aproximação de  $f$  quando  $x$  estiver longe de  $x^k$ , a busca pela minimização de  $m_k$  é feita em uma região em torno de  $x^k$ . Em outras palavras, o passo  $p$  pode ser aproximado resolvendo o seguinte sub-problema:

$$\min_p m_k(x^k + p) \quad (2.5)$$

onde  $x^k + p$  são direções contidas na região de confiança.

Se a solução “candidata” não resultar em um decréscimo suficiente de  $f$ , pode-se concluir que a região de confiança está muito grande, ela deve ser encolhida e o problema (2.5) deve ser novamente resolvido.

Normalmente, a região de confiança é definida por  $\|p\|_2 \leq \delta$ , onde o escalar  $\delta > 0$  é denominado raio da região de confiança. Regiões elípticas e de forma quadrada também podem ser adotadas.

O modelo usado em (2.5), é usualmente definido por uma função quadrática da forma

$$m_k(x^k + p) = f^k + p^T \nabla f^k + \frac{1}{2} p^T B^k p, \quad (2.6)$$

onde  $f^k, \nabla f^k, B^k$  são escalar, vetor e matriz, respectivamente, sendo:  $f^k$  a função no ponto  $x^k$ ,  $\nabla f^k$  o valor do gradiente no ponto  $x^k$ ,  $B^k$  a matriz Hessiana  $\nabla^2 f^k$  ou uma aproximação desta no ponto  $x^k$ .

Note que se  $B^k = 0$  e a região de confiança é definida usando a norma euclidiana, então o sub-problema definido em (2.5) pode ser escrito como:

$$\min_p f^k + p^T \nabla f^k, \quad \text{sujeito a } \|p\|_2 \leq \delta_k \quad (2.7)$$

e a solução deste problema pode ser escrita como

$$p_k = - \frac{\delta_k \nabla f^k}{\|\nabla f^k\|}. \quad (2.8)$$

Nos capítulos posteriores será visto que, neste caso, trata-se do método da descida máxima no qual o comprimento do passo é determinado pelo raio da região de confiança. Assim, estes dois métodos são essencialmente os mesmos, para o caso específico.

Considere o seguinte exemplo:  $f(x) = 10(x_2 - x_1^2)^2 + (1 - x_1)^2$ , cujo gráfico e curvas de nível estão representados na Figura 2.2. Seja o ponto  $x^k = (0, 1)^T$  para o qual  $f^k = 11$ ,

$$\nabla f^k = \begin{bmatrix} -2 \\ 20 \end{bmatrix},$$

$$B^k = \begin{bmatrix} -38 & 0 \\ 0 & 20 \end{bmatrix},$$

$m_k = 11 + p \begin{pmatrix} -2 \\ 20 \end{pmatrix} + \frac{1}{2} p^T \begin{pmatrix} -38 & 0 \\ 0 & 20 \end{pmatrix} p$ . Considerando  $p = \begin{pmatrix} -2 \\ 20 \end{pmatrix}$  tem-se  $m_k = 11$ .

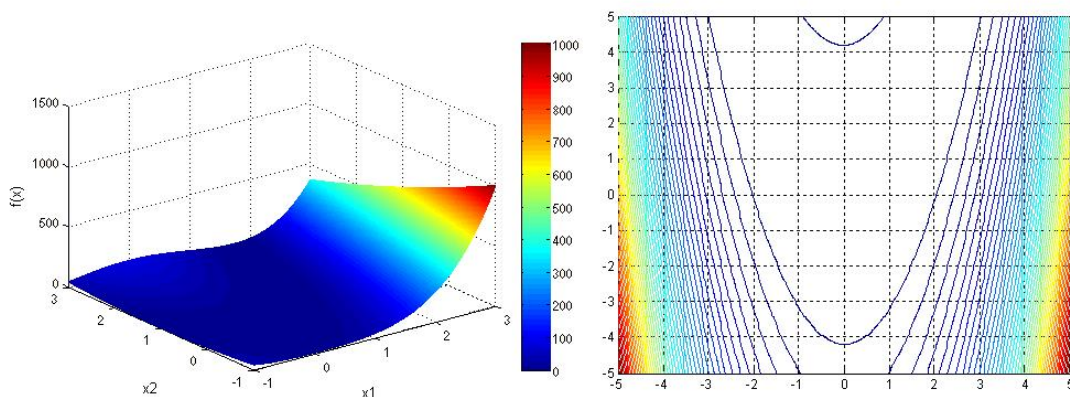


Figura 2.2: Gráfico e curvas de nível da função  $f(x) = 10(x_2 - x_1^2)^2 + (1 - x_1)^2$

Observe na Figura 2.3 que a estratégia da região de confiança pode trabalhar com diferentes direções à partir do ponto  $x^k$ . Assim, difere da busca linear, uma vez que esta trabalha com uma única direção de busca.

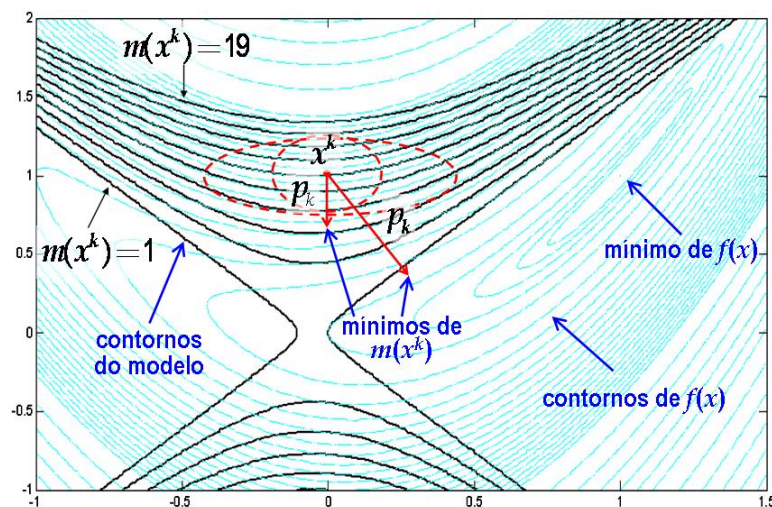


Figura 2.3: Região de confiança para  $f(x) = 10(x_2 - x_1^2)^2 + (1 - x_1)^2$

**Resumindo:** A *busca linear* inicia com uma direção de busca fixa  $p_k$  e procura calcular a distância apropriada (passo)  $\alpha_k$  para se mover nesta direção. Na *região de confiança*,

escolhe-se uma distância máxima (raio de região de confiança  $\delta_k$ ) e então procura-se a direção e o passo de forma a atingir o melhor progresso (minimizar  $f^k$ ) sujeito a uma restrição de distância. Se a solução não se mostrar satisfatória, o raio  $\delta_k$  é reduzido e faz-se nova tentativa.

# Capítulo 3

## Escolha do passo nos Métodos de Busca Linear

Nos métodos de busca linear, a cada iteração é calculado a direção de busca  $p_k$  e então se decide o quanto se deve mover ao longo desta direção. A iteração é dada por:

$$x^{k+1} = x^k + \alpha_k p_k, \quad (3.1)$$

sendo que  $\alpha_k$  é um escalar positivo chamado de *tamanho do passo ou comprimento de passo*. O sucesso do método de busca linear depende de uma escolha eficaz tanto da direção  $p_k$  quanto do tamanho do passo  $\alpha_k$ .

A maioria dos algoritmos requerem que  $p_k$  seja uma direção de descida, isto é, que  $p_k^T \nabla f^k < 0$ , pois esta propriedade garante que a função  $f$  possa ser reduzida ao longo desta direção. Além disso, a direção de busca frequentemente tem a forma

$$p_k = -(B^k)^{-1} \nabla f^k, \quad (3.2)$$

sendo que  $B^k$  é uma matriz simétrica e não-singular. Será visto, nos capítulos seguintes, que no método de descida máxima  $B^k$  é simplesmente a matriz identidade  $I$ . Para o método de Newton  $B^k$  é a matriz Hessiana  $\nabla^2 f(x^k)$ . Nos métodos Quase-Newton  $B^k$  é uma aproximação da Hessiana que deve ser atualizada em toda iteração. Considerando  $p_k$  definido por (3.2) e  $B^k$  simétrica e positiva definida, tem-se

$$p_k^T \nabla f^k = -(\nabla f^k)^T (B^k)^{-1} \nabla f^k < 0, \quad (3.3)$$

e portanto  $p_k$  é uma direção de descida.

Neste capítulo, serão discutidas as direções de busca e também a escolha do comprimento do passo  $\alpha_k$  ao longo da direção de busca  $p_k$  para que o método seja convergente partindo de um ponto inicial qualquer.

### 3.1 Considerações sobre a escolha do comprimento do passo

No cálculo do passo  $\alpha_k$ , deve-se fazer uma negociação, pois a escolha de  $\alpha_k$  que resulte em uma substancial redução de  $f$ , pode custar muito em termos de tempo. A escolha ideal é um mínimo global da função unidimensional definida por

$$\phi(\alpha) = f(x^k + \alpha p_k), \quad \alpha > 0. \quad (3.4)$$

Uma ilustração do mínimo global de uma função qualquer em  $\alpha$  é dada na Figura 3.1.

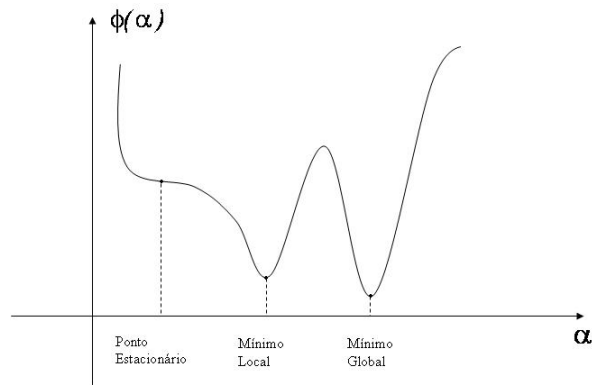


Figura 3.1: Comprimento de passo ideal dado pelo mínimo global

Geralmente é muito caro identificar o passo ideal, pois mesmo para encontrar um mínimo local, usando uma precisão moderada, são requeridas muitas avaliações da função objetivo  $f$  e de seu gradiente  $\nabla f$ . A estratégia prática é uma busca inexata, identificando um comprimento de passo que reduz  $f$ . Normalmente, em uma busca típica, obtém-se uma sequência de valores para  $\alpha$ , finalizando quando alguma condição de parada é satisfeita. A busca é feita em duas fases:

- a primeira fase encontra um intervalo que contém o passo desejado;
- a segunda fase calcula o melhor passo neste intervalo.

Uma simples condição que pode ser imposta em  $\alpha_k$  é requerer a redução de  $f$ , ou seja,

$$f(x^k + \alpha_k p_k) < f(x^k). \quad (3.5)$$

Mas esta condição simples não é suficiente para garantir a convergência para o ótimo  $x^*$ , conforme citado por Nocedal *et al.* (2000) e ilustrado na Figura 3.2. Neste caso, o mínimo da função é  $f(x^*) = -1$ . Observe que tem-se uma redução de  $f$  a cada iteração, mas a redução é insuficiente para garantir que a sequência de iterações convirja rapidamente para o ponto ótimo. Para evitar este comportamento será necessária uma condição de decréscimo suficiente, conforme conceitos que serão vistos a seguir.

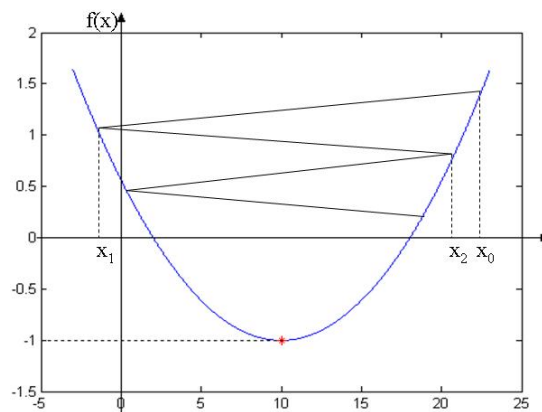


Figura 3.2: Ilustração de que o decréscimo de uma função não garante convergência rápida

## 3.2 Condições de Wolfe

Uma condição popular de busca inexata estipula que  $\alpha_k$  será o primeiro valor que obedecer a um decrescimento suficiente da função objetivo, sendo medida pela seguinte inequação:

$$f(x^k + \alpha p_k) \leq f(x^k) + c_1 \alpha (\nabla f^k)^T p_k \quad (3.6)$$

para alguma constante  $c_1 \in (0, 1)$ . Em outras palavras, a redução de  $f$  deve ser proporcional tanto ao comprimento do passo  $\alpha_k$  quanto à direção  $(\nabla f^k)^T p_k$ . A inequação (3.6) é também chamada de *Condição de Armijo*, sendo ilustrada na Figura 3.3.

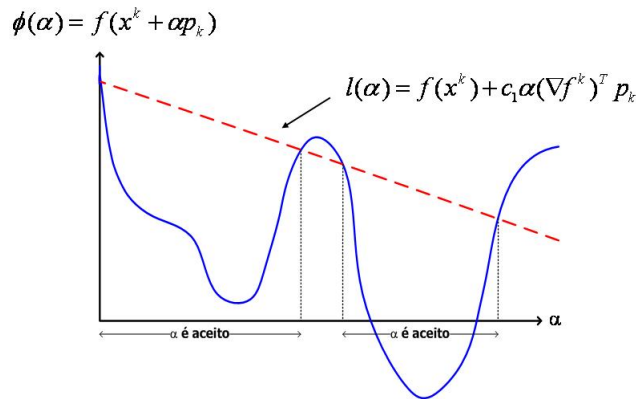


Figura 3.3: Condição de decrescimento suficiente

Observe que  $l(\alpha) = f(x^k) + c_1 \alpha (\nabla f^k)^T p_k$  é uma função linear, que possui inclinação negativa, já que  $(\nabla f^k)^T p_k < 0$  (direção de descida) e  $c_1 \in (0, 1)$ . A condição de decrescimento suficiente estabelece que  $\alpha$  é aceitável somente se  $\phi(\alpha) \leq l(\alpha)$ , conforme os intervalos mostrados na figura. Na prática,  $c_1$  é escolhido muito pequeno, em torno  $c_1 = 10^{-4}$ .

A condição (3.6) não é suficiente, por ela mesma, para assegurar que o algoritmo tenha um progresso razoável. Desta forma, é introduzida um novo requisito, denominado *Condição de Curvatura*, segundo o qual  $\alpha$  deve satisfazer

$$\nabla f(x^k + \alpha_k p_k)^T p_k \geq c_2 (\nabla f^k)^T p_k \quad (3.7)$$

para algum  $c_2 \in (c_1, 1)$ . Observe que o lado esquerdo de (3.7) é simplesmente a derivada  $\phi'(\alpha_k)$ , desta forma, a condição (3.7) garante que a inclinação de  $\phi$  em  $\alpha_k$  é maior que  $c_2$  vezes a inclinação inicial  $\phi'(0)$ . Isto faz sentido se a inclinação  $\phi'(\alpha)$  é fortemente negativa, pois isto indica que é possível reduzir significativamente  $f$  movendo-se ao longo desta direção. Por outro lado, se  $\phi'(\alpha)$  é levemente negativa ou mesmo positiva, isto significa que não se pode esperar um decrescimento importante de  $f$  nesta direção, sendo melhor parar a busca linear. Normalmente, adota-se  $c_2 = 0,9$  para os métodos de Newton e Quase-Newton e  $c_2 = 0,1$  para o método do gradiente conjugado. Para uma melhor compreensão observe a Figura 3.4 que ilustra a condição de curvatura.

As condições de decrescimento suficiente e de curvatura são conhecidas coletivamente como *condições de Wolfe*, assim:

$$f(x^k + \alpha_k p_k) \leq f(x^k) + c_1 \alpha_k (\nabla f^k)^T p_k \quad (3.8)$$

e

$$\nabla f(x^k + \alpha_k p_k)^T p_k \geq c_2 (\nabla f^k)^T p_k \quad (3.9)$$

para  $0 < c_1 < c_2 < 1$ .

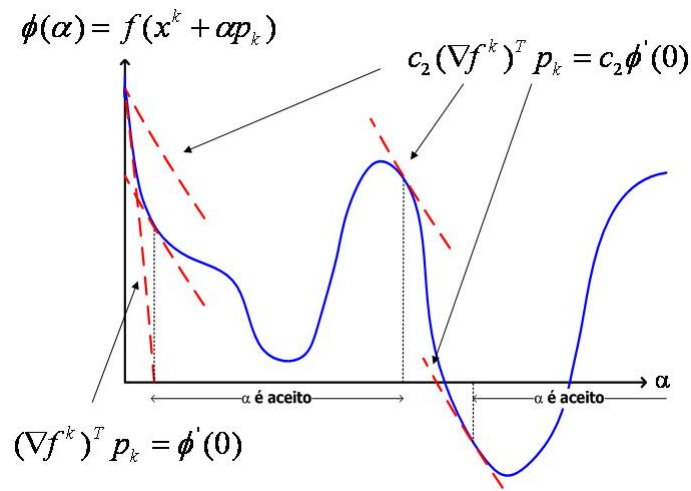


Figura 3.4: Condição de curvatura

O comprimento do passo  $\alpha_k$  pode satisfazer a condição de Wolfe sem estar em uma região particularmente próxima do mínimo de  $\phi$ , como mostrado na Figura 3.5.

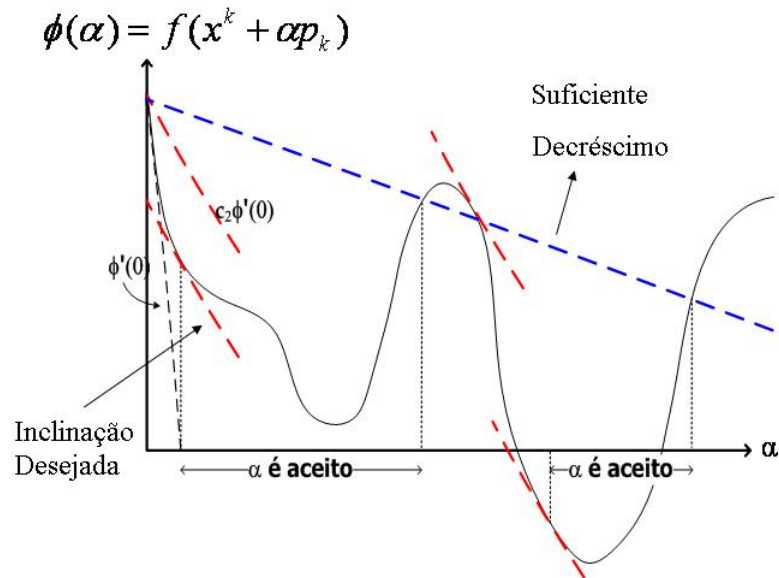


Figura 3.5: Condições de Wolfe

A condição de curvatura pode ser modificada de forma a forçar  $\alpha_k$  a permanecer na região próxima ao mínimo local ou pelo menos de um ponto crítico de  $\phi$ .

Assim na *condição forte de Wolfe (strong condition)* o passo  $\alpha_k$  deve satisfazer

$$f(x^k + \alpha_k p_k) \leq f(x^k) + c_1 \alpha_k (\nabla f^k)^T p_k \quad (3.10)$$

e

$$|\nabla f(x^k + \alpha_k p_k)^T p_k| \leq c_2 |(\nabla f^k)^T p_k| \quad (3.11)$$

com  $0 < c_1 < c_2 < 1$ .

Desta forma, exclui-se os pontos que estão longe dos pontos críticos de  $\phi$ . Esta situação está ilustrada na Figura 3.6.

**Lema 3.1** *Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  seja continuamente diferenciável. Seja  $p_k$  uma direção de descida em  $x^k$ , e assume-se que  $f$  seja limitada ao longo do raio  $\{x^k + \alpha p_k : \alpha > 0\}$ . Então se*

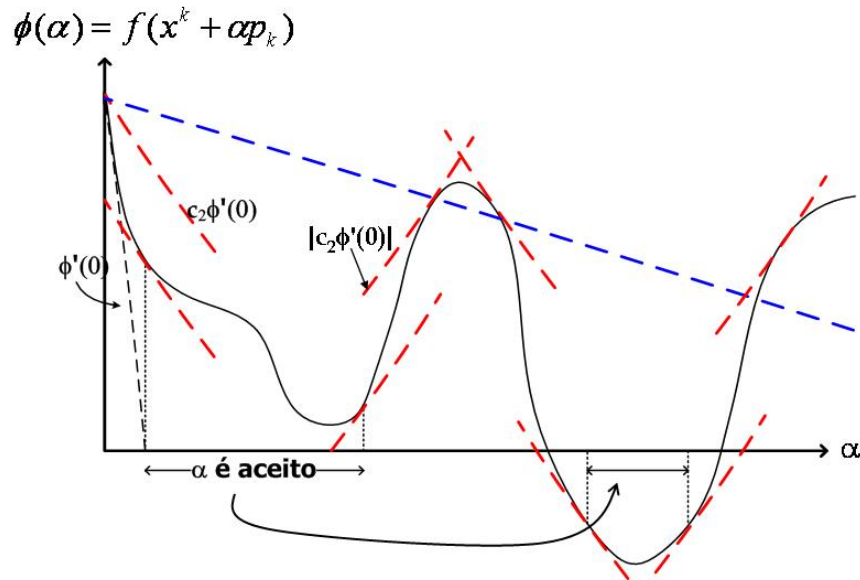


Figura 3.6: Condições forte de Wolfe

$0 < c_1 < c_2 < 1$ , existe um intervalo para o qual o comprimento de passo  $\alpha$  satisfaz a condição de Wolfe (3.8) e (3.9) e a condição forte de Wolfe (3.10) e (3.11).

**Demonstração:** Note que  $\phi(\alpha) = f(x^k + \alpha p_k)$  é limitada para todo  $\alpha > 0$ . Uma vez que  $0 < c_1 < 1$ , a reta  $l(\alpha) = f(x^k) + \alpha' c_1 (\nabla f^k)^T p_k$  não é limitada e deve portanto interceptar o gráfico de  $\phi$  ao menos uma vez. Seja  $\alpha' > 0$  o menor valor interceptador de  $\alpha$ , isto é,

$$f(x^k + \alpha' p_k) = f(x^k) + \alpha' c_1 (\nabla f^k)^T p_k. \quad (3.12)$$

A condição de decrescimento suficiente (3.8) obviamente é válida para todos os comprimentos de passo menores que  $\alpha'$ . Usando o Teorema do Valor Médio, existe  $\alpha'' \in (0, \alpha')$  tal que

$$f(x^k + \alpha' p_k) - f(x^k) = \alpha' \nabla f(x^k + \alpha'' p_k)^T p_k. \quad (3.13)$$

Comparando (3.12) e (3.13), obtém-se

$$\nabla f(x^k + \alpha'' p_k)^T p_k = c_1 \nabla f(x^k)^T p_k > c_2 (\nabla f^k)^T p_k \quad (3.14)$$

uma vez que  $c_1 < c_2$  e  $(\nabla f^k)^T p_k < 0$ . Desta forma,  $\alpha''$  satisfaz a condição de Wolfe tanto para (3.8) quanto (3.9). Assim, pela condição de suavidade de  $f$ , existe um intervalo ao redor de  $\alpha''$  para o qual a condição de Wolfe é obedecida. Além disso, uma vez que o termo do lado direito de (3.14) é negativo, a condição forte de Wolfe é válida no mesmo intervalo.  $\square$

**Observação** A condição de Wolfe é um escalar invariante no seguinte sentido: multiplicando a função objetivo por uma constante ou fazendo pequenos ajustes no valor das variáveis, a função não é alterada.

### 3.3 Busca unidimensional usando o Método da Seção Áurea

Definido o processo iterativo  $x^{k+1} = x^k + \alpha_k p_k$ , a otimização de uma função de várias variáveis recai a cada iteração na busca unidimensional do ponto  $\alpha_k$  que minimiza a função

$$\alpha_k = \min \phi(\alpha) = \min f(x^k + \alpha p_k). \quad (3.15)$$

O problema é definido em um intervalo de incerteza de busca  $I_b = [a_0, b_0]$ . Assumindo que  $\phi(\alpha)$  é unimodal, então  $\phi(\alpha)$  é decrescente em  $[a_0, \alpha_k]$  e crescente em  $[\alpha_k, b_0]$ .

**Definição 3.1** Seja  $\alpha_k$  o ponto mínimo em  $[a_0, b_0]$  de uma função unimodal  $\phi(\alpha)$ . Sejam  $c_0, d_0 \in I_b$ , sendo  $c_0 < d_0$ , define-se:

**Caso 1** se  $\phi(c_0) > \phi(d_0)$ , então  $\alpha_k \in [c_0, b_0]$ ;

**Caso 2** se  $\phi(c_0) < \phi(d_0)$ , então  $\alpha_k \in [a_0, d_0]$ ;

**Caso 3** se  $\phi(c_0) = \phi(d_0)$ , então  $\alpha_k \in [c_0, d_0]$ .

Uma ilustração de cada caso acima é dada nas Figuras 3.7, 3.8 e 3.9, respectivamente. A aplicação iterativa desta definição conduz a intervalos cada vez menores. Além disso, deseja-se que o algoritmo acelere a redução do intervalo e faça o menor número de avaliações de  $\phi(\alpha)$ .

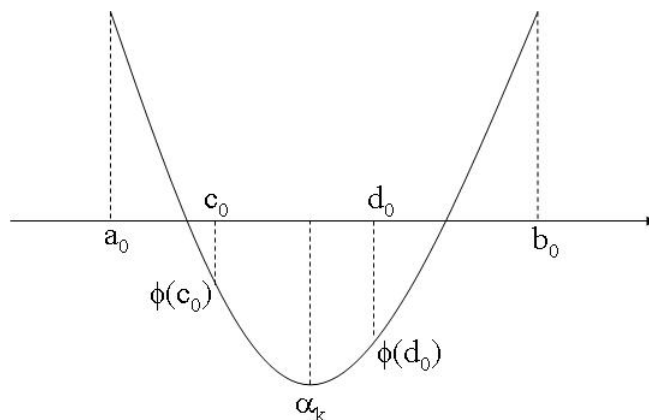


Figura 3.7: *Caso 1* do método da Seção Áurea:  $\phi(c_0) > \phi(d_0) \Rightarrow \alpha_k \in [c_0, b_0]$

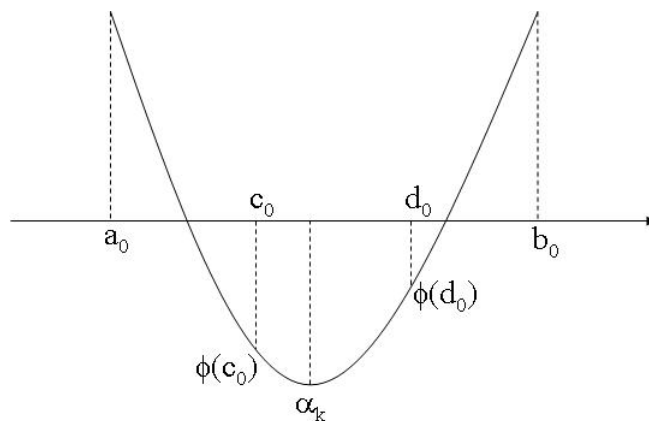


Figura 3.8: *Caso 2* do método da Seção Áurea:  $\phi(d_0) > \phi(c_0) \Rightarrow \alpha_k \in [a_0, d_0]$

Considerando a  $s$ -ésima iteração da aplicação do Método da Seção Áurea, tem-se que os requisitos para a obtenção deste algoritmo são:

1. Que cada novo intervalo  $[a_{s+1}, b_{s+1}]$  seja uma fração fixa do intervalo anterior

$$\begin{aligned} I_{s+1} &= \tau I_s \\ \Leftrightarrow b_s - c_s &= d_s - a_s = \tau(b_s - a_s). \end{aligned} \quad (3.16)$$

De (3.16), vem

$$\begin{aligned} d_s - a_s &= \tau(b_s - a_s) \\ \Rightarrow d_s &= a_s + \tau(b_s - a_s). \end{aligned} \quad (3.17)$$

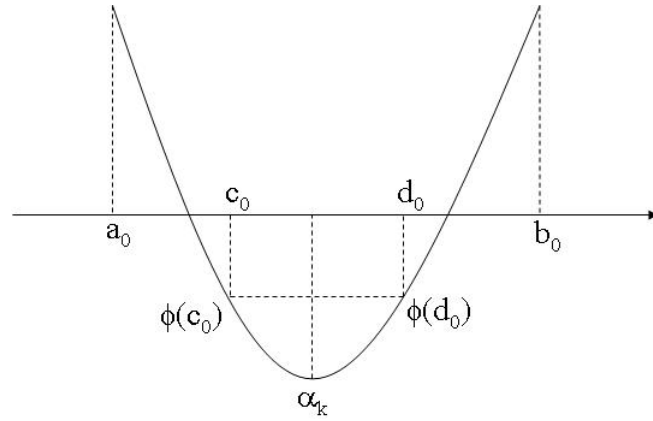


Figura 3.9: *Caso 3* do método da Seção Áurea:  $\phi(c_0) = \phi(d_0) \Rightarrow \alpha_k \in [c_0, d_0]$

De (3.16), vem também

$$\begin{aligned}
 b_s - c_s &= \tau(b_s - a_s) \\
 \Rightarrow c_s &= b_s - \tau(b_s - a_s) + a_s - a_s \\
 \Rightarrow c_s &= (b_s - a_s) - \tau(b_s - a_s) + a_s \\
 \Rightarrow c_s &= a_s + (1 - \tau)(b_s - a_s)
 \end{aligned} \tag{3.18}$$

2. As condições da definição devem ser obedecidas:

(a) Se  $\phi(c_s) \geq \phi(d_s)$ , então  $[a_{s+1}, b_{s+1}] = [c_s, b_s]$ , como na Figura 3.10. Assim,

$$\begin{aligned}
 a_{s+1} &= c_s, \\
 b_{s+1} &= b_s, \\
 c_{s+1} &= d_s.
 \end{aligned} \tag{3.19}$$

Usando a equação (3.17) tem-se:

$$d_{s+1} = a_{s+1} + \tau(b_{s+1} - a_{s+1}). \tag{3.20}$$

Usando a equação (3.18) pode-se escrever  $c_{s+1} = a_{s+1} + (1 - \tau)(b_{s+1} - a_{s+1})$ , e usando (3.19) vem que  $d_s = c_{s+1} = a_{s+1} + (1 - \tau)(b_{s+1} - a_{s+1})$ . Como  $a_{s+1} = c_s$  e  $b_{s+1} = b_s$ , segue que

$$d_s - c_s = (1 - \tau)(b_s - c_s). \tag{3.21}$$

De (3.17) e (3.18) vem que  $d_s - a_s = \tau(b_s - a_s)$  e  $c_s - a_s = (1 - \tau)(b_s - a_s)$ , donde

$$d_s - c_s = (2\tau - 1)(b_s - a_s). \tag{3.22}$$

Comparando (3.21) e (3.22), resulta

$$(1 - \tau)(b_s - c_s) = (2\tau - 1)(b_s - a_s). \tag{3.23}$$

Substituindo (3.16) vem

$$(1 - \tau)\tau(b_s - a_s) = (2\tau - 1)(b_s - a_s) \Rightarrow \tau^2 + \tau - 1 = 0 \Rightarrow \underline{\tau = 0,618034}. \tag{3.24}$$

(b) Se  $\phi(c_s) < \phi(d_s)$ , então  $[a_{s+1}, b_{s+1}] = [a_s, d_s]$ , como na Figura 3.11. Assim,

$$\begin{aligned}
 a_{s+1} &= a_s, \\
 b_{s+1} &= d_s, \\
 d_{s+1} &= c_s.
 \end{aligned} \tag{3.25}$$

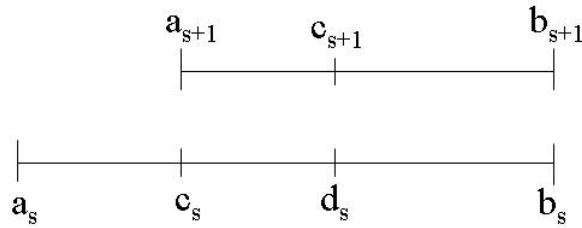


Figura 3.10: Se  $\phi(c_s) \geq \phi(d_s)$ , então  $[a_{s+1}, b_{s+1}] = [c_s, b_s]$

Usando a equação (3.18) tem-se

$$c_{s+1} = a_{s+1} + (1 - \tau)(b_{s+1} - a_{s+1}). \quad (3.26)$$

Usando (3.17) pode-se escrever  $d_{s+1} = a_{s+1} + \tau(b_{s+1} - a_{s+1})$ . De (3.25) segue que  $c_s = d_{s+1} = a_{s+1} + \tau(b_{s+1} - a_{s+1})$ . Como  $a_{s+1} = a_s$  e  $b_{s+1} = d_s$ , segue que

$$c_s - a_s = \tau(d_s - a_s). \quad (3.27)$$

Substituindo (3.18) em (3.27), resulta:

$$(1 - \tau)(b_s - a_s) = \tau(d_s - a_s) \quad (3.28)$$

Substituindo (3.17) em (3.28), obtém-se:

$$(1 - \tau)(b_s - a_s) = \tau\tau(b_s - a_s) \Rightarrow \tau^2 + \tau - 1 = 0 \Rightarrow \underline{\tau = 0,618034}. \quad (3.29)$$

De (3.24) e (3.29), verifica-se que existe uma única taxa de redução que obedece aos requisitos do teorema.

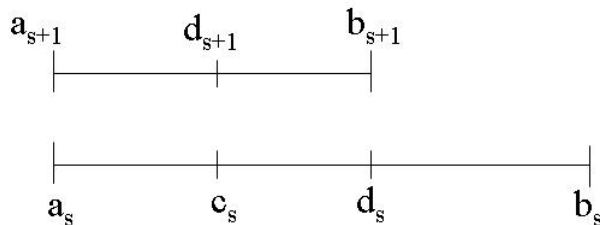


Figura 3.11: Se  $\phi(c_s) < \phi(d_s)$ , então  $[a_{s+1}, b_{s+1}] = [a_s, d_s]$

A cada iteração do Método da Seção Áurea (gold section) o intervalo é reduzido a aproximadamente 62% de seu comprimento.

O valor de  $\alpha_k$  pode ser calculado como o ponto de mínimo do polinômio  $P_n(\alpha)$  que aproxima a função unidimensional  $\phi(\alpha)$ . Normalmente, considerando os 4 pontos do último intervalo, calcula-se uma interpolação cúbica:  $P_3(\alpha) = c_1 + c_2\alpha + c_3\alpha^2 + c_4\alpha^3$ . Como  $P_3(\alpha_i) = \phi(\alpha_i)$ , resulta em

$$\begin{bmatrix} 1 & a_s & a_s^2 & a_s^3 \\ 1 & c_s & c_s^2 & c_s^3 \\ 1 & d_s & d_s^2 & d_s^3 \\ 1 & b_s & b_s^2 & b_s^3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} \phi(a_s) \\ \phi(c_s) \\ \phi(d_s) \\ \phi(b_s) \end{bmatrix}. \quad (3.30)$$

Além disso, fazendo  $\frac{dP_3}{d\alpha} = c_2 + 2c_3\alpha + 3c_4\alpha^2 = 0$ , tem-se

$$\alpha_i = \frac{-c_3 \pm \sqrt{c_3^2 - 3c_2c_4}}{3c_4}, \quad i = 1, 2.$$

Para  $\alpha_1$  e  $\alpha_2$  verifica-se

$$\frac{d^2P_3}{d\alpha^2} = 2c_3 + 6c_4\alpha_i > 0 \quad \Rightarrow \quad \alpha_i = \alpha_k.$$

Uma opção interessante do ponto de vista computacional, para um intervalo  $I_s$  bastante reduzido, é optar por um procedimento simplificado, no qual  $\alpha_k$  será o ponto do último intervalo que corresponder ao menor valor  $\phi(\alpha)$ .

Um algoritmo do Método da Seção Áurea é dado a seguir:

**Dados iniciais:** Defina o intervalo de incerteza inicial  $[a_0, b_0]$ , faça  $c_0 = a_0 + (1-\tau)(b_0 - a_0)$  e  $d_0 = a_0 + \tau(b_0 - a_0)$  e calcule  $\phi(c_0)$  e  $\phi(d_0)$ . Faça  $s = 1$  e vá para o passo 1.

**Passo 1:** Se  $b_s - a_s < \epsilon$ , pare. A solução ótima está no intervalo dado  $[a_s, b_s]$ . Caso contrário, se  $\phi(c_s) \geq \phi(d_s)$ , vá para o passo 2, se  $\phi(c_s) < \phi(d_s)$ , vá para o passo 3.

**Passo 2:** Faça  $a_{s+1} = c_s$ ,  $b_{s+1} = b_s$ ,  $c_{s+1} = d_s$  e  $d_{s+1} = a_{s+1} + \tau(b_{s+1} - a_{s+1})$ . Calcule  $\phi(d_{s+1})$  e vá para o passo 4.

**Passo 3:** Faça  $a_{s+1} = a_s$ ,  $b_{s+1} = d_s$ ,  $d_{s+1} = c_s$  e  $c_{s+1} = a_{s+1} + (1-\tau)(b_{s+1} - a_{s+1})$ . Calcule  $\phi(c_{s+1})$  e vá para o passo 4.

**Passo 4:** Faça  $s = s + 1$  e vá para o passo 1.

**Observação:** A seguir serão apresentados dois exemplos que serão utilizados para mostrar a performance de todos os métodos estudados nesta dissertação. Vale ressaltar que todos os códigos computacionais desenvolvidos neste trabalho foram implementados em Matlab®.

**EXEMPLO 1:** Considere o seguinte problema:

$$\min f(x) = (2x_1 - 3x_2)^2 + (x_2 - 1)^4. \quad (3.31)$$

Note que a solução ótima deste problema é  $x^* = (\frac{3}{2}, 1)^T$  com valor ótimo  $f(x^*) = 0$ . A figura 3.12 mostra o seu gráfico, suas curvas de nível e solução ótima.

Tem-se como justificativas para a escolha deste problema o fato da função  $f$  ter solução analítica, do valor do mínimo da função ser igual a zero, isto é,  $f(x^*) = 0$  e da matriz Hessiana de  $f$  ser facilmente invertível, pois é simétrica e positiva definida.

Dado o ponto inicial  $x^0 = (0, 2)^T$ . Calcular o tamanho do passo  $\alpha^*$  utilizando o Método da Seção Áurea para a primeira iteração. Seja  $I_0 = [0, 2]$  o intervalo inicial. Será adotado como critério de parada que o intervalo final seja reduzido a 0,5% do intervalo inicial, ou seja,  $I_b \leq 0,005I_0$ , donde o comprimento de  $I_b$  deve ser menor ou igual a 0,01. Para este exemplo será considerado direção de busca dada por  $p_k = -\nabla f(x^k)$ , assim:

$$p_k = -\nabla f(x) = \begin{pmatrix} -8x_1 + 12x_2 \\ 12x_1 - 4x_2^3 + 12x_2^2 - 30x_2 + 4 \end{pmatrix}. \quad (3.32)$$

A Tabela 3.1 mostra um resumo dos cálculos. Observe que com 11 iterações obteve-se o intervalo de confiança  $I_b = [0, 0263 ; 0, 0426]$ . Para encontrar  $\alpha^*$  será utilizado um processo

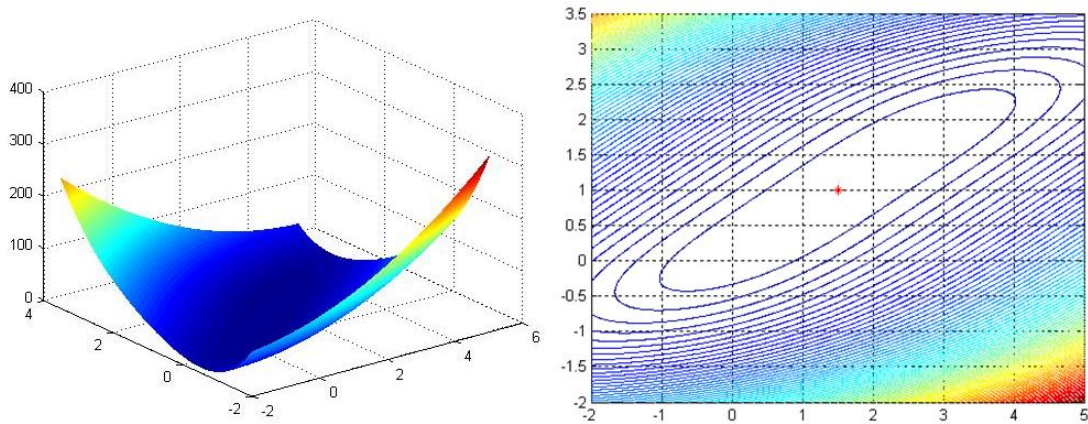


Figura 3.12: Gráfico e curvas de nível da função  $f(x) = (2x_1 - 3x_2)^2 + (x_2 - 1)^4$  do Exemplo 1

simplificado no qual calcula-se os valores da função  $\phi$  na última iteração e escolhe-se o menor deles. Assim,  $\alpha^*$  é o ponto que minimiza a função  $\phi$  aplicado na última iteração do processo, ou seja,

$$\begin{aligned} & \min\{\phi(a_{11}), \phi(c_{11}), \phi(d_{11}), \phi(b_{11})\} \\ & = \min\{\phi(0, 0263), \phi(0, 0325), \phi(0, 0364), \phi(0, 0426)\} \\ & = \min\{2, 4955; 0, 2957; 0, 0545; 1, 5716\} = 0, 0545. \end{aligned}$$

Logo,  $\alpha^* = d_{11} = 0, 0364 \in I_b = [0, 0263 ; 0, 0426]$ .

Iteração k	$a_s$	$c_s$	$d_s$	$b_s$	$\phi(c_s)$	$\phi(d_s)$	$b_s - a_s$
1	0	0,7639	1,2361	2	$7, 7820 \times 10^5$	$5, 5476 \times 10^6$	2
2	0	0,4721	0,7639	1,2361	$1, 0770 \times 10^5$	$7, 7820 \times 10^5$	1,2361
3	0	0,2918	0,4721	0,7639	$1, 4821 \times 10^4$	$1, 0770 \times 10^5$	0,7639
4	0	0,1803	0,2918	0,4721	$2, 0810 \times 10^3$	$1, 4821 \times 10^4$	0,4721
5	0	0,1115	0,1803	0,2918	$3, 0494 \times 10^2$	$2, 0810 \times 10^3$	0,2918
6	0	0,0689	0,1115	0,1803	40,5464	$3, 0494 \times 10^2$	0,1803
7	0	0,0426	0,0689	0,1115	1,5716	40,5464	0,1115
8	0	0,0263	0,0426	0,0689	2,4955	1,5716	0,0689
9	0,0263	0,0426	0,0526	0,0689	1,5716	9,5592	0,0163
10	0,0263	0,0364	0,0426	0,0526	0,0545	1,5716	0,0101
11	0,0263	0,0325	0,0364	0,0426	0,2957	0,0545	0,0062

Tabela 3.1: Método da Seção Áurea aplicado ao Exemplo 1

O processo iterativo  $x^{k+1} = x^k + \alpha_k p_k$  fornece o próximo ponto, isto é,

$$x^1 = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + 0, 0364 \begin{pmatrix} 24 \\ -40 \end{pmatrix} = \begin{pmatrix} 0, 8727 \\ 0, 5455 \end{pmatrix}. \quad (3.33)$$

A Figuras 3.13 mostra o progresso do método aplicado ao Exemplos 1 para a primeira iteração.

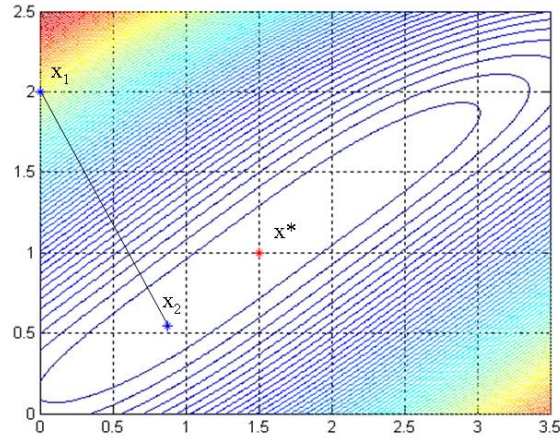


Figura 3.13: Primeira iteração do Método da Seção Áurea aplicado ao Exemplo 1

**EXEMPLO 2:** Considere vapor d'água fluindo a razão de  $N$  moles/hr à pressão de 1 atm, que deve ser comprimido à 64 atm usando um compressor de três estágios conforme Figura 3.14. Assume-se que o processo seja reversível, adiabático e que após cada estágio o vapor esteja a mesma temperatura inicial  $T$ .

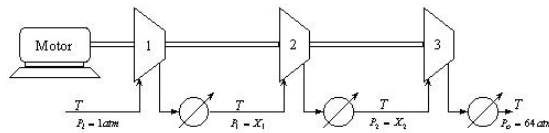


Figura 3.14: Compressor de multi-estágios

Deseja-se determinar as pressões intermediárias que minimize o consumo de energia, ou seja, que o trabalho seja mínimo:

$$W = NRT\beta f(p), \quad (3.34)$$

$$f(p) = \left( p_1^\beta + \left( \frac{p_2}{p_1} \right)^\beta + \left( \frac{64}{p_2} \right)^\beta - 3 \right), \quad (3.35)$$

sendo  $R$  a constante ideal dos gases,  $\beta = \frac{k-1}{k}$ , onde  $k = \frac{C_p}{C_v} = 1,33$  para vapor d'água, resultando em  $\beta = 0,2481$ .

Para que o trabalho seja mínimo basta minimizar  $f(p)$ . Como restrições laterais tem-se  $p_1 \geq 1,0$  e  $1 \leq p_2 \leq 64$ . Portanto, para encontrar os valores das pressões  $p_1$  e  $p_2$  que minimize o consumo de energia basta resolver o seguinte problema:

$$\min f(p) = \left( \left( \frac{p_1}{1} \right)^{0,2481} + \left( \frac{p_2}{p_1} \right)^{0,2481} + \left( \frac{64}{p_2} \right)^{0,2481} - 3 \right) \quad (3.36)$$

A Figura 3.15 representa a função objetivo  $f(p)$  e as suas curvas de nível.

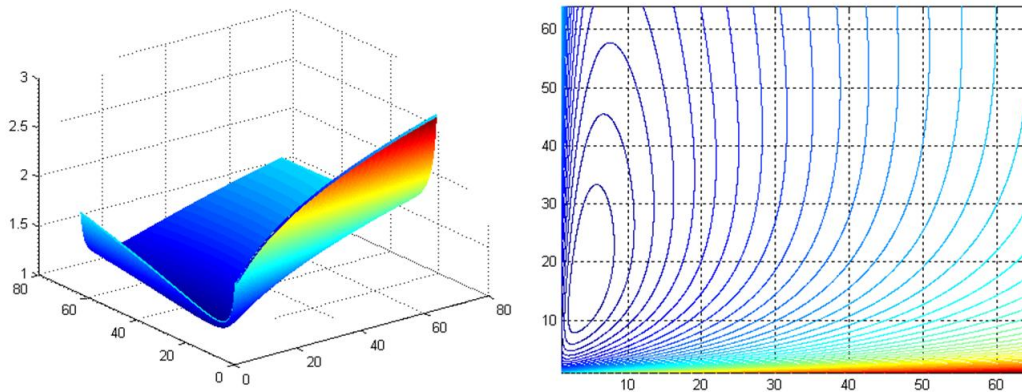


Figura 3.15: Gráfico e curvas de nível da função  $f(p) = \left( \left( \frac{p_1}{1} \right)^\beta + \left( \frac{p_2}{p_1} \right)^\beta + \left( \frac{64}{p_2} \right)^\beta - 3 \right)$

**Observação:** A função que permite calcular o trabalho para o compressor de multi-estágio, dada pela equação (3.36), possui vários mínimos locais, conforme pode ser visto na Figura 3.16 que mostra suas curvas de nível para  $-60 \leq p_i \leq 60$  atm,  $i = 1, 2$ . Por isso, sempre que o algoritmo caminhar para um mínimo que não pertença a região viável, será feita uma adaptação no algoritmo para que este obedeça a região viável  $p_1 \geq 1$  e  $1 \leq p_2 \leq 64$ . Assim, quando o algoritmo encontrar um valor que não pertença à região viável, a seguinte correção deve ser feita:

$$\begin{cases} \text{se } p_1 < 1, & \text{faça } p_1 = 1 \\ \text{se } p_2 > 64, & \text{faça } p_2 = 64 \end{cases} .$$

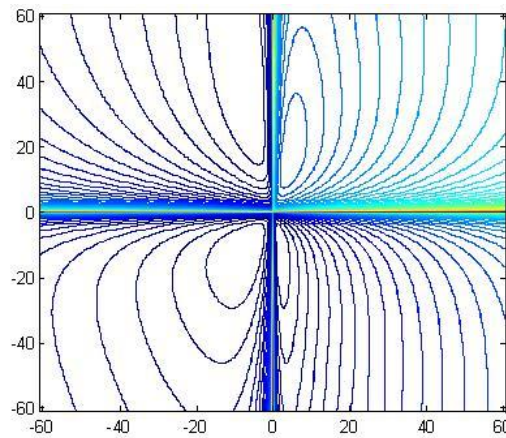


Figura 3.16: Curvas de nível da função  $f(p)$  para pontos fora da região viável

Dado  $x^0 = (10; 20)^T$ , calcular o comprimento do passo  $\alpha^*$  para a primeira iteração, usando o Método da Seção Áurea. A direção de busca  $p_k = -\nabla f(x^k)$ , será dada por:

$$p_k = -\nabla f(p) = \begin{pmatrix} -0,2481 p_1^{-0,7519} + \frac{0,2481 p_2^{0,2481}}{p_1^{1,2481}} \\ -\frac{0,2481}{p_1^{0,2481} p_2^{0,7519}} + \frac{0,2481 \cdot 64^{0,2481}}{p_2^{1,2481}} \end{pmatrix}. \quad (3.37)$$

A Tabela 3.2 resume os cálculos para o Exemplo 2, na qual é encontrado  $\alpha^* = 2$ , donde  $x^1 = (9,9710; 20,0036)^T$  e  $f(x^1) = 1,292260$ .

Iteração k	$a_s$	$c_s$	$d_s$	$b_s$	$\phi(c_s)$	$\phi(d_s)$	$b_s - a_s$
1	0	0,7639	1,2361	2	1,2924234	1,2923613	2
2	0,7639	1,2361	1,5279	2	1,2923613	1,2923229	1,2061
3	1,2361	1,5279	1,7082	2	1,2923229	1,292299	0,7639
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
9	1,9574	1,9737	1,9837	2	1,2922643	1,2922622	0,0426
10	1,9737	1,9837	1,9899	2	1,2922630	1,2922622	0,0263

Tabela 3.2: Método da Seção Áurea aplicado ao Exemplo 2

# Capítulo 4

## Métodos de Otimização Determinística Independentes de Derivadas

Este capítulo apresenta alguns métodos que minimizam uma função de várias variáveis sem o uso de derivadas. São métodos muito simples e conhecidos a muitos anos. A principal importância destes métodos é que foram fundamentais para o desenvolvimento de métodos mais elaborados. Normalmente apresentam um progresso significativo nas primeiras iterações mas podem se tornarem lentos após alguns passos. Segundo Nocedal *et al* (2000) se a função é diferenciável, então os métodos convergem para um ponto estacionário.

### 4.1 O Método das Coordenadas Cíclicas

Este método utiliza a direção dos eixos coordenados como direções de busca. Mais especificamente, o método caminha através das direções  $d_1, \dots, d_n$  onde  $d_j$  é um vetor de zeros, exceto na  $j$ -ésima posição onde assume o valor 1. Desta forma, caminhando na direção de busca  $d_j$ , a variável (ou  $j$ -ésima entrada)  $x_j$  é modificada, enquanto todas as outras permanecem fixas.

Assume-se que a minimização é feita nas componentes  $x_1, \dots, x_n$ , seguindo esta ordem, ou seja, primeiro na variável  $x_1$ , depois  $x_2$  e assim, sucessivamente, conforme esquematizado na Figura 4.1.

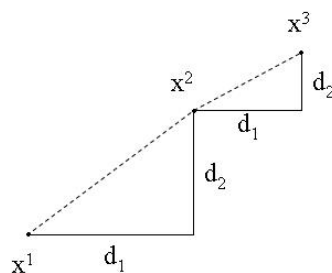


Figura 4.1: Primeiras duas iterações do Método das Coordenadas Cíclicas

A Figura 4.2 mostra que existem alguns casos onde o método pode ser finalizado prematuramente devido a não diferenciabilidade da função. Note que a partir do ponto  $x_2$  o método não consegue progredir, pois ao longo da direção  $(0, 1)$  a função escolhida não diminui. Logo, busca-se melhorar este método escolhendo uma direção de busca alternativa.

A seguir será apresentado o algoritmo básico do método das coordenadas cíclicas:

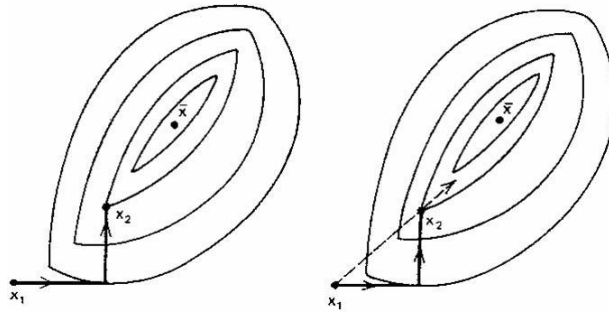


Figura 4.2: Caso onde o método das coordenadas cíclicas foi finalizado prematuramente (Bazaraa *et al*, 1979)

**Dados Iniciais:** Escolha um  $\epsilon > 0$  para interromper o algoritmo e seja  $d_1, \dots, d_n$  as direções coordenadas. Escolha um ponto inicial  $x^1$  faça  $y^1 = x^1$  e  $k = j = 1$  e inicie o processo iterativo.

**Passo 1:** Seja  $\alpha_j$  uma solução ótima do problema:  $\min f(y^j + \alpha d_j)$ , com  $\alpha \in \mathbb{R}$ . Seja  $y^{j+1} = y^j + \alpha_j d_j$ . Se  $j < n$ , troque  $j$  por  $j + 1$  e repita o passo 1. Caso contrário, se  $j = n$ , vá para o Passo 2.

**Passo 2:** Seja  $x^{k+1} = y^{n+1}$ . Se  $\|x^{k+1} - x^k\| < \epsilon$  então pare. Caso contrário, seja  $y^1 = x^{k+1}$  e  $j = 1$ , substitua  $k$  por  $k + 1$  e repita o passo 1.

**Observação:** Nos exemplos que serão apresentados nas próximas seções, para todos os métodos estudados, a obtenção do comprimento do passo  $\alpha$  será por meio da utilização do programa de busca uni-dimensional do toolbox do Matlab<sup>®</sup>.

**EXEMPLO 1:** Considere o problema dado em (3.31). Adote  $x^1 = (0, 2)^T$  e aplique o método das coordenadas cíclicas. Para este exemplo, como o mínimo de  $f(x)$  é zero, será adotado o critério  $|f(x^k)| < \epsilon$ . Seja  $\epsilon = 0,01$ . O processo é interrompido na 18ª iteração com ponto ótimo  $x^* = (1, 975; 1, 310)^T$  e  $f(x^*) = 0,0096 < 0,01$ . Observando a Tabela 4.1 é possível notar que nas primeiras iterações obtém-se um progresso significativo, pois  $f(x^1) = 37$  enquanto que  $f(x^2) = 0,723$  e após alguns passos o progresso torna-se muito lento já que, por exemplo,  $f(x^{12}) = 0,037$  e  $f(x^{13}) = 0,032$ . A cada iteração, os vetores  $y^1$  e  $y^2$  são obtidos por uma busca linear nas direções  $d_1 = (1, 0)^T$  e  $d_2 = (0, 1)^T$ , respectivamente. Pode-se provar que a taxa de convergência deste método é similar ao método da descida máxima (Nocedal *et al*, (2000)). A Figura 4.3 mostra o progresso do método.

**EXEMPLO 2:** Considere o problema de otimização definido em (3.36). Os resultados obtidos partindo do ponto  $x^1 = (10, 20)^T$  para resolver o problema 3.36 são apresentados na Tabela 4.2. Para este exemplo foi adotado o critério de parada  $|f^{k+1} - f^k| < \epsilon = 0,001$ . O processo foi interrompido na 3ª iteração devido a  $|f^3 - f^2| = 0,0008 < \epsilon$  com  $p_1 = 4,1132$  atm e  $p_2 = 16,2249$  atm, no qual a função objetivo vale  $f(p) = 1,2315$ . A Figura 4.4 apresenta o progresso do Método das Coordenadas Cíclicas aplicado ao Exemplo 2.

Pode-se citar como vantagens do Método das Coordenadas Cíclicas o baixo esforço computacional e, principalmente, o fato de não usar derivadas. As desvantagens do método são que no caso de funções que não sejam diferenciáveis o método pode levar a uma solução que não seja ótima, conforme representado na Figura 4.2. Além disso, observa-se que nas primeiras iterações obtém-se um progresso significativo, mas após alguns passos o progresso torna-se muito lento, o que pode ser observado na Figura 4.3.

Iteração k	$(x^k)^T$ $f(x^k)$	j	$d_j^T$	$(y^j)^T$	$\alpha_j$	$(y^{j+1})^T$
1	(0; 2) 37	1	(1; 0)	(0; 2)	3	(3; 2)
		2	(0; 1)	(3; 2)	-0,141	(3; 1,859)
2	(3; 1,859) 0,723	1	(1; 0)	(3; 1,859)	-0,211	(2,789; 1,859)
		2	(0; 1)	(2,789; 1,859)	-0,098	(2,789; 1,761)
3	(2,789; 1,761) 0,422	1	(1; 0)	(2,789; 1,761)	-0,147	(2,642; 1,761)
		2	(0; 1)	(2,642; 1,761)	-0,073	(2,642; 1,688)
4	(2,642; 1,688) 0,272	1	(1; 0)	(2,642; 1,688)	-0,109	(2,533; 1,688)
		2	(0; 1)	(2,533; 1,688)	-0,056	(2,533; 1,632)
5	(2,533; 1,632) 0,188	1	(1; 0)	(2,533; 1,632)	-0,084	(2,448; 1,632)
		2	(0; 1)	(2,448; 1,632)	-0,0450	(2,448; 1,587)
6	(2,448; 1,587) 0,137	1	(1; 0)	(2,448; 1,587)	-0,067	(2,382; 1,587)
		2	(0; 1)	(2,381; 1,587)	-0,037	(2,381; 1,550)
7	(2,381; 1,550) 0,104	1	(1; 0)	(2,381; 1,550)	-0,056	(2,325; 1,550)
		2	(0; 1)	(2,325; 1,550)	-0,031	(2,325; 1,519)
8	(2,325; 1,519) 0,081	1	(1; 0)	(2,325; 1,519)	-0,046	(2,278; 1,519)
		2	(0; 1)	(2,279; 1,519)	-0,027	(2,279; 1,492)
9	(2,279; 1,492) 0,065	1	(1; 0)	(2,279; 1,492)	-0,040	(2,239; 1,492)
		2	(0; 1)	(2,239; 1,492)	-0,023	(2,239; 1,469)
10	(2,239; 1,469) 0,053	1	(1; 0)	(2,239; 1,469)	-0,034	(2,204; 1,469)
		2	(0; 1)	(2,204; 1,469)	-0,020	(2,204; 1,449)
11	(2,204; 1,449) 0,044	1	(1; 0)	(2,204; 1,449)	-0,031	(2,173; 1,449)
		2	(0; 1)	(2,173; 1,449)	-0,018	(2,173; 1,431)
12	(2,173; 1,431) 0,037	1	(1; 0)	(2,173; 1,431)	-0,026	(2,146; 1,431)
		2	(0; 1)	(2,146; 1,431)	-0,016	(2,146; 1,415)
13	(2,146; 1,415) 0,032	1	(1; 0)	(2,146; 1,415)	-0,023	(2,122; 1,415)
		2	(0; 1)	(2,122; 1,415)	-0,015	(2,122; 1,400)
14	(2,122; 1,4) 0,027	1	(1; 0)	(2,122; 1,400)	-0,022	(2,010; 1,400)
		2	(0; 1)	(2,010; 1,400)	-0,068	(2,010; 1,332)
15	(2,010; 1,332) 0,013	1	(1; 0)	(2,010; 1,332)	-0,012	(1,998; 1,332)
		2	(0; 1)	(1,998; 1,332)	-0,008	(1,998; 1,324)
16	(1,998; 1,324) 0,012	1	(1; 0)	(1,998; 1,324)	-0,012	(1,986; 1,324)
		2	(0; 1)	(1,986; 1,324)	-0,007	(1,986; 1,317)
17	(1,986; 1,317) 0,010	1	(1; 0)	(1,986; 1,317)	-0,011	(1,975; 1,317)
		2	(0; 1)	(1,975; 1,317)	-0,007	(1,975; 1,310)
18	(1,975; 1,310) 0,0096					

Tabela 4.1: Método das Coordenadas Cíclicas aplicado ao Exemplo 1

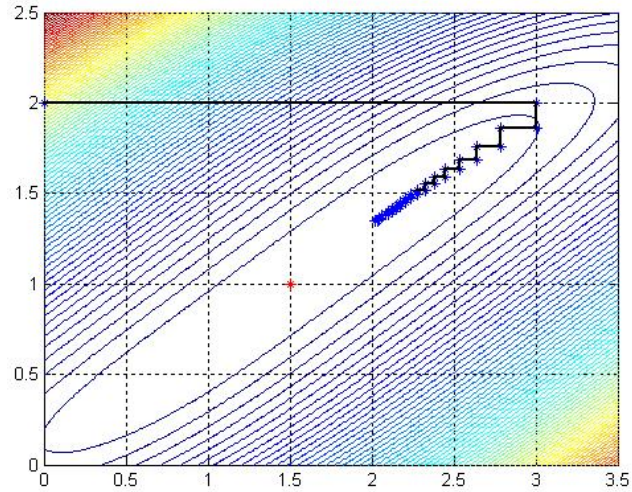


Figura 4.3: Progresso do Método das Coordenadas Cíclicas aplicado ao Exemplo 1

Iteração k	$(x^k)^T$ $f(x^k)$	j	$d_j^T$	$(y^j)^T$	$\alpha_j$	$(y^{j+1})^T$
1	(10; 20)	1	(1; 0)	(10; 20)	-5,5278	(4,4721; 20)
	1,2927	2	(0; 1)	(4,4721; 20)	-3,0821	(4,4721; 16,9179)
2	(4,4721; 16,9179)	1	(1; 0)	(4,4721; 16,9179)	-0,3589	(4,1132; 16,9179)
	1,2323	2	(0; 1)	(4,1132; 16,9179)	-0,6930	(4,1132; 16,2249)
3	(4,1132; 16,2249)					
	1,2315					

Tabela 4.2: Método das Coodenadas Cíclicas aplicado ao Exemplo 2

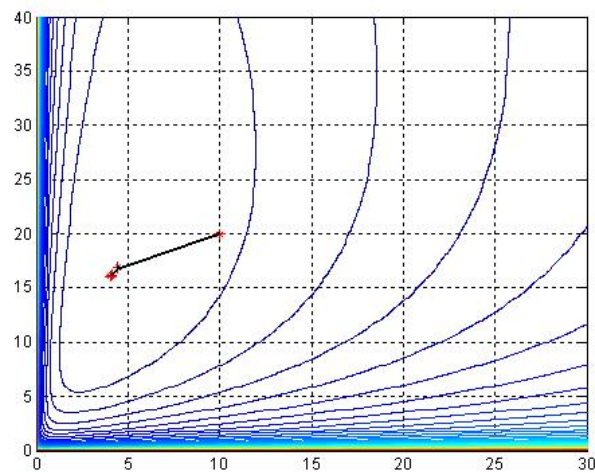


Figura 4.4: Progresso do Método das Coordenadas Cíclicas aplicado ao Exemplo 2

## 4.2 O Método de Hooke e Jeeves com busca linear

O método de Hooke e Jeeves realiza dois tipos de busca - busca exploratória e busca padrão. As primeiras duas iterações do procedimento são ilustradas na Figura 4.5.

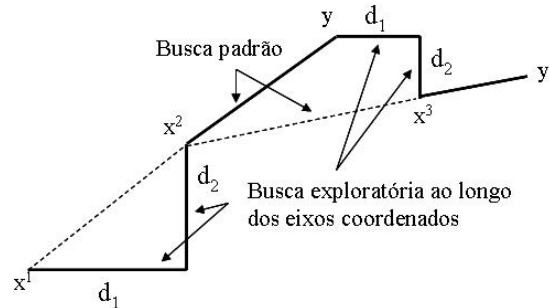


Figura 4.5: Primeiras duas iterações do Método de Hooke e Jeeves, (Bazaraa *et al*, 1979)

Dado  $x^1$ , uma busca exploratória ao longo de uma direção coordenada produz o ponto  $x^2$ . Agora, a busca padrão ao longo da direção  $x^2 - x^1$  conduz ao ponto  $y$ . Uma outra busca exploratória começando em  $y$  resulta no ponto  $x^3$ . A próxima busca padrão ao longo da direção  $x^3 - x^2$ , produz  $y'$ . O processo é então repetido.

O método proposto originalmente por Hooke e Jeeves, não realizava qualquer busca linear, mas assumia um determinado número de passos discretos ao longo de buscas direcionais, como será discutido posteriormente. Aqui, será apresentado uma versão contínua do método usando buscas lineares ao longo das direções  $d_1, d_2, \dots, d_n$  e as direções padrões, conforme o algoritmo:

**Dados iniciais:** Escolha um escalar  $\epsilon > 0$  para critério de parada do algoritmo. Escolha um ponto inicial  $x^1$ , faça  $y^1 = x^1$ ,  $k = j = 1$  e inicie o processo iterativo.

**Passo 1:** Seja  $\alpha_j$  uma solução ótima para o problema de minimização  $f(y^j + \alpha d_j)$  sujeito a  $\alpha \in \mathbb{R}$ , e seja  $y^{j+1} = y^j + \alpha_j d_j$ . Se  $j < n$ , troque  $j$  por  $j + 1$  e repita o passo 1. Caso contrário, se  $j = n$ , seja  $x^{k+1} = y^{n+1}$ . Se  $\|x^{k+1} - x^k\| < \epsilon$  pare, se não, vá para o passo 2.

**Passo 2:** Seja  $d = x^{k+1} - x^k$ , e  $\hat{\alpha}$  uma solução ótima para o problema de minimização  $f(x^{k+1} + \alpha d)$  sujeito a  $\alpha \in \mathbb{R}$ . Seja  $y^1 = x^{k+1} + \hat{\alpha}d$ , seja  $j = 1$ ; troque  $k$  por  $k + 1$  e repita o passo 1.

**EXEMPLO 1:** O método foi aplicado ao problema em (3.31). A tabela 4.3 resume os cálculos do método de Hooke e Jeeves, começando do ponto inicial  $x^1 = (0, 2)^T$  com o seguinte critério de parada:  $|f(x^k)| < 0,01$ . Como  $f(x^5) = 0,005 < 0,01$ , o processo iterativo é interrompido na quinta iteração e toma-se  $x^* = (1, 913; 1, 271)^T$ .

A cada iteração, uma busca exploratória ao longo da coordenada direcional fornece os pontos  $y^2$  e  $y^3$ , e a busca padrão ao longo da direção  $x^{k+1} - x^k$  dá o ponto  $y^k$ , exceto na iteração  $k = 1$ ; onde  $y^1 = x^1$ . Note que são necessárias cinco iterações para mover do ponto inicial para se aproximar do ponto ótimo. Nessa iteração,  $|f(x^5)| = 0,005 < 0,01 = \epsilon$  donde o processo é finalizado. A Figura 4.6 ilustra a geração de pontos pelo Método Hooke e Jeeves usando buscas lineares.

Iteração k	$(x^k)^T$ $f(x^k)$	j	$(y^j)^T$	$d_j^T$	$\alpha_j$	$(y^{j+1})^T$	$d^T$	$\hat{\alpha}$	$(y^3 + \hat{\alpha}d)^T$
1	(0; 2) 37	1 2	(0; 2) (3; 2)	(1; 0) (0; 1)	3 -0,141	(3; 2) (3; 1,859)	- (3; -0,141)	- -0,061	- (2,816; 1,868)
2	(3; 1,859) 0,723	1 2	(2,816; 1,868) (2,801; 1,868)	(1; 0) (0; 1)	-0,014 -0,100	(2,801; 1,868) (2,801; 1,767)	- (-0,198; -0,092)	- 3,545	- (2,097; 1,442)
3	(2,801; 1,767) 0,437	1 2	(2,097; 1,442) (2,162; 1,442)	(1; 0) (0; 1)	0,065 -0,017	(2,162; 1,442) (2,162; 1,425)	- (-0,639; -0,343)	- 0,435	- (1,885; 1,276)
4	(2,162; 1,425) 0,035	1 2	(1,885; 1,276) (1,913; 1,276)	(1; 0) (0; 1)	0,029 -0,004	(1,913; 1,276) (1,913; 1,271)	- (-0,249; -0,153)	- 0,881	- (1,694; 1,136)
5	(1,913; 1,271) 0,005								

Tabela 4.3: Método de Hooke e Jeeves utilizando buscas lineares aplicado ao Exemplo 1

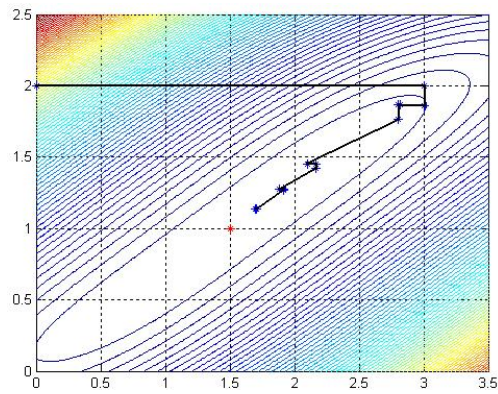


Figura 4.6: Progresso de Hooke e Jeeves utilizando buscas lineares aplicado ao Exemplo 1

**EXEMPLO 2:** O Método de Hooke e Jeeves utilizando buscas lineares será, agora, aplicado ao problema em (3.36). Adotou-se o critério de parada  $|f^{k+1} - f^k| < 0,001$ . Os resultados obtidos estão expostos na Tabela 4.4 e o progresso do método pode ser observado na Figura 4.7. Observe que na primeira iteração foi encontrado um valor negativo para a pressão  $p_1 = -4,5132$ . Como  $p_1 \geq 1$ , foi feito um ajuste no valor de  $p_1$  na segunda iteração afim de que o algoritmo continuasse dentro da região viável do problema. Assim, adotou-se  $p_1 = 1$  resultando em  $y^1 = (1; 11,9087)^T$  ao invés de  $y^1 = (-4,5132; 11,9087)^T$ . O algoritmo é interrompido na terceira iteração, já que  $|f^3 - f^2| = 0,0006 < 0,001$ , com ponto ótimo  $p_1^* = 3,4509$  atm e  $p_2^* = 14,8611$  atm e  $f(p^*) = 1,2329$ .

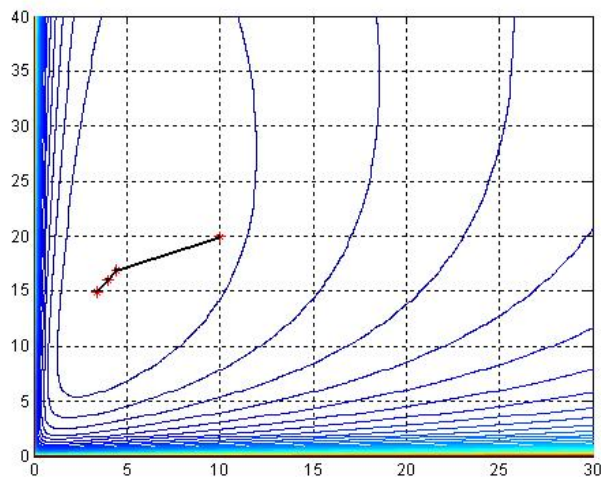


Figura 4.7: Progresso de Hooke e Jeeves utilizando buscas lineares aplicado ao Exemplo 2

Iteração k	$(x^k)^T$ $f(x^k)$	j	$(y^j)^T$	$d_j^T$	$\alpha_j$	$(y^{j+1})^T$	$d^T$	$\hat{\alpha}$	$(y^3 + \hat{\alpha}d)^T$
1	(10; 20) 1,2927	1	(10; 20) (4,4722; 20)	(1; 0) (0; 1)	-5,5278 -3,0818	(4,4722; 20) (4,4722; 16,9182)	- (-5,5278; -3,0818)	- 1,6255	- (-4,5132; 11,9087)
2	(4,4722; 16,9182) 1,2323	1	(1; 11,9087) (3,4509; 11,9087)	(1; 0) (0; 1)	2,4509 2,9524	(3,4509; 11,9087) (3,4509; 14,8611)	- (-1,0213; -2,0571)	- -0,5376	- (4; 15,9671)
3	(3,4509; 14,8611) 1,2329								

Tabela 4.4: Método de Hooke e Jeeves utilizando buscas lineares aplicado ao Exemplo 2

### 4.3 O Método de Hooke e Jeeves com passos discretos

Como mencionado anteriormente, o método de Hooke e Jeeves, não era composto de buscas lineares, mas de um esquema simples envolvendo valores da função. Um algoritmo do método simplificado é dado a seguir:

**Dados iniciais:** Seja  $d_1, \dots, d_n$  as direções coordenadas. Escolha um escalar  $\epsilon > 0$  como critério de parada do algoritmo. Além disso, escolha um comprimento para o passo inicial  $\Delta \geq \epsilon$ , e um fator de aceleração  $\lambda > 0$ . Escolha um ponto inicial  $x^1$ , seja  $y^1 = x^1$  e  $k = j = 1$  e vá para o processo iterativo.

**Passo 1:** Se  $f(y^j + \Delta d_j) < f(y^j)$  o processo é considerado um bom resultado. Seja  $y^{j+1} = y^j + \Delta d_j$ ; e vá para o passo 2. Se, contudo,  $f(y^j + \Delta d_j) > f(y^j)$ , o processo é considerado falho. Nesse caso, se  $f(y^j - \Delta d_j) < f(y^j)$  seja  $y^{j+1} = y^j - \Delta d_j$  e vá para o passo 2.

**Passo 2:** Se  $j < n$  troque  $j$  por  $j + 1$  e repita o passo 1. Caso contrário, vá para o passo 3 se  $f(y^{n+1}) < f(x^k)$  e vá para o passo 4 se  $f(y^{n+1}) \geq f(x^k)$ .

**Passo 3:** Seja  $x^{k+1} = y^{n+1}$  e  $y^1 = x^{k+1} + \lambda(x^{k+1} - x^k)$ . Troque  $k$  por  $k + 1$  e faça  $j = 1$  e vá para o passo 1.

**Passo 4:** Se  $\Delta < \epsilon$  pare,  $x^k$  é a solução. Caso contrário, troque  $\Delta$  por  $\frac{\Delta}{2}$ . Seja  $x^{k+1} = x^k$ ,  $y^1 = x^k$ , troque  $k$  por  $k + 1$ , faça  $j = 1$  e repita o passo 1.

Observe que os passos 1 e 2 acima descritos são buscas exploratórias. Além disso, o passo 3 é um passo acelerado ao longo da direção  $x^{k+1} - x^k$ . Note que a decisão sobre aceitar ou rejeitar o passo acelerado não é feita após a realização de uma busca exploratória. No passo 4, o tamanho  $\Delta$  é reduzido. O procedimento pode facilmente ser modificado pelos diferentes tamanhos  $\Delta$  que são usados ao longo de diferentes direções. Isto é, por vezes, adotado para efeito de escalonamentos.

Nos exemplos seguintes, ( $S$ ) denotará que a tentativa obteve sucesso e ( $F$ ) denotará que a tentativa falhou. Na primeira iteração, e nas iterações subsequentes sempre que  $f(y^3) \geq f(x^k)$ , o vetor  $y^1$  é tomado como  $x^k$ . Caso contrário,  $y^1 = 2x^{k+1} - x^k$ .

**EXEMPLO 1:** Esta técnica será aplicada ao problema em (3.31). Adota-se os parâmetros  $\lambda = 1$  e  $\Delta = 0,2$ . Para dar uma ilustração melhor a Tabela 4.5 resume os cálculos começando do ponto inicial  $x^1 = (0; 2)^T$ . Note que na iteração final  $k = 4$ , o ponto  $x^4 = (1, 2; 0, 8)^T$  é encontrado com valor objetivo  $f^4 = 0,002$ . O procedimento é então interrompido visto que  $|f^4| < 0,01$  e toma-se  $x^* = x^4$ .

A Figura 4.8 mostra o caminho feito pelo algoritmo partindo do ponto  $x^1 = (0; 2)^T$ . Os pontos (em verde) fora do caminho são os passos rejeitados pelo método.

**EXEMPLO 2:** Agora, o método Hooke e Jeeves com passos discretos será aplicado ao problema em (3.36). Toma-se  $x^1 = (10, 20)^T$  como ponto inicial,  $\Delta = 0,2$  e  $\lambda = 1$ . A Tabela 4.6 resume os cálculos. Na nona iteração o algoritmo é interrompido pois  $|f^9 - f^8| = 0,0000 < 0,001$ . O valor ótimo encontrado é  $p^* = (4, 8; 23)^T$  e  $f(p^*) = 1,2399$ . A Figura 4.9 mostra o progresso do método.

Iteração k	$\Delta$	$(x^k)^T$ $f(x^k)$	j	$(y^j)^T$ $f(y^j)$	$d_j^T$	$(y^j + \Delta d_j)^T$ $f(y^j + \Delta d_j)$	$(y^j - \Delta d_j)^T$ $f(y^j - \Delta d_j)$
1	0,2	(0; 2) 37	1	(0; 2) 37	(1; 0)	(0,2; 2) 32,360 (S)	-
			2	(0,2; 2) 32,360	(0; 1)	(0,2; 2,2) 40,514 (F)	(0,2; 1,8) 25,410 (S)
2	0,2	(0,2; 1,8) 25,410	1	(0,4; 1,6) 16,130	(1; 0)	(0,6; 1,6) 13,090 (S)	-
			2	(0,6; 1,6) 13,090	(0; 1)	(0,6; 1,8) 40,514 (F)	(0,6; 1,4) 9,026 (S)
3	0,2	(0,6; 1,4) 9,026	1	(1; 1) 1	(1; 0)	(1,2; 1) 0,360 (S)	-
			2	(1,2; 1) 0,360	(0; 1)	(1,2; 1,2) 1,442 (F)	(1,2; 0,8) 0,002 (S)
4	0,2	(1,2; 0,8) 0,002					

Tabela 4.5: Método Hooke e Jeeves com passos discretos aplicado ao Exemplo 1

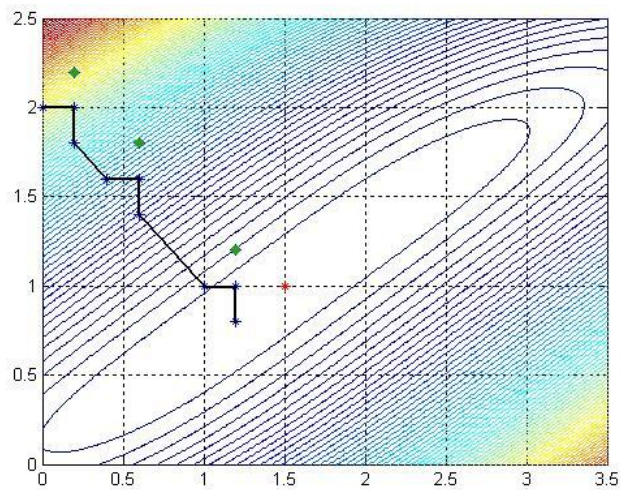


Figura 4.8: Progresso de Hooke e Jeeves com passos discretos aplicado ao Exemplo 1

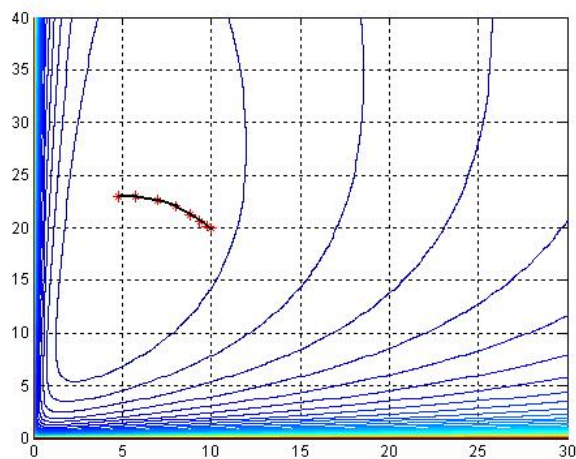


Figura 4.9: Progresso de Hooke e Jeeves com passos discretos aplicado ao Exemplo 2

Iteração k	$\Delta$	$(x^k)^T$ $f(x^k)$	j	$(y^j)^T$ $f(y^j)$	$d_j^T$	$(y^j + \Delta d_j)^T$ $f(y^j + \Delta d_j)$	$(y^j - \Delta d_j)^T$ $f(y^j - \Delta d_j)$
1	0,2	(10; 20) 1,2927	1	(10; 20) 1,2927	(1; 0)	(10,2; 20) 1,2955 (F)	(9,8; 20) 1,2898 (S)
			2	(9,8; 20) 1,2898	(0; 1)	(9,8; 20,2) 1,2895 (S)	-
2	0,2	(9,8; 20,2) 1,2895	1	(9,6; 20,4) 1,2863	(1; 0)	(9,8; 20,4) 1,2891 (F)	(9,4; 20,4) 1,2835 (S)
			2	(9,4; 20,4) 1,2835	(0; 1)	(9,4; 20,6) 1,2832 (S)	-
3	0,2	(9,4; 20,6) 1,2832	1	(9; 21) 1,2772	(1; 0)	(9,2; 21) 1,2780 (F)	(8,80; 21) 1,2746 (S)
			2	(8,8; 21) 1,2746	(0; 1)	(8,8; 21,2) 1,2744 (S)	-
4	0,2	(8,8; 21,2) 1,2744	1	(8,2; 21,8) 1,2663	(1; 0)	(8,4; 21,8) 1,2688 (F)	(8; 21,8) 1,2638 (S)
			2	(8; 21,8) 1,263832	(0; 1)	(8; 22) 1,263784 (S)	-
5	0,2	(8; 22) 1,26378	1	(7,2; 22,8) 1,2548	(1; 0)	(7,4; 22,8) 1,2570 (F)	(7; 22,8) 1,2528 (S)
			2	(7; 22,8) 1,2528	(0; 1)	(7; 23) 1,2529 (F)	(7; 22,6) 1,2527 (S)
6	0,2	(7; 22,6) 1,2527	1	(6; 23,2) 1,2447	(1; 0)	(6,2; 23,2) 1,2461 (F)	(5,8; 23,2) 1,2435 (S)
			2	(5,8; 23,2) 1,2435	(0; 1)	(5,8; 23,4) 1,2437 (F)	(5,8; 23) 1,2432 (S)
7	0,2	(5,8; 23) 1,2432	1	(4,6; 23,2) 1,2405	(1; 0)	(4,8; 23,2) 1,2403 (S)	-
			2	(4,8; 23,2) 1,2403	(0; 1)	(4,8; 23,4) 1,2407 (F)	(4,8; 23) 1,2399 (S)
8	0,2	(4,8; 23) 1,2399	1	(3,8; 23) 1,2448	(1; 0)	(4; 23) 1,2429 (S)	-
			2	(4; 23) 1,2429	(0; 1)	(4; 23,2) 1,2435 (F)	(4; 22,8) 1,2424 (F)
9	0,1	(4,8; 23) 1,2399					

Tabela 4.6: Método Hooke e Jeeves com passos discretos aplicado ao Exemplo 2

## 4.4 O Método de Rosenbrock

O Método de Rosenbrock não utiliza busca lineares, mas adota passos discretos ao longo das direções de busca. Para construir as direções de busca considere  $d_1, \dots, d_n$  vetores linearmente independentes, cada um com norma igual a 1. Suponha  $d_i^T d_j = 0, \forall i \neq j$ . Começando por um vetor inicial  $x^k$ , a função objetivo é minimizada ao longo de cada uma das direções iterativamente, resultando no ponto  $x^{k+1}$ .

Em particular

$$x^{k+1} - x^k = \sum_{j=1}^n \alpha_j d_j, \quad (4.1)$$

sendo que  $\alpha_j$  é a distância movida ao longo de  $d_j$ . A nova coleção de direções  $\bar{d}_1, \bar{d}_2, \dots, \bar{d}_n$  é dada pelo lema:

**Lema 4.1** *Suponha que os vetores  $d_1, \dots, d_n$  são linearmente independentes e mutuamente ortogonais. Então as direções  $\bar{d}_1, \bar{d}_2, \dots, \bar{d}_n$  definidas por:*

$$a_j = \begin{cases} d_j & \alpha_j = 0 \\ \sum_{i=j}^n \alpha_i d_i & \alpha_j \neq 0 \end{cases} \quad (4.2)$$

$$b_j = \begin{cases} a_j & j = 1 \\ a_j - \sum_{i=1}^{j-1} (a_j^T \bar{d}_i) \bar{d}_i & j \geq 2 \end{cases} \quad (4.3)$$

$$\bar{d}_j = \frac{b_j}{\|b_j\|}, \quad (4.4)$$

são também linearmente independentes e mutuamente ortogonais para qualquer conjunto  $\alpha_1, \dots, \alpha_n$ . Ainda, se  $\alpha_j = 0$  então  $\bar{d}_j = d_j$ .

**Demonstração** Primeiro será mostrado que  $a_1, \dots, a_n$  são linearmente independentes. Suponha que

$$\sum_{j=1}^n \omega_j a_j = 0.$$

Seja  $I = \{j : \alpha_j = 0\}$ , e seja  $J(j) = \{i : i \notin I, i \leq j\}$ . Pela definição de  $a_j$  tem-se

$$\begin{aligned} 0 &= \sum_{j=1}^n \omega_j a_j = \sum_{j \in I} \omega_j d_j + \sum_{j \notin I} \omega_j \left( \sum_{i=j}^n \alpha_i d_i \right) \\ &= \sum_{j \in I} \omega_j d_j + \sum_{j \notin I} \left( \alpha_j \sum_{i \in J(j)} \omega_i \right) d_j. \end{aligned}$$

Como  $d_1, \dots, d_n$  são linearmente independentes, segue que  $\omega_j = 0$  para  $j \in I$  e  $\alpha_j \sum_{i \in J(j)} \omega_i = 0$  para  $j \notin I$ . Mas,  $\alpha_j \neq 0$  para  $j \notin I$ , logo,  $\sum_{i \in J(j)} \omega_i = 0$  para cada  $j \notin I$ . Da definição de  $J(j)$ , tem-se portanto que  $\omega_1, \dots, \omega_n = 0$ , donde  $a_1, \dots, a_n$  são linearmente independentes.

A demonstração de que  $b_1, \dots, b_n$  são linearmente independentes seguirá por indução. Já que  $b_1 = a_1 \neq 0$  é suficiente mostrar que se  $b_1, \dots, b_k$  são linearmente independentes, então  $b_1, \dots, b_k, b_{k+1}$  serão também linearmente independentes. Suponha que

$$\sum_{j=1}^{k+1} \omega_j b_j = 0.$$

Da definição de  $b_{k+1}$  tem-se

$$\begin{aligned} 0 &= \sum_{j=1}^k \omega_j b_j + \omega_{k+1} b_{k+1} \\ &= \sum_{j=1}^k \left[ \omega_j - \frac{\omega_{k+1} (a_{k+1}^T \bar{d}_j)}{\|b_j\|} \right] b_j + \omega_{k+1} a_{k+1}. \end{aligned} \quad (4.5)$$

Novamente, da definição dada no teorema, segue que cada vetor  $b_j$  é uma combinação linear de  $a_1, \dots, a_j$ . Visto que  $a_1, \dots, a_{k+1}$  são linearmente independentes, segue de (4.5) que  $\omega_{k+1} = 0$ . Já que  $b_1, \dots, b_k$  são assumidos serem linearmente independentes pela hipótese de indução, de (4.5) tem-se que

$$\omega_j - \frac{\omega_{k+1} (a_{k+1}^T \bar{d}_j)}{\|b_j\|} = 0$$

para  $j = 1, \dots, k$ . Visto que  $\omega_{k+1} = 0$ ,  $\omega_j = 0$  para cada  $j$ . Portanto,  $b_1, \dots, b_k, b_{k+1}$  são linearmente independentes. Como  $\bar{d}_j = \frac{b_j}{\|b_j\|}$ , segue que  $\bar{d}_1, \dots, \bar{d}_n$  são linearmente independentes.

Agora, será mostrado que  $b_1, \dots, b_n$  são ortogonais. Da definição dada,  $b_1^T b_2 = 0$ , assim, é suficiente mostrar que se  $b_1, \dots, b_k$  são mutuamente ortogonais, então  $b_1, \dots, b_k, b_{k+1}$  também serão. De (4.5) e observando que  $b_j^T \bar{d}_i = 0$  para cada  $i \neq j$ ,  $i, j = 1, \dots, k$ , segue que

$$\begin{aligned} b_j^T b_{k+1} &= b_j^T \left[ a_{k+1} - \sum_{i=1}^k (a_{k+1}^T \bar{d}_i) \bar{d}_i \right] \\ &= b_j^T a_{k+1} - (a_{k+1}^T \bar{d}_j) b_j^T \bar{d}_j = 0, \quad i, j = 1, \dots, k. \end{aligned}$$

Assim,  $b_1, \dots, b_k, b_{k+1}$  são mutuamente ortogonais, logo  $\bar{d}_1, \dots, \bar{d}_n$  são mutuamente ortogonais.

Para completar a demonstração, será mostrado que  $\bar{d}_j = d_j$  se  $\alpha_j = 0$ . Da definição, se  $\alpha_j = 0$  então

$$b_j = d_j - \sum_{i=1}^{j-1} \frac{1}{\|b_i\|} (d_j^T b_i) \bar{d}_i. \quad (4.6)$$

Observe que  $b_i$  é uma combinação linear de  $a_1, \dots, a_i$ , e então

$$b_i = \sum_{r=1}^i v_{ir} a_r.$$

Logo,

$$b_i = \sum_{r \in R} v_{ir} d_r + \sum_{r \in \bar{R}} v_{ir} \left( \sum_{s=r}^n \alpha_s d_s \right), \quad (4.7)$$

onde  $R = \{r : r \leq i, \alpha_r = 0\}$  e  $\bar{R} = \{r : r \leq i, \alpha_r \neq 0\}$ . Considere  $i < j$  e observe que  $d_j^T d_v = 0$  para  $v \neq j$ . Para  $r \in R$ ,  $r \leq i < j$  e portanto,  $d_j^T d_r = 0$ . Para  $r \notin R$ ,

$$d_j^T \left( \sum_{s=r}^n \alpha_s d_s \right) = \alpha_j d_j^T d_j = \alpha_j.$$

Pela hipótese  $\alpha_j = 0$  e multiplicando (4.7) por  $d_j^T$  tem-se  $d_j^T b_i = 0$  para  $i < j$ . De (4.6), segue que  $b_j = d_j$ , e portanto,  $\bar{d}_j = d_j$ . O que completa a demonstração.  $\square$

O Método de Rosenbrock utilizando busca linear é então dado pelo algoritmo:

**Passo 1:** Escolha  $x^1$  como ponto inicial e seja  $y^1 = x^1$ ,  $k = j = 1$ . Tome  $\alpha_j$  como solução ótima do problema de minimizar  $f(y^j + \alpha d_j)$  sujeito a  $\alpha \in \mathbb{R}$  e seja  $y^{j+1} = y^j + \alpha_j d_j$ . Se  $j < n$  substitua  $j$  por  $j + 1$  e repita o passo 1. Caso contrário, vá para o passo 2.

**Passo 2:** Seja  $x^{k+1} = y^{n+1}$ . Se  $\|x^{k+1} - x^k\| < \epsilon$ , então pare. Se não, seja  $y^1 = x^{k+1}$ , substitua  $k$  por  $k + 1$ , tome  $j = 1$  e vá para o passo 3.

**Passo 3:** Forme um novo conjunto de direções linearmente independentes e ortogonais utilizando o Lema 4.1 e repita o passo 1.

**EXEMPLO 1:** Considere, novamente o problema em (3.31). Adote  $\epsilon = 0,01$  como critério de parada e escolha  $d_1 = (1; 0)^T$  e  $d_2 = (0; 1)^T$  como as direções coordenadas. Seja  $x^1 = (0, 2)^T$  o ponto inicial. A Tabela 4.7 resume as etapas do método.

Iteração k	$(x^k)^T$ $f(x^k)$	j	$(y^j)^T$ $f(y^j)$	$d_j^T$	$\alpha_j$	$(y^{j+1})^T$ $f(y^{j+1})$
1	(0; 2) 37	1	(0; 2) 37	(1; 0)	3	(3; 2) 1
		2	(3; 2) 1	(0; 1)	-0,141	(3; 1,859) 0,723
2	(3; 1,859) 0,723	1	(3; 1,859) 0,723	(0,999; -0,047)	-0,184	(2,816; 1,868) 0,568
		2	(2,816; 1,868) 0,568	(-0,047; -0,999)	0,098	(2,811; 1,770) 0,449
3	(2,811; 1,770) 0,449	1	(2,811; 1,770) 0,449	(-0,904; -0,428)	0,827	(2,064; 1,415) 0,044
		2	(2,064; 1,415) 0,044	(0,428; -0,904)	0,041	(2,082; 1,379) 0,021
4	(2,082; 1,379) 0,021	1	(2,082; 1,379) 0,021	(-0,881; -0,472)	0,233	(1,876; 1,269) 0,008
		2	(1,876; 1,269) 0,008	(0,472; -0,881)	0,017	(1,884; 1,254) 0,004
5	(1,884; 1,254) 0,004	1	(1,884; 1,254) 0,004	(-0,845; -0,535)	0,279	(1,651; 1,105) $3,53 \times 10^{-4}$
		2	(1,651; 1,105) $3,53 \times 10^{-4}$	(0,535; -0,845)	0,004	(1,653; 1,102) $1,08 \times 10^{-4}$
6	(1,653; 1,102) $1,08 \times 10^{-4}$					

Tabela 4.7: Método Rosenbrock aplicado ao Exemplo 1

Pode-se notar que, para este exemplo, o método de Rosenbrock é o mais prático, com apenas seis iterações o ótimo foi alcançado com precisão superior à requerida, já que,  $x^* = (1,653; 1,102)^T$  e  $f(x^*) = 1,08 \times 10^{-4}$ . A Figura 4.10 mostra o caminho percorrido pelo método de Rosenbrock para alcançar o ponto ótimo partindo do ponto inicial.

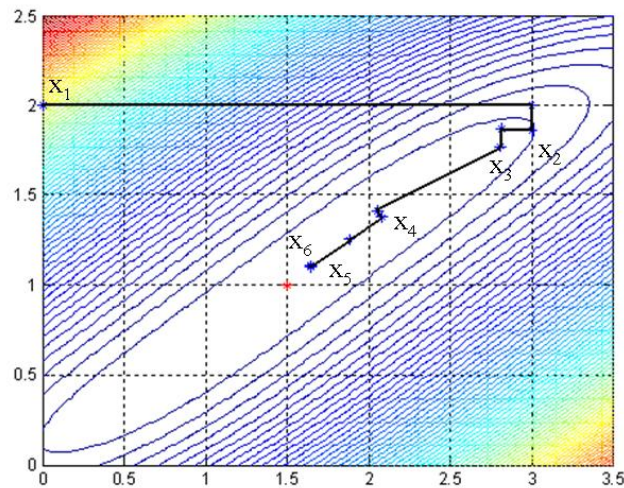


Figura 4.10: Progresso do Método Rosenbrok aplicado ao Exemplo 1

**EXEMPLO 2:** Será aplicado, agora, o método Rosenbrock ao problema em 3.36. Partindo do ponto  $x^1 = (10, 20)^T$  a Tabela 4.8 resume os cálculos para se alcançar o ponto ótimo, o que acontece na terceira iteração visto que  $|f^3 - f^2| = 7,22 \times 10^{-4} < 0,001$ . Tem-se  $p^* = (4, 1159; 16, 6040)^T$  e  $f(p^*) = 1,2316$ . A Figura 4.11 mostra o progresso do método.

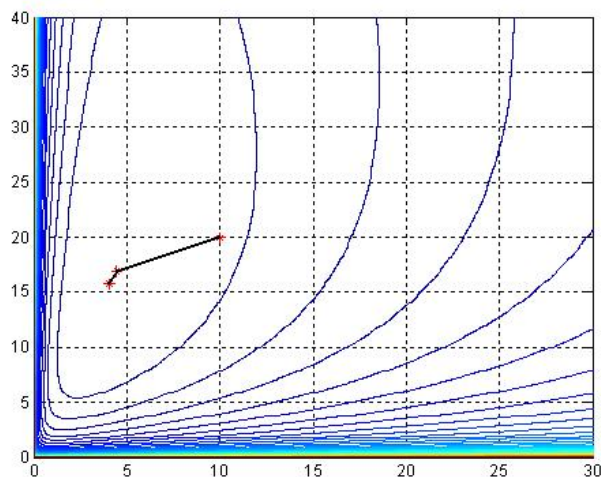


Figura 4.11: Progresso do Método Rosenbrok aplicado ao Exemplo 2

Há, ainda, uma versão contínua do Método de Rosenbrock que utiliza passos discretos ao longo de uma direção. Nesta variação do método, em cada iteração, uma aceleração é incorporada ao método para aumentar ou diminuir adequadamente os tamanhos dos passos até que se consiga diminuir o valor da função. Porém, a escolha do tamanho destes passos é um tanto aleatória o que torna o método muito lento. Por este motivo, esta versão do Método de Rosenbrock não será apresentada aqui. Para mais informações, leia em Bazaraa *et al.* (1979).

Iteração k	$(x^k)^T$ $f(x^k)$	j	$(y^j)^T$ $f(y^j)$	$d_j^T$	$\alpha_j$	$(y^{j+1})^T$ $f(y^{j+1})$
1	(10; 20) 1,2927	1	(10; 20) 1,2927	(1; 0)	-5,5278	(4,4722; 20) 1,2347
		2	(4,4722; 20) 1,2347	(0; 1)	-3,0819	(4,4722; 16,9180) 1,2323
2	(4,4722; 16,9180) 1,2323	1	(4,4722; 16,9180) 1,2323	(-0,8734; -0,4870)	0,4641	(4,0668; 16,6920) 1,2316
		2	(4,0668; 16,6920) 1,2316	(0,4870; -0,8734)	0,1007	(4,1159; 16,6040) 1,2316
3	(4,1159; 16,6040) 1,2316					

Tabela 4.8: Método Rosenbrock aplicado ao Exemplo 2

# Capítulo 5

## Método da Descida Máxima

O método de descida máxima é um método de busca linear que se move ao longo da direção

$$p_k = -\nabla f^k \quad (5.1)$$

em todos os passos. É intuitivo pensar que devemos escolher uma direção ao longo da qual a função  $f$  decresce mais rapidamente.

### 5.1 Direção de descida máxima (steepest descent direction)

Para escolher uma direção ao longo da qual a função  $f$  decresce rapidamente, utiliza-se novamente o Teorema de Taylor, equação (1.7), segundo o qual dada uma direção de busca  $p$  e um comprimento de passo  $\alpha$ , tem-se

$$f(x^k + \alpha p) = f(x^k) + \alpha p^T \nabla f^k + \frac{1}{2} \alpha^2 p^T \nabla^2 f(x^k + tp) p, \quad t \in (0, \alpha). \quad (5.2)$$

A razão com que a função  $f$  muda ao longo da direção  $p$  é o coeficiente de  $\alpha$ , ou seja,  $p^T \nabla f^k$ . Desta forma, a direção unitária  $p$  de decrescimento mais rápido é obtida solucionando-se o problema

$$\min_p p^T \nabla f^k, \quad \text{sujeito a } \|p\| = 1. \quad (5.3)$$

Uma vez que  $p^T \nabla f^k = \|p\| \cdot \|\nabla f^k\| \cdot \cos\theta$ , onde  $\theta$  é o ângulo entre  $p$  e  $\nabla f^k$ , é fácil ver que a minimização ocorre para  $\cos\theta = -1$  e que

$$p = -\frac{\nabla f^k}{\|\nabla f^k\|} \quad (5.4)$$

como pretendido. Como pode ser visto na Figura 5.1, a direção de busca é oposta ao gradiente e ortogonal às curvas de nível da função.

Uma das vantagens deste método é que requer apenas o cálculo do gradiente  $\nabla f^k$ , e não depende das segundas derivadas. Entretanto, ele pode se tornar *exaustivamente lento* para alguns problemas mais difíceis.

Além da direção oposta ao gradiente, outras direções podem ser escolhidas. De forma geral, qualquer direção que faça um ângulo estritamente menor que  $90^\circ$  com relação  $-\nabla f^k$ , garante o decréscimo da função  $f$ , desde que o comprimento do passo seja suficientemente pequeno, conforme ilustrado na Figura 5.2.

Isto pode ser verificado usando novamente o Teorema de Taylor, equação (1.7):

$$f(x^k + \epsilon p_k) = f(x^k) + \epsilon p_k^T \nabla f^k + 0(\epsilon^2) \approx f(x^k) + \epsilon p_k^T \nabla f^k, \quad (5.5)$$

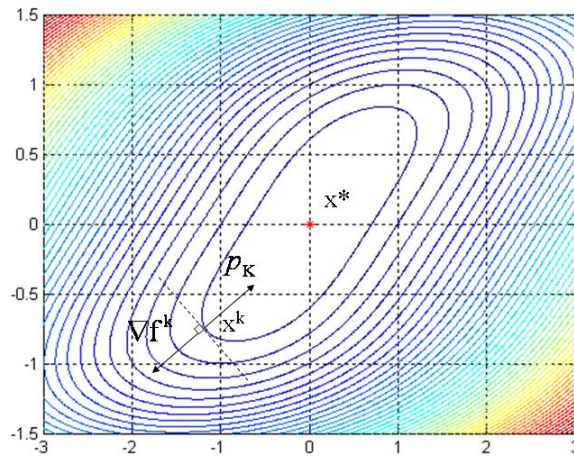


Figura 5.1: Direção de descida máxima

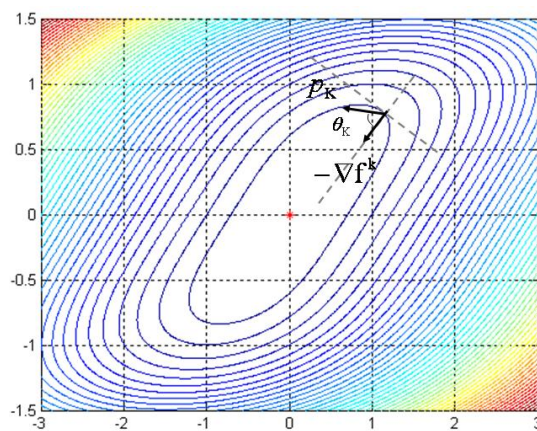


Figura 5.2: Uma direção de descida  $p_k$

onde,  $p_k$  é a direção de descida e o ângulo entre  $p_k$  e o gradiente  $\nabla f^k$  é dado por  $\theta_k$ . Considerando  $\cos\theta_k < 0$ , vem

$$p_k^T \nabla f^k = \|p_k\| \cdot \|\nabla f^k\| \cdot \cos\theta_k < 0, \quad (5.6)$$

isto significa que  $f(x^k + \epsilon p_k) < f(x^k)$  para todo valor positivo mas suficientemente pequeno de  $\epsilon$ .

## 5.2 Convergência dos Métodos de Descida Máxima

Para que o Método de Descida Máxima seja globalmente convergente, não se deve apenas escolher bem o comprimento do passo  $\alpha_k$ , mas também ter uma boa escolha da direção de busca  $p_k$ . Nesta seção serão estudadas os requisitos para escolher a direção de busca, focando em uma propriedade chave: o ângulo  $\theta_k$  entre a direção e  $-\nabla f^k$ . Conforme a equação (5.6), tem-se

$$\cos\theta_k = \frac{-(\nabla f^k)^T p_k}{\|\nabla f^k\| \cdot \|p_k\|}. \quad (5.7)$$

O teorema de Zoutendijk que será apresentado a seguir tem consequência de longo alcance, pois seus resultados são abrangentes podendo ser estendidos a outros métodos. Ele quantifica o efeito de uma escolha apropriada do comprimento de passo  $\alpha_k$  e mostra que o método da descida máxima é globalmente convergente. Para os outros algoritmos, este descreve o quanto  $p_k$  pode se desviar da direção de descida máxima e ainda produzir uma direção globalmente

convergente. Várias condições de parada da busca linear podem ser usadas para estabelecer este resultado, mas para usar um exemplo será considerada somente a condição de Wolfe (3.8) e (3.9). O resultado deste teorema tornar-se evidente ao longo do desenvolvimento do texto.

**Teorema 5.1** (*Zoutendijk*) *Considere uma iteração da forma (3.1) onde  $p_k$  é a direção de descida e  $\alpha_k$  satisfaz as condições de Wolfe (3.8) e (3.9). Suponha que  $f$  é limitada em  $\mathbb{R}^n$  e que  $f$  é continuamente diferenciável em um conjunto aberto  $V$  contendo o conjunto de nível  $L := \{x \in \mathbb{R}^n; f(x) \leq f(x^0)\}$ , onde  $x^0$  é o ponto inicial da iteração. Assuma também que o gradiente  $\nabla f$  é Lipschitz contínuo em  $V$ , isto é, existe uma constante  $L > 0$  tal que*

$$\|\nabla f(x) - \nabla f(\tilde{x})\| \leq L\|x - \tilde{x}\| \quad (5.8)$$

para todo  $x, \tilde{x} \in V$ . Então

$$\sum_{k \geq 0} \cos^2 \theta_k \|\nabla f^k\|^2 < \infty. \quad (5.9)$$

**Demonstração:** De (3.9) e (3.1) tem-se que

$$(\nabla f^{k+1} - \nabla f^k)^T p_k \geq (c_2 - 1) \nabla (f^k)^T p_k, \quad \alpha_k p_k = x^{k+1} - x^k, \quad (5.10)$$

enquanto que a condição de Lipschitz (5.8) implica que

$$(\nabla f^{k+1} - \nabla f^k)^T p_k \leq \alpha_k L \|p_k\|^2. \quad (5.11)$$

Combinando as duas relações anteriores, resulta:

$$\alpha_k \geq \frac{(c_2 - 1)}{L} \cdot \frac{(\nabla f^k)^T p_k}{\|p_k\|^2}. \quad (5.12)$$

Substituindo esta inequação na primeira condição de Wolfe (3.8) obtém-se

$$f^{k+1} \leq f^k - c_1 \frac{(1 - c_2)}{L} \cdot \frac{((\nabla f^k)^T p_k)^2}{\|p_k\|^2}. \quad (5.13)$$

da definição (5.7), a relação acima pode ser escrita como:

$$f^{k+1} \leq f^k - c \cdot \cos^2 \theta_k \|\nabla f^k\|^2 \quad (5.14)$$

onde  $c = c_1(1 - c_2)/L$ . Assumindo que esta expressão é válida para todos os índices menores ou iguais a  $k$ , obtém-se

$$f^{k+1} \leq f^0 - c \sum_{j=0}^k \cos^2 \theta_j \|\nabla f^j\|^2. \quad (5.15)$$

Uma vez que  $f$  é limitada, tem-se que  $f^0 - f^{k+1}$  é menor que alguma constante positiva, para todo  $k$ . Então, considerando o limite em (5.15), obtém-se

$$\sum_{k=0}^{\infty} \cos^2 \theta_k \|\nabla f^k\|^2 < \infty, \quad (5.16)$$

concluindo a demonstração.  $\square$

Observe que as hipóteses do Teorema 5.1 não são muito restritivas. Se a função  $f$  não for limitada, o problema de otimização não será bem definido. A hipótese de suavidade (continuidade

do gradiente - condição de Lipschitz) são assumidas em muitos teoremas de convergência local e são frequentemente satisfeitas na prática.

A condição (5.9) implica que

$$\cos^2\theta_k \|\nabla f^k\|^2 \rightarrow 0 \quad (5.17)$$

este limite pode ser usado para obter o resultado de convergência global dos algoritmos de busca linear.

Se o método escolhido para a direção de busca no processo iterativo (3.1) garante que o ângulo  $\theta_k$  definido em (5.7) é limitado por  $90^\circ$ , existe uma constante positiva  $\delta$  tal que

$$\cos\theta_k \geq \delta > 0 \text{ para todo } k \quad (5.18)$$

De (5.17) resulta imediatamente que:

$$\lim_{k \rightarrow \infty} \|\nabla f^k\| = 0. \quad (5.19)$$

Em outras palavras, pode-se garantir que a norma do gradiente converge para zero, desde que a direção de busca nunca seja muito próxima da ortogonalidade com o gradiente. No caso do *Método da Descida Máxima*, no qual a direção de busca é paralela à direção negativa (oposta) ao gradiente da equação (5.1), resulta que a sequência de gradientes converge para zero, provando que a busca linear satisfaz a condição de Wolfe.

### 5.3 Taxas de convergência do Método de Descida Máxima

Em uma análise elementar, poderia parecer que projetar algoritmos de otimização com boas propriedades de convergência é fácil, uma vez que basta garantir que a direção de busca  $p_k$  não se torne ortogonal ao gradiente  $\nabla f^k$  ou tomar o método da descida máxima regularmente. Poderia-se simplesmente calcular  $\cos\theta_k$  em cada iteração e adotar  $p_k$  como a direção de descida máxima se  $\cos\theta_k$  for inferior a alguma constante  $\delta > 0$  pré-estabelecida. Este tipo de *teste do ângulo* garante a convergência global, mas não é desejável por duas razões:

1. pode impedir uma taxa rápida de convergência, pois devido a problemas de mal-condicionamento da matriz Hessiana, algumas vezes torna-se necessário adotar uma direção de busca quase ortogonal ao gradiente e uma escolha não adequada do parâmetro  $\delta$  pode causar a rejeição de vários passos; além disso,
2. o teste do ângulo pode prejudicar a característica de invariância dos métodos Quase-Newton.

Estratégias computacionais para obter convergência rápida podem algumas vezes ser conflitantes com o requisito de convergência global e vice-versa. Por exemplo, o método da descida máxima é essencialmente um algoritmo globalmente convergente, mas é muito lento na prática.

Por outro lado, o método de Newton converge rapidamente quando adota um ponto inicial suficientemente próximo da solução, mas seus passos podem levar a uma direção que resulta em pontos distantes da solução. O desafio é projetar um algoritmo que incorpore as duas propriedades: garantia de uma boa convergência global e uma rápida taxa de convergência.

Para fins acadêmicos, visando auxiliar a compreensão deste método, será considerado um caso ideal, no qual a função objetivo é quadrática e as direções de busca exatamente lineares.

Suponha que

$$f(x) = \frac{1}{2}x^T Qx - b^T x, \quad (5.20)$$

onde  $Q$  é uma matriz simétrica positiva definida. O gradiente é dado por:

$$\nabla f(x) = Qx - b \quad (5.21)$$

e o mínimo  $x^*$  é a solução única do sistema linear:

$$Qx = b. \quad (5.22)$$

Seja  $r = b - Qx = -\nabla f(x)$  o resíduo. É fácil calcular o comprimento do passo que minimiza  $f(x^k - \alpha_k \nabla f^k)$ . Da série de Taylor de 1º ordem, tem-se:

$$f(x^k + \alpha_k r_k) = f(x^k) - \alpha_k r_k^T r_k + \frac{1}{2} \alpha_k^2 r_k^T Q r_k, \quad (5.23)$$

onde  $x^{k+1} - x^k = \alpha_k r_k$ ,  $\nabla f = -r_k$  e  $\nabla^2 f = Q$ .

Diferenciando a função em (5.23) em relação a  $\alpha_k$ , e fazendo a derivada igual a zero, obtém-se

$$\frac{\partial f}{\partial \alpha_k} = -r_k^T r_k + \alpha_k r_k^T Q r_k = 0 \implies \alpha_k (\nabla f^k)^T Q \nabla f^k - (\nabla f^k)^T \nabla f^k = 0, \quad (5.24)$$

donde

$$\alpha_k = \frac{(\nabla f^k)^T \nabla f^k}{(\nabla f^k)^T Q \nabla f^k}. \quad (5.25)$$

Assim, a iteração do método da descida máxima para a função quadrática (5.20) é dada por:

$$x^{k+1} = x^k - \left( \frac{(\nabla f^k)^T \nabla f^k}{(\nabla f^k)^T Q \nabla f^k} \right) \nabla f^k. \quad (5.26)$$

Uma vez que  $\nabla f^k = Qx^k - b$ , esta equação leva a uma expressão da forma fechada para  $x^{k+1}$  em termos de  $x^k$ .

A Figura 5.3 mostra uma sequência típica gerada por este método para uma função quadrática. As curvas de nível de  $f$  são elipsoides cujas linhas axiais são ortogonais aos auto-vetores de  $Q$ . Observe que as iterações caminham em zigue-zague na direção da solução.

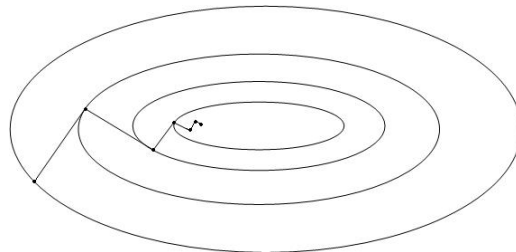


Figura 5.3: Passos do Método da Descida Máxima

Para quantificar a taxa de convergência considere a seguinte norma  $\|x\|_Q^2 = x^T Q x$ . Utilizando a relação  $Qx^* = b$ , mostra-se que:

$$\frac{1}{2} \|x - x^*\|_Q^2 = f(x) - f(x^*). \quad (5.27)$$

De fato, desenvolvendo o lado direito da equação (5.27) tem-se:

$$\begin{aligned} \frac{1}{2} \|x - x^*\|_Q^2 &= \frac{1}{2} (x - x^*)^T Q (x - x^*) = \frac{1}{2} (x - x^*)^T (Qx - Qx^*) \\ &= \frac{1}{2} (x - x^*)^T (Qx - b) = \frac{1}{2} (x - x^*)^T Qx - \frac{1}{2} (x - x^*)^T b \\ &= \frac{1}{2} x^T Qx - \frac{1}{2} x^{*T} Qx - \frac{1}{2} x^T b + \frac{1}{2} x^{*T} b \\ &= \frac{1}{2} x^T Qx - \frac{1}{2} b^T x - \frac{1}{2} x^T b + \frac{1}{2} x^{*T} b \\ &= \frac{1}{2} x^T Qx - b^T x + \frac{1}{2} x^{*T} b. \end{aligned} \quad (5.28)$$

Desenvolvendo o lado esquerdo da equação (5.27) vem que:

$$\begin{aligned}
f(x) - f(x^*) &= \frac{1}{2}x^T Qx - b^T x - \frac{1}{2}x^{*T} Qx^* + b^T x^* \\
&= \frac{1}{2}x^T Qx - b^T x - \frac{1}{2}x^{*T} b + b^T x^* \\
&= \frac{1}{2}x^T Qx - b^T x + \frac{1}{2}b^T x^* \\
&= \frac{1}{2}x^T Qx - b^T x + \frac{1}{2}x^{*T} b.
\end{aligned} \tag{5.29}$$

De (5.28) e (5.29) segue que  $\frac{1}{2}\|x - x^*\|_Q^2 = f(x) - f(x^*)$ .

Assim, esta norma mede a diferença entre o valor da função objetivo atual e da função objetivo no ponto ótimo. Usando a equação (5.21) pode-se observar que  $\nabla f^k = Qx^k - b$ . Mas  $b = Qx^*$ , donde

$$\nabla f^k = Q(x^k - x^*). \tag{5.30}$$

Deseja-se, agora, obter uma expressão que descreve o decréscimo de  $f(x)$  em cada iteração. Com esta finalidade, usando a equação (5.26), pode-se escrever:

$$\begin{aligned}
&\|x^k - x^*\|_Q^2 - \|x^{k+1} - x^*\|_Q^2 = \\
&= (x^k - x^*)^T Q(x^k - x^*) - (x^{k+1} - x^*)^T Q(x^{k+1} - x^*) \\
&= (x^k - x^*)^T Q(x^k - x^*) - \left(x^k - x^* - \frac{(\nabla f^k)^T \nabla f^k}{(\nabla f^k)^T Q \nabla f^k} \nabla f^k\right)^T Q \left(x^k - x^* - \frac{(\nabla f^k)^T \nabla f^k}{(\nabla f^k)^T Q \nabla f^k} \nabla f^k\right) \\
&= (x^k - x^*)^T Q(x^k - x^*) - (x^k - x^* - \alpha_k \nabla f^k)^T Q(x^k - x^* - \alpha_k \nabla f^k) \\
&= (x^k - x^*)^T Q(x^k - x^*) - (x^k - x^*)^T Q(x^k - x^* - \alpha_k \nabla f^k) + \alpha_k (\nabla f^k)^T Q(x^k - x^* - \alpha_k \nabla f^k) \\
&= (x^k - x^*)^T Q(x^k - x^*) - (x^k - x^*)^T Q(x^k - x^*) + (x^k - x^*)^T Q \alpha_k \nabla f^k + \alpha_k \nabla f^k Q(x^k - x^*) - \\
&\quad - \alpha_k (\nabla f^k)^T Q \alpha_k \nabla f^k \\
&= 2\alpha_k (\nabla f^k)^T Q(x^k - x^*) - \alpha_k^2 (\nabla f^k)^T Q \nabla f^k.
\end{aligned} \tag{5.31}$$

Substituindo o comprimento do passo, dado pela equação (5.25), lembrando que  $\nabla f^k = Q(x^k - x^*)$ , tem-se

$$\begin{aligned}
&\|x^k - x^*\|_Q^2 - \|x^{k+1} - x^*\|_Q^2 = \\
&= 2 \frac{(\nabla f^k)^T \nabla f^k}{(\nabla f^k)^T Q \nabla f^k} (\nabla f^k)^T Q(x^k - x^*) - \left(\frac{(\nabla f^k)^T \nabla f^k}{(\nabla f^k)^T Q \nabla f^k}\right)^2 (\nabla f^k)^T Q \nabla f^k \\
&= 2 \frac{(\nabla f^k)^T \nabla f^k}{(\nabla f^k)^T Q \nabla f^k} (\nabla f^k)^T \nabla f^k - \frac{((\nabla f^k)^T \nabla f^k)^2}{(\nabla f^k)^T Q \nabla f^k} \\
&= \frac{2((\nabla f^k)^T \nabla f^k)^2 - ((\nabla f^k)^T \nabla f^k)^2}{(\nabla f^k)^T Q \nabla f^k} \\
&= \frac{((\nabla f^k)^T \nabla f^k)^2}{(\nabla f^k)^T Q \nabla f^k}.
\end{aligned} \tag{5.32}$$

Logo,

$$\frac{\|x^k - x^*\|_Q^2 - \|x^{k+1} - x^*\|_Q^2}{\|x^k - x^*\|_Q^2} = \frac{((\nabla f^k)^T \nabla f^k)^2}{(\nabla f^k)^T Q \nabla f^k \cdot (x^k - x^*)^T Q (x^k - x^*)} \quad (5.33)$$

Já que  $Q$  é invertível,

$$\nabla f^k = Q(x^k - x^*) \quad \Rightarrow \quad (x^k - x^*) = Q^{-1} \nabla f^k. \quad (5.34)$$

Mais ainda,

$$(x^k - x^*)^T Q (x^k - x^*) = (\nabla f^k)^T Q^{-1} \nabla f^k. \quad (5.35)$$

Substituindo (5.35) em (5.33) segue que

$$\frac{\|x^k - x^*\|_Q^2 - \|x^{k+1} - x^*\|_Q^2}{\|x^k - x^*\|_Q^2} = \frac{((\nabla f^k)^T \nabla f^k)^2}{(\nabla f^k)^T Q \nabla f^k (\nabla f^k)^T Q^{-1} \nabla f^k}, \quad (5.36)$$

Donde,

$$\|x^k - x^*\|_Q^2 - \|x^{k+1} - x^*\|_Q^2 = \left( \frac{((\nabla f^k)^T \nabla f^k)^2}{(\nabla f^k)^T Q \nabla f^k (\nabla f^k)^T Q^{-1} \nabla f^k} \right) \|x^k - x^*\|_Q^2. \quad (5.37)$$

Portanto,

$$\|x^{k+1} - x^*\|_Q^2 = \left\{ 1 - \frac{((\nabla f^k)^T \nabla f^k)^2}{(\nabla f^k)^T Q \nabla f^k (\nabla f^k)^T Q^{-1} \nabla f^k} \right\} \|x^k - x^*\|_Q^2. \quad (5.38)$$

Finalmente, obtém-se a expressão (5.38) que descreve o decaimento exato de  $f$  para cada iteração. Mas, como o termo entre chaves é difícil de ser interpretado, torna-se mais útil limitar em termos do *fator de condição*  $k(Q) = \lambda_n/\lambda_1$  do problema, sendo que  $\lambda_i$  são os auto-valores de  $Q$ .

**Teorema 5.2** *Quando o método da descida máxima é aplicado a uma função quadrática fortemente convexa conforme (5.20) e direções exatamente lineares (5.25), a norma do erro (5.27) satisfaz*

$$\|x^{k+1} - x^*\|_Q^2 \leq \left( \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} \right)^2 \|x^k - x^*\|_Q^2, \quad (5.39)$$

onde  $0 \leq \lambda_1 \leq \lambda_2 \leq \dots < \lambda_n$  são os auto-valores de  $Q$ .

A prova deste resultado é dada em Luenberger (1984).

As desigualdades (5.39) e (5.27) mostram que a função converge para o mínimo  $f^k$  em uma razão linear. Como um caso especial deste resultado, pode-se ver que a convergência é obtida com apenas uma iteração se todos os auto-valores são iguais ( $\lambda_1 = \lambda_n$ ). Assim,  $Q$  é uma matriz múltipla da matriz identidade, e todas as curvas de nível são círculos e a direção de busca sempre aponta para a solução. Em geral, se o fator de condição  $k(Q) = \lambda_n/\lambda_1$  aumenta, as curvas de nível tornam-se mais alongadas, e o zigue-zague mostrado na Figura 5.3 torna-se mais nítido e segundo (5.39) a taxa de convergência decai.

O comportamento da taxa de convergência para o método de descida máxima é essencialmente o mesmo para funções não-lineares. No resultado seguinte, apresentado por Nocedal *et al* (2000), assume-se que o comprimento do passo é um minimizador global ao longo da direção de busca.

Iteração k	$(x^k)^T$	$f(x^k)$	$\alpha_k$	$p_k^T$	$\ \nabla f(x^k)\ $
1	(0; 2)	37	0,0355	(24; -40)	46,6476
2	(0,8521 ; 0,5799)	0,0324	2,5591	(0,1418; 0,0839)	0,1648
⋮	⋮	⋮	⋮	⋮	
5	(1,2922 ; 0,8563)	$6,6777 \times 10^{-4}$	0,0383	(-0,0622; 0,1051)	0,1218
10	(1,3534 ; 0,9024)	$9,0954 \times 10^{-5}$	9,2262	(0,0018; 0,0010)	0,0021
16	(1,3873 ; 0,9249)	$3,1791 \times 10^{-5}$	0,5	$(0,92; 0,31) \times 10^{-3}$	$9,71 \times 10^{-4}$

Tabela 5.1: Método da Descida Máxima aplicado ao Exemplo 1

**Teorema 5.3** *Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é duas vezes continuamente diferenciável e que as iterações geradas pelo método da descida máxima com direção exatamente lineares convergem para o ponto  $x^*$  para o qual a matriz Hessiana  $\nabla^2 f(x^*)$  seja positiva definida. Seja  $r$  um escalar que satisfaça  $r \in \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}, 1\right)$ , sendo  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  os auto-valores de  $\nabla^2 f(x^*)$ . Então para todo  $k$  suficientemente grande tem-se*

$$f(x^{k+1}) - f(x^*) \leq r^2[f(x^k) - f(x^*)]. \quad (5.40)$$

Em geral, pode-se esperar que a taxa de convergência melhore se uma direção não linear for usada. Entretanto, o Teorema 5.3 mostra que o método da descida máxima pode ter uma taxa de convergência lenta (inaceitável), mesmo quando a matriz Hessiana for razoavelmente bem-comportada, isto é, bem condicionada (apta de ser invertível).

Por exemplo, considere uma situação hipotética em que  $K(Q) = 800$ ,  $f(x^1) = 1$  e  $f(x^*) = 0$ . O Teorema 5.3 sugere que o valor da função ainda estará em torno de 0,08 após um milhão de iterações do método da descida máxima com busca linear.

**EXEMPLO 1:** Considere o problema definido pela em (3.31). Minimizar a função  $f$  utilizando o Método da Descida Máxima. A direção de descida para o Exemplo 1 é dada por:

$$p_k = -\nabla f(x) = \begin{pmatrix} -8x_1 + 12x_2 \\ 12x_1 - 4x_2^3 + 12x_2^2 - 30x_2 + 4 \end{pmatrix}. \quad (5.41)$$

A Tabela 5.1 apresenta os cálculos para otimizar a função  $f$  utilizando o método da descida máxima. Observe que o processo foi interrompido na 16ª iteração com o critério de parada  $\|\nabla f(x^k)\| < 0,001$ , visto que  $\|\nabla f(x^{16})\| = 9,71 \times 10^{-4}$ . O ponto  $x^{16}$  é escolhido como uma boa aproximação para  $x^*$  e  $f(x^{16}) = 3,1791 \times 10^{-5}$ . A Figura 5.4 mostra o progresso do método da Descida Máxima aplicado ao Exemplo 1.

**EXEMPLO 2:** Agora, considere o problema de otimização do compressor de multi-estágios definido em (3.36), para o qual tem-se a seguinte direção de descida:

$$p_k = -\nabla f(p) = \begin{pmatrix} -\frac{0,2481}{p_1^{0,7519}} + \frac{0,2481 p_2^{0,2481}}{p_1^{1,2481}} \\ -\frac{0,2481}{p_1^{0,2481} p_2^{0,7519}} + \frac{0,2481 \cdot 64^{0,2481}}{p_2^{1,2481}} \end{pmatrix}. \quad (5.42)$$

A Tabela 5.2 <sup>1</sup> resume os cálculos do método da Descida Máxima para a minimização deste problema.

<sup>1</sup>Foi feita uma adaptação no algoritmo para se obedecer a região viável  $p_1 \geq 1$ . O valor  $p_1 = -7,5392$  obtido na 2ª iteração foi substituído por 1

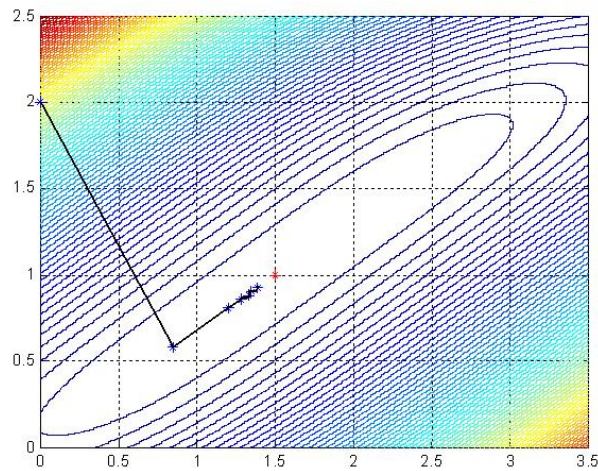


Figura 5.4: Progresso do Método da Descida Máxima aplicado ao Exemplo 1

Iteração k	$(x^k)^T$	$f(x^k)$	$\alpha_k$	$p_k^T$	$\ \nabla f(x^k)\ $
1	(10; 20)	1,2927	$1,2096 \times 10^3$	(-0,0145; 0,0018)	0,0146
2	(1 ; 22,2040)	1,4583	12,9003	(0,2873; -0,0096)	0,2874
3	(4,7062 ; 22,0804)	1,2381	$2,6474 \times 10^3$	(-0,0001; -0,0019)	0,0019
4	(4,5498 ; 17,1642)	1,2326	$1,1958 \times 10^2$	(-0,0036; -0,0001)	0,0036
5	(4,1185 ; 17,1572)	1,2318	0,5	$(0,24; -0,56) \times 10^{-3}$	$6,08 \times 10^{-4}$

Tabela 5.2: Método da Descida Máxima aplicado ao Exemplo 2

Para o Exemplo 2, o Método da Descida Máxima com apenas cinco iterações encontrou o ponto  $p^* = (4,1185; 17,1572)^T$ , no qual a função objetivo vale 1,2318. Note ainda, que  $\|\nabla f(x^5)\| = 6,08 \times 10^{-4}$ , a precisão para este caso é superior ao estabelecido como critério de parada. A Figura 5.5 mostra o progresso do Método da Descida Máxima aplicado ao Exemplo 2 onde verifica-se a rápida convergência para o ótimo.

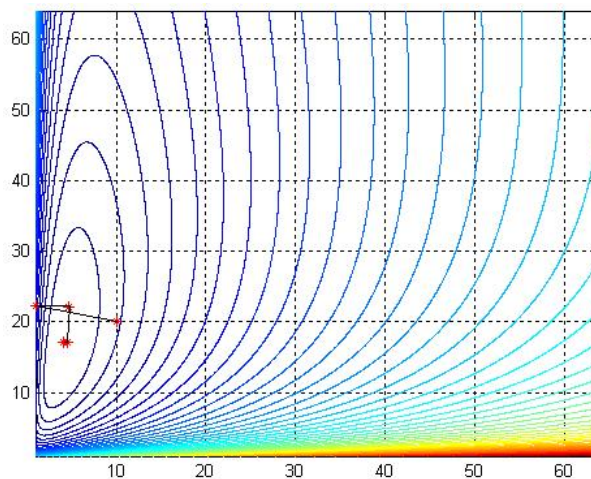


Figura 5.5: Progresso do Método da Descida Máxima aplicado ao Exemplo 2

# Capítulo 6

## Método de Newton

A direção de busca dos Métodos de Newton é muito importante, talvez a mais importante de todas, sendo obtida da aproximação de  $f(x^k + p)$  por uma Série de Taylor de 2º ordem, ou seja:

$$f(x^k + p) \approx f^k + p^T \nabla f^k + \frac{1}{2} p^T \nabla^2 f^k p := m_k(p). \quad (6.1)$$

Assumindo-se que a matriz  $\nabla^2 f^k$  é positiva definida, a direção de Newton é obtida encontrando-se o vetor  $p$  que minimiza  $m_k(p)$ .

Fazendo com que a derivada de  $m_k(p)$  seja zero, obtém-se a seguinte fórmula explícita:

$$\frac{\partial m_k(p)}{\partial p} = \nabla f^k + \nabla^2 f^k p_k = 0, \quad (6.2)$$

ou seja,

$$p_k = -(\nabla^2 f^k)^{-1} \nabla f^k. \quad (6.3)$$

A utilização da direção de Newton torna-se viável quando o erro entre a função verdadeira  $f(x^k + p)$  e sua aproximação dada pelo modelo quadrático  $m_k(p)$  não for muito grande. Comparando as equações (6.1) e (1.7) verifica-se que a única diferença entre essas funções é que a matriz  $\nabla^2 f(x^k + tp)$  foi substituída por  $\nabla^2 f$ , no 3º termo da expansão da série.

Se  $\nabla^2 f$  for suficientemente suave, esta diferença introduz uma perturbação apenas da ordem  $O(\|p\|^3)$  na expansão em série. Assim, quando  $\|p\|$  é pequena, o modelo  $m_k(p)$  aproxima-se de  $f(x^k + p)$  com precisão.

A direção de Newton pode ser usada nos métodos de busca linear quando a matriz  $\nabla^2 f$  é positiva definida. Neste caso, de (6.2) vem:

$$(\nabla f^k)^T p_k = -p_k^T \nabla^2 f^k p_k \leq -\sigma_k \|p_k\|^2 \quad (6.4)$$

para algum  $\sigma_k > 0$ . A menos que o gradiente  $\nabla f^k$  (ou o passo  $p_k$ ) seja nulo, tem-se que  $(\nabla f^k)^T p_k < 0$ . Assim a direção de Newton é uma direção de descida.

Diferentemente da direção de descida máxima, o comprimento do passo  $\alpha$  pode ser associado a 1 na direção de Newton. Desta forma, em muitas implementações do método de Newton para busca linear adota-se  $\alpha = 1$ . O valor de  $\alpha$  é ajustado apenas nos casos em que não ocorre uma redução satisfatória de  $f$ .

Pode-se, usar também, o método da região de confiança de Newton, pois este tem-se mostrado muito eficiente nas aplicações práticas. Ao invés de se escolher a matriz  $B^k = 0$  no modelo  $m_k$  apresentado na equação (2.6), escolhe-se  $B^k$  de forma a ser igual a matriz Hessiana  $\nabla^2 f^k$ . Considerando a restrição da região de confiança  $\|p\|_2 \leq \Delta_k$ , pode-se garantir que o sub-problema (2.5) possui solução, mesmo quando  $\nabla^2 f^k$  não é positiva definida.

Métodos que usam a direção de Newton possuem uma taxa de convergência rápida, tipicamente quadrática. Depois que a vizinhança da solução é alcançada por alguma iteração, a convergência ocorre com alta precisão em poucas iterações. A principal desvantagem deste método é necessitar da inversa da matriz Hessiana  $\nabla^2 f(x)$ . Além disso, conforme a aplicação, a computação das matrizes de segunda ordem podem ser onerosas, propícias a erro e de elevado custo computacional.

## 6.1 Convergência do Método de Newton

Para os métodos do tipo Newton (ou Quase-Newton) a direção de busca é dada na forma

$$p_k = -(B^k)^{-1} \nabla f^k, \quad (6.5)$$

onde  $B^k$  é uma matriz simétrica e não singular. Para o método de Newton,  $B^k$  é exatamente a matriz Hessiana  $\nabla^2 f(x^k)$ .

Considere um método dado por (6.5) e assumamos que a matriz  $B^k$  é positiva definida com uma condição de ser numericamente limitada. Isto é, existe uma constante  $M$  tal que

$$\|B^k\| \|(B^k)^{-1}\| \leq M \text{ para todo } k. \quad (6.6)$$

Pode-se mostrar (Nocedal *et al.* (2000)) que  $\|Bx\| \geq \|x\|/\|B^{-1}\|$  e da condição (5.7) resulta que

$$\cos \theta_k \geq \frac{1}{M}. \quad (6.7)$$

Substituindo este resultado em (5.17) vem

$$\lim_{k \rightarrow \infty} \|\nabla f^k\| = 0. \quad (6.8)$$

Portanto, verifica-se que o método de Newton e os métodos Quase-Newton são globalmente convergentes se as matrizes  $B^k$  são positiva definidas, se possuem uma condição de serem numericamente limitadas e se o comprimento do passo satisfaz a condição de Wolfe.

## 6.2 Taxa de convergência do Método de Newton

Considere a iteração de Newton para a qual a busca é dada por

$$p_k = -(\nabla^2 f^k)^{-1} \nabla f^k. \quad (6.9)$$

Um dos problemas deste método é que uma vez que a matriz Hessiana  $\nabla^2 f^k$  nem sempre é positiva definida, a direção  $p_k$  nem sempre será uma direção de descida e a maioria dos conceitos apresentados na seção anterior não podem ser aplicados. No entanto, nesta seção será discutida a taxa de convergência do método de Newton, considerando que para todo  $x$  vizinho da solução  $x^*$  (sendo  $\nabla^2 f(x^*)$  positiva definida) a matriz  $\nabla^2 f(x)$  também será positiva definida. Nestas condições, o método de Newton será bem definido nesta região e convergirá quadraticamente.

**Teorema 6.1** *Suponha que a função  $f$  seja duas vezes diferenciável e que a Hessiana  $\nabla^2 f(x)$  seja Lipschitz contínua. Então na vizinhança da solução  $x^*$  as condições suficientes de 2º ordem definidas no Teorema 2.3 são satisfeitas. Considere a iteração  $x^{k+1} = x^k + p_k$ , sendo  $p_k$  dado pela equação (6.9), então:*

1. se o ponto de partida  $x^0$  é suficientemente próximo de  $x^*$ , a sequência de iterações  $x^k$  converge para  $x^*$ ;

2. a razão de convergência de  $\{x^k\}$  é quadrática;

3. a sequência de normas do gradiente  $\{\|\nabla f^k\|\}$  converge quadraticamente para zero.

**Demonstração:** A afirmação 1 já foi considerada. Serão mostradas as outras afirmações. Pela definição da iteração de Newton tem-se

$$x^k + p_k - x^* = x^k - x^* - (\nabla^2 f^k)^{-1} \nabla f^k. \quad (6.10)$$

Agora, lembrando que  $\nabla f^* = 0$  e que  $(\nabla^2 f^k)^{-1} \nabla^2 f^k = I$  e substituindo em (6.10) segue que

$$x^k + p_k - x^* = (\nabla^2 f^k)^{-1} [\nabla^2 f^k (x^k - x^*) - (\nabla f^k - \nabla f^*)]. \quad (6.11)$$

Donde,

$$\|x^k + p_k - x^*\| = \|x^{k+1} - x^*\| \leq \|(\nabla^2 f^k)^{-1}\| \|\nabla^2 f^k (x^k - x^*) - (\nabla f^k - \nabla f^*)\|. \quad (6.12)$$

Tem-se

$$\begin{aligned} & \left\| \int_0^1 [\nabla^2 f(x^k) - \nabla^2 f(x^k + t(x^* - x^k))](x^k - x^*) dt \right\| \\ &= \left\| \int_0^1 \nabla^2 f(x^k)(x^k - x^*) dt - \int_0^1 \nabla^2 f(x^k + t(x^* - x^k))(x^k - x^*) dt \right\|. \end{aligned} \quad (6.13)$$

Considerando a seguinte mudança de variável  $x^k + t(x^* - x^k) = m$ , e substituindo em (6.13) resulta:

$$\begin{aligned} & \left\| \int_0^1 [\nabla^2 f(x^k) - \nabla^2 f(x^k + t(x^* - x^k))](x^k - x^*) dt \right\| \\ &= \left\| \nabla^2 f(x^k)(x^k - x^*) + \int_{x^k}^{x^*} \nabla^2 f(m) dm \right\| \\ &= \left\| \nabla^2 f(x^k)(x^k - x^*) + \nabla f(x^*) - \nabla f(x^k) \right\|. \end{aligned} \quad (6.14)$$

Assim, observando (6.14), verifica-se que

$$\begin{aligned} & \left\| \nabla^2 f(x^k)(x^k - x^*) - (\nabla f(x^k) - \nabla f(x^*)) \right\| \\ &= \left\| \int_0^1 [\nabla^2 f(x^k) - \nabla^2 f(x^k + t(x^* - x^k))](x^k - x^*) dt \right\| \\ &\leq \int_0^1 \|\nabla^2 f(x^k) - \nabla^2 f(x^k + t(x^* - x^k))\| \|x^k - x^*\| dt \\ &\leq \int_0^1 L \|x^k + t(x^* - x^k) - x^k\| \|x^k - x^*\| dt \\ &\leq \int_0^1 L \|x^* - x^k\| \|x^k - x^*\| t dt \\ &\leq L \|x^k - x^*\|^2 \int_0^1 t dt, \end{aligned}$$

sendo  $L$  a constante de Lipschitz para  $\nabla^2 f(x)$  para  $x$  próximo  $x^*$ . Portanto,

$$\|\nabla^2 f^k(x^k - x^*) - (\nabla f^k - \nabla f^*)\| \leq \frac{1}{2} L \|x^k - x^*\|^2. \quad (6.15)$$

Substituindo (6.15) em (6.12) resulta

$$\|x_{k+1} - x^*\| \leq \frac{1}{2} L \|(\nabla^2 f^k)^{-1}\| \|x^k - x^*\|^2. \quad (6.16)$$

Uma vez que  $\nabla^2 f(x^*)$  é não-singular, existe um raio  $r > 0$  tal que

$$\|(\nabla^2 f^k)^{-1}\| \leq 2\|\nabla^2 f(x^*)^{-1}\| \quad (6.17)$$

para todo  $x^k$  com  $\|x^k - x^*\| \leq r$ .

Assim, a expressão (6.16) pode ser escrita como:

$$\|x^{k+1} - x^*\| \leq L\|\nabla^2 f(x^*)^{-1}\| \|x^k - x^*\|^2 = \tilde{L}\|x^k - x^*\|^2, \quad (6.18)$$

sendo que  $\tilde{L} = L\|\nabla^2 f(x^*)^{-1}\|$ .

Usando a inequação (6.18), deduz-se que a seqüência  $x^k$  converge para  $x^*$  e a razão de convergência é quadrática.

Sabe-se que,  $x^{k+1} - x^k = p_k$  e  $\nabla f^k + \nabla^2 f^k p_k = 0$ . Então pode-se escrever:

$$\begin{aligned} \|\nabla f^{k+1}\| &= \|\nabla f^{k+1} - (\nabla f^k + \nabla^2 f^k p_k)\| \\ &= \|\nabla f^{k+1} - \nabla f^k - \nabla^2 f^k p_k\| \\ &= \left\| \int_0^1 \nabla^2 f(x^k + tp_k)(x^{k+1} - x^k) dt - \nabla^2 f(x^k) p_k \right\| \\ &= \left\| \int_0^1 \nabla^2 f(x^k + tp_k)(x^{k+1} - x^k) dt - \int_0^1 (\nabla^2 f(x^k) p_k) dt \right\| \\ &= \left\| \int_0^1 [\nabla^2 f(x^k + tp_k) - \nabla^2 f(x^k)] p_k dt \right\| \\ &\leq \int_0^1 \|\nabla^2 f(x^k + tp_k) - \nabla^2 f(x^k)\| \|p_k\| dt \\ &\leq \int_0^1 L\|x^k + tp_k - x^k\| \|p_k\| dt \\ &\leq \int_0^1 L\|p_k\|^2 t dt \\ &\leq L\|p_k\|^2 \int_0^1 t dt \\ &\leq \frac{1}{2} L\|p_k\|^2 \\ &\leq \frac{1}{2} L\| -(\nabla^2 f^k)^{-1} \nabla f^k \|^2 \\ &\leq \frac{1}{2} L\|(\nabla^2 f^k)^{-1}\|^2 \|\nabla f^k\|^2. \end{aligned} \quad (6.19)$$

Aqui, novamente, foi usado a mudança de variável  $x^k + t(x^{k+1} - x^k) = m$  para chegar na igualdade  $\nabla f^{k+1} - \nabla f^k = \int_0^1 \nabla^2 f(x^k + tp_k)(x^{k+1} - x^k) dt$ .

Logo, tem-se que

$$\|\nabla f^{k+1}\| \leq \frac{1}{2} L\|(\nabla^2 f^k)^{-1}\|^2 \|\nabla f^k\|^2. \quad (6.20)$$

Mas da desigualdade (6.17) resulta

$$\|\nabla f(x^{k+1})\| \leq L\|\nabla^2 f(x^*)^{-1}\|^2 \|\nabla f(x^k)\|^2, \quad (6.21)$$

ou seja,

$$\|\nabla f^{k+1}\| \leq c_1 \|\nabla f^k\|^2, \quad (6.22)$$

provando que a norma do gradiente converge quadraticamente para zero.  $\square$

Iteração k	$(x^k)^T$	$f(x^k)$	$\alpha_k$	$p_k^T$	$\ \nabla f(x^k)\ $
1	(0; 2)	37	1,0054	(2,5; -0,3333)	46,6476
2	(2,5136 ; 1,6648)	0,1961	2,3344	(-0,3487; -0,2216)	0,9872
3	(1,6995 ; 1,1415)	0,0024	1,0459	(-0,0520; -0,0492)	0,1934
4	(1,6451 ; 1,0961)	$8,9242 \times 10^{-5}$	1,8117	(-0,0490; -0,0320)	0,0109
5	(1,5563 ; 1,0381)	$4,7299 \times 10^{-6}$	1,1077	(-0,0182; -0,0127)	0,0124
6	(1,5361 ; 1,0240)	$3,6270 \times 10^{-7}$	0,5	(-0,0121; -0,0080)	0,0014
7	(1,5300 ; 1,0200)	$1,6782 \times 10^{-7}$	0,5	(-0,0100; -0,0067)	$3,2 \times 10^{-5}$

Tabela 6.1: Método de Newton aplicado ao Exemplo 1

Nos seguintes exemplos o valor de  $\alpha$  será obtido fazendo uma busca unidimensional a cada iteração ao invés de se adotar  $\alpha = 1$ .

**EXEMPLO 1:** Considere o problema definido em (3.31). Deseja-se minimizar  $f$  usando o método de Newton e partindo do ponto inicial  $x^0 = (0; 2)^T$ . A direção de busca é dada por

$$p_k = - \begin{bmatrix} 8 & -12 \\ -12 & 12x_2^2 - 24x_2 + 30 \end{bmatrix}^{-1} \begin{bmatrix} 8x_1 - 12x_2 \\ -12x_1 + 4x_2^3 - 12x_2^2 + 30x_2 - 4 \end{bmatrix}. \quad (6.23)$$

A Tabela 6.1 resume os cálculos do método e a Figura 6.1 mostra o progresso do método. Na sétima iteração o processo foi interrompido com o critério de parada  $\|\nabla f(x^k)\| < 0,001$  obtendo-se aproximação ótima  $x^7 = (1,53; 1,02)^T$  no qual a função objetivo vale  $1,6782 \times 10^{-7}$ . Tem-se que  $\nabla f(x^7) = (0,3494; -0,4920) \times 10^{-3}$ , donde  $\|\nabla f(x^7)\| = 3,2 \times 10^{-5}$ .

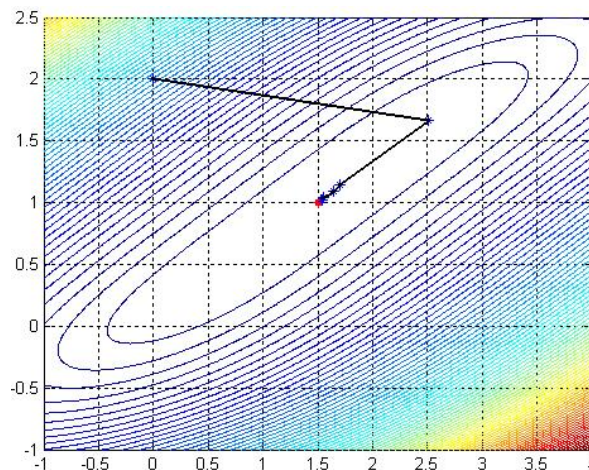


Figura 6.1: Progresso do Método de Newton aplicado ao Exemplo 1

**EXEMPLO 2:** No caso da otimização das pressões intermediárias do compressor multi-estágio, problema (3.36), usando o Método de Newton e o ponto inicial  $x^0 = (10; 20)^T$ , a direção de busca é dada por:

$$p_k = - \begin{bmatrix} \frac{-0,1865}{p_1^{1,7519}} + \frac{0,3096p_2^{0,2481}}{p_1^{2,2481}} & -\frac{0,0615}{p_1^{1,2481}p_2^{0,7519}} \\ -\frac{0,0615}{p_1^{1,2481}p_2^{0,7519}} & -\frac{0,1865}{p_1^{0,2481}p_2^{1,7519}} + \frac{0,8689}{p_2^{2,2481}} \end{bmatrix}^{-1} \begin{bmatrix} \frac{0,2481}{p_1^{0,7519}} - \frac{0,2481p_2^{0,2481}}{p_1^{1,2481}} \\ \frac{0,2481}{p_1^{0,2481}p_2^{0,7519}} - \frac{0,6962}{p_2^{1,2481}} \end{bmatrix}. \quad (6.24)$$

Iteração k	$(x^k)^T$	$f(x^k)$	$\alpha_k$	$p_k^T$	$\ \nabla f(x^k)\ $
1	(10; 20)	1,2927	0,1761	(0,4413; 4,1377)	0,0146
2	(10,0777 ; 20,7286)	1,2926	-0,0784	(0,4411; 3,8537)	0,0143
⋮	⋮	⋮	⋮	⋮	
10	(10,0534 ; 20,5199)	1,2926	$-8,7423 \times 10^{-5}$	(0,4410; 3,9374)	0,0144
50	(10,0534 ; 20,5196)	1,2926	$7,7246 \times 10^{-6}$	(0,4410; 3,9375)	0,0144
100	(10,0534 ; 20,5196)	1,2926	$-2,1372 \times 10^{-6}$	(0,4410; 3,9375)	0,0144
500	(10,0534 ; 20,5196)	1,2926	$7,4550 \times 10^{-6}$	(0,4410; 3,9375)	0,0144
1000	(10,0534 ; 20,5196)	1,2926	$-1,1561 \times 10^{-5}$	(0,4410; 3,9375)	0,0144

Tabela 6.2: Método de Newton aplicado ao Exemplo 2

A Tabela 6.2 resume os cálculos. Note que para 100, 500 e 1000 iterações o método não consegue sair do lugar, isto ocorre devido ao mal condicionamento da matriz Hessiana relacionada à função objetivo. Verifica-se que o valor do determinante da matriz Hessiana da função  $f$  no ponto  $(10,0534; 20,5196)^T$  vale  $-1,3178 \times 10^{-5}$  que é muito próximo de zero, donde a obtenção da inversa da matriz Hessiana se torna muito difícil e o algoritmo encontra dificuldade para obter uma boa aproximação. Para outros pontos iniciais o problema do mal condicionamento da matriz Hessiana persistiu e por isso o Método de Newton não é uma boa escolha para resolver este problema.

Será visto no próximo capítulo que o Método Quase-Newton utilizando uma aproximação da matriz Hessiana evita o problema relacionado ao mal condicionamento da matriz Hessiana e será uma boa opção para este tipo de problema.

# Capítulo 7

## Métodos Quase-Newton

Em torno de 1950, W.C. Davidon, um físico do Argonne National Laboratory, usava o método das descidas coordenadas para solucionar um grande problema de otimização que exigia muitos cálculos. Nesta época os computadores eram muito instáveis e para frustração de Davidon, os computadores sempre falhavam antes de concluir os cálculos. Desta forma, Davidon decidiu encontrar uma forma para acelerar as iterações. O algoritmo que ele desenvolveu (o primeiro algoritmo Quase-Newton) mostrou-se como uma das idéias mais criativas da otimização não-linear. Como foi demonstrado por Fletcher e Powell (1963) este novo algoritmo era muito mais rápido e eficiente do que os outros métodos. Este dramático avanço transformou a otimização não-linear de um dia para outro. Durante os 20 anos seguintes, várias modificações para o método foram propostas e centenas de artigos escritos sobre este assunto. Uma história interessante e irônica é que o artigo de Davidon (1959) *não* foi aceito para publicação, permanecendo como um “relatório técnico” por mais de 30 anos até que foi publicado no primeiro número do SIAM Journal on Optimization em 1991.

Os métodos Quase-Newton, assim como a descida máxima, requerem que apenas o gradiente da função objetivo seja fornecido em cada iteração. Através de medida de mudanças no gradiente, contruiu-se um modelo de função objetivo boa o suficiente para produzir convergência super-linear. O aperfeiçoamento obtido sobre o método da descida máxima é dramático, especialmente para problemas difíceis.

Além disso, uma vez que as segundas derivadas não são necessárias, os métodos Quase-Newton algumas vezes são mais eficientes do que os métodos de Newton. Hoje em dia, os códigos comerciais de otimização contém bibliotecas contendo vários tipos deste método para solucionar problemas irrestritos, restritos, de grande escala, etc.

Com o avanço das técnicas numéricas de diferenciação, o método de Newton tornou-se novamente interessante (pois permitem obter as segundas derivadas). Mas, estas ferramentas de diferenciação não podem ser aplicadas em muitas situações, além de exigir alto custo computacional. Por estas razões, os métodos Quase-Newton continuam bastante atrativos.

### 7.1 Direção de busca dos Métodos de Quase-Newton

A direção de busca dos métodos de Quase-Newton tornou-se de uma alternativa atrativa para o método de Newton, pois não requerem o cálculo de Hessiana e ainda mantém uma taxa de convergência superior à linear (“quase-quadrática”). No lugar de calcular  $\nabla^2 f^k$ , utiliza-se a aproximação  $B^k$  (matriz pseudo-hessiana), a qual é atualizada após cada passo levando-se em conta o conhecimento adicional adquirido durante o passo. Na atualização de  $B^k$  utiliza-se o fato de que as variações do gradiente traz informações sobre a segunda derivada de  $f$  ao longo da direção de busca.

Usando a equação (1.6), adicionando e subtraindo o termo  $\nabla^2 f(x)p$ , vem:

$$\nabla f(x+p) = \nabla f(x) + \nabla^2 f(x)p + \int_0^1 [\nabla^2 f(x+tp) - \nabla^2 f(x)]p dt. \quad (7.1)$$

Como  $\nabla f(\cdot)$  é contínuo, o tamanho do termo final de integração é de ordem  $\theta(\|p\|)$ . Fazendo  $x = x^k$  e  $p = x^{k+1} - x^k$ , obtém-se:

$$\nabla f^{k+1} = \nabla f^k + \nabla^2 f^k(x^{k+1} - x^k) + \theta(\|x^{k+1} - x^k\|). \quad (7.2)$$

Quando  $x^k$  e  $x^{k+1}$  estão alinhados em uma região próxima à solução  $x^*$ , na qual  $\nabla^2 f$  é positiva definida, o termo final da expansão é eventualmente dominado por  $\nabla^2 f^k(x^{k+1} - x^k)$  e pode ser escrito como:

$$\nabla^2 f^k(x^{k+1} - x^k) \approx \nabla f^{k+1} - \nabla f^k. \quad (7.3)$$

Considere que seja adotada uma aproximação  $B^{k+1}$  para a matriz Hessiana, de forma que ela obedeça a propriedade (7.3) da matriz Hessiana verdadeira. Isto requer que a seguinte condição, conhecida por *Equação da Secante*, seja satisfeita:

$$B^{k+1} s_k = y_k \quad (7.4)$$

sendo,

$$s_k = x^{k+1} - x^k \quad e \quad y_k = \nabla f^{k+1} - \nabla f^k. \quad (7.5)$$

A direção Quase-Newton é obtida substituindo  $B^k$  na equação(6.3), isto é,

$$p_k = -(B^k)^{-1} \nabla f^k. \quad (7.6)$$

Em algumas implementações práticas os métodos de Quase-Newton atualizam, a cada passo, a inversa  $(B^k)^{-1}$  no lugar de atualizar  $B^k$ , reduzindo o custo computacional. Assim, uma fórmula que utiliza uma aproximação para a inversa da Hessiana (matriz pseudo-Hessiana inversa) é dada por:

$$H^k := (B^k)^{-1} \quad e \quad H^{k+1} = (I - \rho_k s_k y_k^T) H^k (I - \rho_k y_k s_k^T) + \rho_k y_k y_k^T, \quad (7.7)$$

sendo  $\rho_k = \frac{1}{y_k^T s_k}$  (Nocedal *et al.* (2000)).

Observe que a implementação da equação (7.7) é simples, necessitando apenas do produto de matrizes e vetores (muito mais fácil do que implementar o cálculo de matrizes inversas), e a direção de busca é dada por:

$$p_k = -H^k \nabla f^k. \quad (7.8)$$

A direção Quase-Newton possui convergência super-linear e suas vantagens são que, além de não necessitar do cálculo das segundas derivadas, a inversa da matriz pseudo-Hessiana é obtida por meio de um processo iterativo.

Observe que, novamente pode-se utilizar região de confiança nos métodos Quase-Newton, sendo sua aplicação análoga ao uso da região de confiança para o método de Newton, porém, a matriz  $B^k$ , agora é definida pela matriz pseudo-Hessiana.

## 7.2 Convergência dos Métodos Quase-Newton

Como foi comentado na Seção 6.1 os métodos Quase-Newton são globalmente convergentes se as matrizes  $B^k$  são positivas definidas e possuem a condição de serem numericamente limitadas, e se o comprimento do passo satisfaz a condição de Wolfe.

### 7.3 Taxas de convergência dos Métodos Quase-Newton

Suponha então a seguinte direção de busca

$$p_k = -(B^k)^{-1} \nabla f^k \quad (7.9)$$

onde a matriz simétrica e positiva  $B^k$  é atualizada a cada iteração pela fórmula Quase-Newton que será apresentada na próxima seção.

Os resultados seguintes, apresentados em Nocedal *et al* (2000), mostram que se a direção de busca Quase-Newton aproxima bem a direção de Newton, a condição de comprimento de passo  $\alpha_k$  satisfará a condição de Wolfe e as iterações convergirão para a solução.

**Teorema 7.1** *Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  seja duas vezes continuamente diferenciável. Considere a iteração  $x^{k+1} = x^k + \alpha_k p_k$ , onde  $p_k$  é a direção de descida máxima e  $\alpha_k$  satisfaz as condições de Wolfe (3.8) e (3.9) com  $c_1 \leq \frac{1}{2}$ . Se a seqüência  $\{x^k\}$  converge para o ponto  $x^*$  tal que  $\nabla f(x^*) = 0$  e  $\nabla^2 f(x^*)$  é positiva definida, e se a direção de busca satisfaz*

$$\lim_{k \rightarrow \infty} \frac{\|\nabla f^k + \nabla^2 f^k p_k\|}{\|p_k\|} = 0 \quad (7.10)$$

então

1. o comprimento de passo  $\alpha_k = 1$  é admissível para todo  $k$  maior que algum índice  $k_0$ ;
2. se  $\alpha_k = 1$  para todo  $k > k_0$ ,  $\{x^k\}$  converge para  $x^*$  super-linearmente.

Note que se  $c_1 > \frac{1}{2}$ , então a busca linear excluiu a minimização quadrática, e o comprimento de passo unitário pode não ser admissível.

Se  $p_k$  é uma direção Quase-Newton da forma (7.9), então (7.10) é equivalente a

$$\lim_{k \rightarrow \infty} \frac{\|(B^k - \nabla^2 f(x^*))p_k\|}{\|p_k\|} = 0, \quad (7.11)$$

visto que,  $\nabla f^k = -B^k p_k$ , logo  $\|\nabla f^k + \nabla^2 f^k p_k\| = \|(B^k - \nabla^2 f^k)p_k\|$ .

Assim, tem-se o surpreendente (e feliz) resultado de que a taxa de convergência super-linear pode ser obtida mesmo se a seqüência de matrizes Quase-Newton  $B^k$  não seja convergente para  $\nabla^2 f(x^*)$  ao longo da direção  $p_k$ . É importante observar que a condição (7.11) é suficiente e necessária para garantir a super-linearidade dos métodos Quase-Newton.

**Teorema 7.2** *Suponha que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  seja duas vezes continuamente diferenciável. Considere a iteração  $x^{k+1} = x^k + p_k$  e que  $p_k = -(B^k)^{-1} \nabla f^k$ . Assuma também que  $\{x^k\}$  converge para o ponto  $x^*$  no qual  $\nabla f(x^*) = 0$  e  $\nabla^2 f(x^*)$  é positiva definida. Então  $\{x^k\}$  converge super-linearmente se e somente se (7.11) é válida.*

**Demonstração:** Mostra-se inicialmente que (7.11) é equivalente a

$$p_k - p_k^N = O(\|p_k\|), \quad (7.12)$$

sendo  $p_k^N = -(\nabla^2 f^k)^{-1} \nabla f^k$  a direção de Newton. Assumindo que (7.11) seja válida e lembrando que  $\nabla f^k = -B^k p_k$ , tem-se

$$\begin{aligned}
p_k - p_k^N &= p_k + (\nabla^2 f^k)^{-1} \nabla f^k \\
&= (\nabla^2 f^k)^{-1} (\nabla^2 f^k p_k + \nabla f^k) \\
&= (\nabla^2 f^k)^{-1} (\nabla^2 f^k p_k - B^k p_k) \\
&= (\nabla^2 f^k)^{-1} (\nabla^2 f^k - B^k) p_k \\
&= O(\|(\nabla^2 f^k - B^k) p_k\|) \\
&= O(\|p_k\|)
\end{aligned}$$

onde foi utilizado o fato de que  $\|(\nabla^2 f^k)^{-1}\|$  é limitada para  $x$  suficientemente próximo de  $x^*$ , uma vez que a Hessiana  $\nabla^2 f(x^*)$  é positiva definida. A recíproca aparece ao se multiplicar ambos os lados de (7.12) por  $\nabla^2 f^k$  e reutilizar (7.9).

Combinando (6.18) e (7.12), obtém-se

$$\|x^k + p_k - x^*\| = \|x^k + p_k^N - x^* + p_k - p_k^N\| \leq \|x^k + p_k^N - x^*\| + \|p_k - p_k^N\| = O(\|x^k - x^*\|^2) + O(\|p_k\|).$$

Uma simples manipulação desta desigualdade mostra que  $\|p_k\| = O(\|x^k - x^*\|)$ , assim

$$\|x^k + p_k - x^*\| \leq O(\|x^k - x^*\|)$$

mostrando a convergência super-linear. □

## 7.4 Os métodos BFGS e DFP

O método BFGS é o mais popular dos métodos Quase-Newton, em homenagem aos criadores Broydon, Fletcher, Goldfarb e Shanno (Broydon (1970), Fletcher (1970), Goldfarb (1970) e Shanno (1970)). Este algoritmo é muito similar ao DFP (Davidon-Fletcher-Powell) apresentado em Davidon (1959) e Fletcher e Powell (1963).

Considere o seguinte modelo quadrático da função objetivo na iteração atual  $x^k$ :

$$m_k(p) = f^k + (\nabla f^k)^T p + \frac{1}{2} p^T B^k p, \quad (7.13)$$

onde  $B^k$  é uma matriz  $n \times n$  simétrica positiva definida a qual deve ser atualizada a cada iteração. Observe que, para  $p = 0$ ,  $m_k = f^k$  e  $\nabla m_k = \nabla f^k$ . O minimizador  $p_k$  deste modelo quadrático convexo pode ser escrito explicitamente como

$$p_k = -(B^k)^{-1} \nabla f^k, \quad (7.14)$$

o qual é utilizado como direção de busca, de modo que a nova iteração é:

$$x^{k+1} = x^k + \alpha_k p_k, \quad (7.15)$$

sendo que o passo  $\alpha_k$  é escolhido de forma a satisfazer as condições de Wolfe.

Este processo é similar ao método de Newton, onde a principal diferença é que a aproximação da matriz Hessiana  $B^k$  é utilizada no lugar da verdadeira Hessiana  $\nabla^2 f^k$ .

No lugar de calcular  $B^k$  em todas as iterações, Davidon propôs que ela fosse atualizada de uma maneira simples levando em consideração a medida de curvatura durante o último passo.

Suponha que foi calculada a iteração  $x^{k+1}$  e deseja-se construir um novo modelo quadrático de forma

$$m_{k+1}(p) = f^{k+1} + (\nabla f^{k+1})^T p + \frac{1}{2} p^T B^{k+1} p. \quad (7.16)$$

Que condições podem ser impostas sobre  $B^{k+1}$  baseadas no conhecimento acumulado durante o último passo?

Uma condição razoável é que o gradiente para  $m_{k+1}$  deve ser equivalente ao gradiente da função objetivo ao menos para as duas iterações  $x^k$  e  $x^{k+1}$ . Uma vez que  $\nabla m_{k+1} = \nabla f^{k+1}$  para  $p = 0$ , a segunda condição é satisfeita automaticamente. A primeira condição (em  $x^k$ ) pode ser escrita matematicamente como:

$$\nabla m_{k+1}(-\alpha_k p_k) = \nabla f^{k+1} - \alpha_k B^{k+1} p_k = \nabla f^k. \quad (7.17)$$

Reescrevendo, obtém-se:

$$B^{k+1} \alpha_k p_k = \nabla f^{k+1} - \nabla f^k. \quad (7.18)$$

Para simplificar a equação, utiliza-se os vetores definidos em (7.5):

$$s_k = x^{k+1} - x^k = \alpha_k p_k \quad e \quad y_k = \nabla f^{k+1} - \nabla f^k \quad (7.19)$$

de modo que, (7.18) se torna:

$$B^{k+1} s_k = y_k. \quad (7.20)$$

Esta fórmula, como mencionado anteriormente, é conhecida como *equação da secante*, na qual a aproximação da Hessiana depende de um deslocamento  $s_k$  e da variação do gradiente  $y_k$ .

Dado um deslocamento  $s_k$  e uma variação do gradiente  $y_k$ , a equação da secante requer que  $y_k$  seja uma aplicação da matriz simétrica positiva definida  $B^{k+1}$  em  $s_k$ . Isto só será possível se  $s_k$  e  $y_k$  satisfizerem a condição de curvatura

$$s_k^T y_k > 0, \quad (7.21)$$

como é comprovado pré-multiplicando (7.20) por  $s_k^T$ , o que fornece  $s_k^T y_k = s_k^T B^{k+1} s_k > 0$ .

Quando  $f$  é fortemente convexa, a desigualdade (7.21) é satisfeita para quaisquer dois pontos  $x^k$  e  $x^{k+1}$ . Entretanto, esta condição pode não ser sempre válida no caso de funções não-convexas.

Neste caso, para forçar que (7.21) seja obedecida explicitamente, são impostas condições sobre o procedimento de busca linear para escolha do passo  $\alpha$ . De fato, a condição (7.21) é garantida como válida se são impostas as condições de Wolfe. Para verificar esta afirmação, de (7.19) e (3.9) tem-se

$$(\nabla f^{k+1})^T s_k \geq c_2 (\nabla f^k)^T s_k \quad e \quad assim, \quad y_k^T s_k \geq (c_2 - 1) \alpha_k (\nabla f^k)^T p_k. \quad (7.22)$$

Uma vez que  $c_2 < 1$  e  $p_k$  é uma direção de descida, o termo à direita é positivo e a condição de curvatura (7.21) é válida.

Quando a condição de curvatura é satisfeita, a equação da secante (7.20) sempre possui uma solução  $B^{k+1}$ . De fato, ela admite um número infinito de soluções, já que  $n(n+1)/2$  graus de liberdade em uma matriz simétrica positiva definida excede as  $n$  condições impostas pela equação da secante. A condição de ser positiva definida impõe  $n$  desigualdades adicionais, mas estas condições não absorvem os graus de liberdade restantes.

Para determinar uma solução única  $B^{k+1}$ , impõe-se a condição adicional de que entre todas as matrizes simétricas que satisfazem a equação da secante,  $B^{k+1}$ , é, de alguma forma, a mais próxima da matriz atual  $B^k$ . Em outras palavras, deve-se resolver o problema

$$\min_B \|B - B^k\| \quad (7.23)$$

sujeito a

$$B = B^T, \quad B s_k = y_k, \quad (7.24)$$

sendo que  $s_k$  e  $y_k$  satisfazem (7.21) e  $B^k$  é simétrica positiva definida. Diferentes normas de matrizes podem ser usadas em (7.23), e cada norma gera um método Quase-Newton diferente. A norma que permite a solução mais fácil do problema de minimização (7.23) e resulta em um método de otimização que não depende da escala (scale-invariant) é a *norma de Frobenius* com pesos:

$$\|A\|_W = \|W^{1/2} A W^{1/2}\|_F, \quad (7.25)$$

onde  $\|\cdot\|_F$  é definida por  $\|C\|_F^2 = \sum_{i=1}^n \sum_{j=1}^n C_{ij}^2$ . A matriz peso  $W$  pode ser escolhida como alguma matriz que satisfaz a relação  $W y_k = s_k$ . De forma prática, o usuário pode assumir  $W = G_k^{-1}$ , onde  $G_k$  é a *Hessiana média* definida por:

$$G_k = \left[ \int_0^1 \nabla^2 f(x^k + \sigma \alpha_k p_k) d\sigma \right] \quad (7.26)$$

que possui a propriedade

$$y_k = G_k \alpha_k p_k = G_k s_k, \quad (7.27)$$

segundo o Teorema de Taylor 1.1. Com esta escolha para a matriz peso, a norma (7.25) é adimensional, o que é uma propriedade desejável uma vez que não queremos que a solução do problema (7.23) e (7.24) dependa das unidades do problema.

No método *DFP*, utilizando a matriz peso e a norma definidos no parágrafo anterior, obtém-se a solução única do problema (7.23) e (7.24) dado por:

$$B^{k+1} = (I - \rho_k y_k s_k^T) B^k (I - \rho_k s_k y_k^T) + \rho_k y_k y_k^T \quad (7.28)$$

com

$$\rho_k = \frac{1}{y_k^T s_k}. \quad (7.29)$$

Esta fórmula conhecida como fórmula de atualização *DFP*, foi proposta originalmente por Davidon em 1959, sendo posteriormente estudada e popularizada por Fletcher e Powell.

A matriz inversa de  $B^k$ , aproximação da inversa da Hessiana, que é representada por

$$H^k = (B^k)^{-1}, \quad (7.30)$$

é útil na implementação do método uma vez que permite calcular a direção de busca (7.14) usando apenas multiplicações simples de matrizes e vetores. Para a aproximação numérica desta inversa considere a observação abaixo:

**Observação 7.1** Segundo *Sherman-Morrison-Woodberry* se uma matriz quadrada não-singular  $A$  é atualizada para tornar-se

$$\hat{A} = A + ab^T,$$

sendo que  $a, b \in \mathbb{R}^n$ , então se  $\hat{A}$  é não-singular, tem-se

$$\hat{A}^{-1} = A^{-1} - \frac{A^{-1} a b^T A^{-1}}{1 + b^T A^{-1} a}.$$

De fato, pois,

$$\begin{aligned} \hat{A} \cdot \hat{A}^{-1} &= (A + ab^T) \left( A^{-1} - \frac{A^{-1} a b^T A^{-1}}{1 + b^T A^{-1} a} \right) \\ &= I - \frac{ab^T A^{-1}}{1 + b^T A^{-1} a} + ab^T A^{-1} - \frac{ab^T A^{-1} ab^T A^{-1}}{1 + b^T A^{-1} a} \\ &= I + \frac{-ab^T A^{-1} + ab^T A^{-1} + ab^T A^{-1} b^T A^{-1} a - ab^T A^{-1} ab^T A^{-1}}{1 + b^T A^{-1} a} \\ &= I. \end{aligned}$$

Assim, utilizando a fórmula de Sherman-Morrison-Woodberry (para matrizes não singulares) é possível escrever uma expressão para atualizar uma aproximação para a inversa de Hessiana, como:

$$H^{k+1} = H^k - \frac{H^k y_k y_k^T H^k}{y_k^T H^k y_k} + \frac{s_k s_k^T}{y_k^T s_k}. \quad (7.31)$$

A matriz (7.31) é utilizada pelo método *DFP* para atualizar  $H^{k+1}$ .

Observe que os últimos dois termos do lado direito da equação (7.31) são matrizes de posto-um, de modo que  $H^k$  é submetida a modificações de posto-dois. É fácil ver que (7.28) também é uma modificação de posto-dois sobre  $B^k$ .

A idéia fundamental de atualização dos métodos Quase-Newton é que no lugar de calcular a inversa da matriz Hessiana a cada iteração, é feita uma modificação simples que combina as informações observadas recentemente na função objetivo com o conhecimento embutido na aproximação atual da matriz Hessiana.

A fórmula de atualização do *DFP* é eficiente, mas foi rapidamente suplantada pela fórmula *BFGS*, a qual é considerada a mais eficiente entre todos os métodos Quase-Newton.

A fórmula *BFGS* pode ser obtida fazendo uma modificação simples no conceito que levou a (7.28). No lugar de impor condições sobre a aproximação de Hessiana  $B^k$ , impõe-se condições similares à sua inversa  $H^k$ .

A aproximação de inversa da Hessiana  $H^k$ , deve ser simétrica positiva definida e deve obedecer a equação da secante (7.20), agora escrita como

$$H^{k+1} y_k = s_k. \quad (7.32)$$

A condição da solução ser a mais próxima de  $H^k$ , é definida de forma análoga a (7.23) e (7.24), por

$$\min_H \|H - H^k\| \quad \text{sujeito a } H = H^T, \quad H y_k = s_k. \quad (7.33)$$

A norma é novamente considerada como a norma de Frobenius com pesos, sendo que a matriz peso  $W$  é agora uma matriz qualquer que satisfaça  $W s_k = y_k$ , (na prática, assume-se que  $W$  é dado pela Hessiana média  $G_k$  definida em (7.26)).

Desta forma, a solução única  $H^{k+1}$  do problema (7.33), pelo método *BFGS* é dado por

$$H^{k+1} = (I - \rho_k s_k y_k^T) H^k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T, \quad (7.34)$$

com  $\rho_k$  definido em (7.29).

Existe uma questão importante a ser resolvida antes de definir o algoritmo *BFGS* completo: como escolher a aproximação inicial  $H^0$ ? Não existe uma resposta mágica que funcione bem para todos os casos. São necessários informações específicas sobre o problema. Por exemplo, pode-se calcular a inversa da matriz Hessiana por diferenças finitas para o ponto inicial  $x^0$ .

Outra alternativa usual é simplesmente escolher  $H^0$  como a matriz identidade ou múltiplos da matriz identidade, onde estes múltiplos refletem a ordem da escala das variáveis.

O algoritmo *BFGS* é aplicado com um custo de ordem  $O(n^2)$  relativo a operações aritméticas, é um algoritmo robusto e sua taxa de convergência é super-linear, sendo rápido o suficiente para a grande maioria das aplicações. Apesar de o método de Newton possuir convergência quadrática, seu custo por iteração é elevado, pois necessita das segundas derivadas e da solução de um sistema linear.

Uma versão do método *BFGS* trabalhando com a aproximação, da Hessiana  $B^k$  e não com  $H^k$ , pode ser obtida através de uma simples aplicação da fórmula Sherman-Morrison-Woodberry:

$$B^{k+1} = B^k - \frac{B^k s_k s_k^T B^k}{s_k^T B^k s_k} + \frac{y_k y_k^T}{y_k^T s_k}. \quad (7.35)$$

Uma implementação simples desta variante do método não é eficiente para minimização irrestrita, uma vez que necessita de resolver o sistema  $B^k p_k = -\nabla f^k$  no ponto  $p_k$  e desta forma, o custo computacional do passo é acrescido para  $O(n^3)$ .

Segundo Vanderplaats (1999) os métodos Quase-Newton, também conhecidos como Método de Variável Métrica, são os mais utilizados na otimização sequencial.

Sua grande vantagem é melhorar a eficiência do processo ao considerar informações das iterações anteriores. Os métodos mais aplicados *DFP* e *BFGS* podem ser resumidos na expressão:

$$H^{k+1} = H^k + \left[ \frac{s_k^T y_k + \theta (y_k^T H^k y_k)}{(s_k^T y_k)^2} \right] s_k s_k^T + \frac{(\theta - 1)}{y_k^T H^k y_k} H^k y_k (H^k y_k)^T - \frac{\theta}{s_k^T y_k} [H^k y_k s_k^T + s_k (H^k y_k)^T], \quad (7.36)$$

sendo que,

$$y_k = \nabla f^{k+1} - \nabla f^k, \quad s_k = x^{k+1} - x^k, \quad H^0 = I, \quad H \cong (\nabla^2 f^k)^{-1}, \quad x^{k+1} = x^k + \alpha^* (-H^k \nabla f^k). \quad (7.37)$$

Para  $\theta = 0$  tem-se o método *DFP* e para  $\theta = 1$  tem-se o método *BFGS*.

Para exemplificar todos os passos do método descrito em (7.36), considere  $f = x^2 - 3xy + 4y^2 + x - y$ . Deseja-se aproximar o ponto mínimo até que  $\|\nabla f^k\| < 0,5 \cdot 10^{-2}$ , usando o método *DFP* e a seção áurea para a busca unidimensional, dado  $x^0 = (2; 2)^T$ .

$$\text{Tem-se que } f^0 = 8, \quad \nabla f = \begin{pmatrix} 2x - 3y + 1 \\ -3x + 8y - 1 \end{pmatrix}, \quad \nabla f^0 = \begin{pmatrix} -1 \\ 9 \end{pmatrix} \text{ e } H^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

1º Iteração

$$x^1 = \begin{pmatrix} 2 \\ 2 \end{pmatrix} + \alpha^* \begin{pmatrix} 1 \\ -9 \end{pmatrix}.$$

Usando o método da seção áurea, dados  $x^0, \nabla f^0$  e  $\alpha \in [0, 4]$ , encontramos  $\alpha^* = 0,1165$ , donde  $x^1 = \begin{pmatrix} 2,1165 \\ 0,9515 \end{pmatrix}$ ,  $f^1 = 3,2244$  e  $\nabla f^1 = \begin{pmatrix} 2,3785 \\ 0,2625 \end{pmatrix}$ .

2º Iteração

$$s_0 = x^1 - x^0 = \begin{pmatrix} 0,1165 \\ -1,0485 \end{pmatrix} \text{ e } y_0 = \nabla f^1 - \nabla f^0 = \begin{pmatrix} 3,3785 \\ -8,7375 \end{pmatrix}.$$

Logo

$$\bullet \quad s_0 s_0^T = \begin{pmatrix} 0,1165 \\ -1,0485 \end{pmatrix} \begin{pmatrix} 0,1165 & -1,0485 \end{pmatrix} = \begin{pmatrix} 0,0136 & -0,1222 \\ -0,1222 & 1,0994 \end{pmatrix},$$

$$\bullet \quad s_0^T y_0 = \begin{pmatrix} 0,1165 & -1,0485 \end{pmatrix} \begin{pmatrix} 3,3785 \\ -8,7375 \end{pmatrix} = 9,5549,$$

$$\bullet \quad \frac{s_0 s_0^T}{s_0^T y_0} = \frac{1}{9,5549} \begin{pmatrix} 0,0136 & -0,1222 \\ -0,1222 & 1,0994 \end{pmatrix} = \begin{pmatrix} 0,00142 & -0,01279 \\ -0,01279 & 0,11506 \end{pmatrix},$$

$$\bullet \quad y_0 y_0^T = \begin{pmatrix} 3,3785 \\ -8,7375 \end{pmatrix} \begin{pmatrix} 3,3785 & -8,7375 \end{pmatrix} = \begin{pmatrix} 11,4143 & -29,5196 \\ -29,5196 & 76,3439 \end{pmatrix},$$

$$\bullet \quad y_0^T y_0 = \begin{pmatrix} 3,3785 & -8,7375 \end{pmatrix} \begin{pmatrix} 3,3785 \\ -8,7375 \end{pmatrix} = 87,7582,$$

$$\bullet \quad \frac{-y_0 y_0^T}{y_0^T y_0} = \frac{-1}{87,7582} \begin{pmatrix} 11,4143 & -29,5196 \\ -29,5196 & 76,3439 \end{pmatrix} = \begin{pmatrix} -0,1301 & 0,3363 \\ 0,3363 & -0,8699 \end{pmatrix}.$$

Iteração k	$(x^k)^T$	$f(x^k)$	$\alpha_k$	$p_k^T$	$\ \nabla f(x^k)\ $
1	(0; 2)	37	0,0355	(24; -40)	46,6476
2	(0,8521 ; 0,5799)	0,0324	0,9919	(0,5348; 0,0205)	0,1663
3	(1,5305 ; 1,0203)	$1,69 \times 10^{-7}$	0,5	$(-0,13; -0,09) \times 10^{-4}$	$6,93 \times 10^{-4}$

Tabela 7.1: Método Quase-Newton (DFP) aplicado ao Exemplo 1

Aproximação da inversa da pseudo-Hessiana:  $H^1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0,00142 & -0,01279 \\ -0,01279 & 0,11506 \end{pmatrix} + \begin{pmatrix} -0,1301 & 0,3363 \\ 0,3363 & -0,8699 \end{pmatrix} = \begin{pmatrix} 0,87132 & 0,32351 \\ 0,32351 & 0,24516 \end{pmatrix}$ ,

- $x^2 = x^1 + \alpha^*(-H^1\nabla f^1) = x^1 + \alpha^*p_1$
- $x^2 = \begin{pmatrix} 2,1165 \\ 0,9515 \end{pmatrix} - \alpha^* \begin{pmatrix} 0,87132 & 0,32351 \\ 0,32351 & 0,24516 \end{pmatrix} \begin{pmatrix} 2,3785 \\ 0,2625 \end{pmatrix} = \begin{pmatrix} 2,1165 \\ 0,9515 \end{pmatrix} - \alpha^* \begin{pmatrix} 2,15746 \\ 0,8340 \end{pmatrix}$ .

Pelo método da seção áurea, com  $\alpha \in [0, 4]$ , encontramos  $\alpha^* = 1,3121$ . Portanto,

- $x^2 = \begin{pmatrix} 2,1165 \\ 0,9515 \end{pmatrix} + 1,3121 \begin{pmatrix} -2,15746 \\ -0,8340 \end{pmatrix} = \begin{pmatrix} -0,7142 \\ -0,1433 \end{pmatrix}$ ,
- $f^2 = -0,2857$ ,
- $\nabla f^2 = \begin{pmatrix} 0,0015 \\ -0,0038 \end{pmatrix}$  e
- $\|\nabla f^2\|_L = 0,0038 < 0,5 \cdot 10^{-2}$ .

Assim, após duas iterações do método DFP, obteve-se a aproximação  $x^2 = (-0,7142; -0,1433)^T$  para o ponto ótimo.

**EXEMPLO 1:** Considere o problema definido em (3.31). A função  $f$  será minimizada pelo Método Quase-Newton partindo do ponto inicial  $x^0 = (0; 2)^T$ . Primeiro será aplicado o método DFP e em seguida o método BFGS. A Tabela 7.1 resume os cálculos do Método DFP na qual a terceira iteração do processo foi interrompido com o critério de parada  $\|\nabla f(x^k)\| < 0,001$ , visto que  $\nabla f(x^3) = (0,4; -0,56) \times 10^{-3}$ , donde  $\|\nabla f(x^3)\| = 6,93 \times 10^{-4}$ . O valor ótimo encontrado é  $x^3 = (1,5305; 1,0203)^T$  no qual a função objetivo vale  $1,6995 \times 10^{-7}$ . A Figura 7.1 mostra o progresso do Método DFP.

A Tabela 7.2 resume os cálculos do Método BFGS. Observe que algoritmo do método também foi interrompido na terceira iteração com o critério de parada  $\|\nabla f(x^k)\| < 0,001$ , visto que  $\nabla f(x^3) = (0; 0,1545) \times 10^{-3}$ , logo  $\|\nabla f(x^3)\| = 1,54 \times 10^{-4}$ . Como aproximação ótima obteve-se o ponto  $x^3 = (1,5507; 1,0338)^T$ , no qual a função objetivo vale  $1,3097 \times 10^{-6}$ . A Figura 7.2 mostra o progresso do Método BFGS.

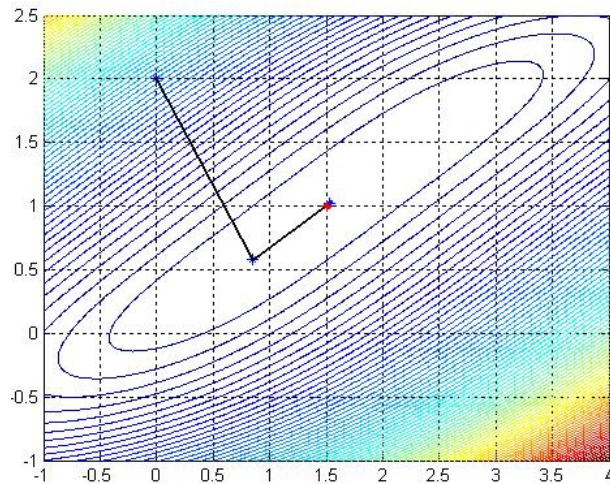


Figura 7.1: Progresso do Método Quase-Newton (DFP) aplicado ao Exemplo 1

Iteração k	$(x^k)^T$	$f(x^k)$	$\alpha_k$	$p_k^T$	$\ \nabla f(x^k)\ $
1	(0; 2)	37	0,0355	(24; -40)	46,6476
2	(0,8521 ; 0,5799)	0,0324	0,9865	(0,5583; 0,0343)	0,1663
3	(1,5507 ; 1,0338)	$1,30 \times 10^{-6}$	0,5	$(-0,67; -0,45) \times 10^{-4}$	$1,54 \times 10^{-4}$

Tabela 7.2: Método Quase-Newton (BFGS) aplicado ao Exemplo 1

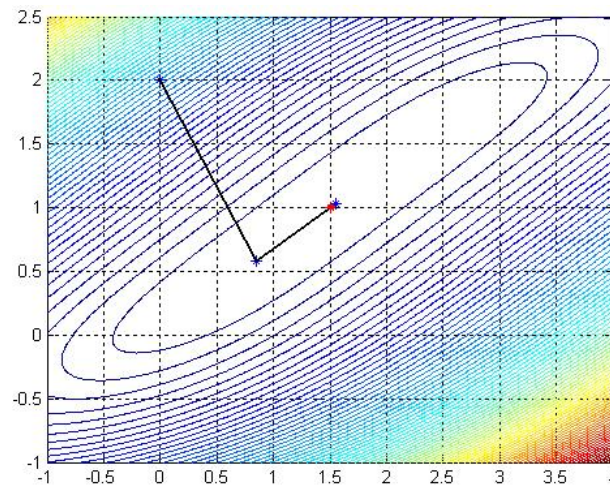


Figura 7.2: Progresso do Método Quase-Newton (BFGS) aplicado ao Exemplo 1

Iteração k	$(x^k)^T$	$f(x^k)$	$\alpha_k$	$p_k^T$	$\ \nabla f(x^k)\ $
1	(10; 20)	1,2927	$1,20 \times 10^3$	(-0,0145; 0,0018)	0,0146
2	(1 ; 22,2040)	1,4583	0,4282	(8,5595; -2,0949)	0,2875
3	(4,6653 ; 21,3070)	1,2368	$2,29 \times 10^{-3}$	(-0; -0,0017)	0,0017
4	(4,5586 ; 17,4256)	1,2326	0,6966	(-0,0228; -0,8043)	0,0034
5	(4,5427 ; 16,8653)	1,2325	$6,2861 \times 10^{-2}$	$(-0,66; 0,45) \times 10^{-3}$	0,0039
6	(4,1237 ; 17,1489)	1,2318	0,9785	(0,0502; -0,0396)	$5,74 \times 10^{-4}$

Tabela 7.3: Método Quase-Newton (DFP) aplicado ao Exemplo 2

Iteração k	$(x^k)^T$	$f(x^k)$	$\alpha_k$	$p_k^T$	$\ \nabla f(x^k)\ $
1	(10; 20)	1,2927	$1,2096 \times 10^3$	(-0,0145; 0,0018)	0,0146
2	(1 ; 22,2040)	1,4583	0,4282	(8,5595; -2,0949)	0,2875
3	(4,6653 ; 21,3070)	1,2368	$2,3232 \times 10^{-3}$	(-0; -0,0018)	0,0017
4	(4,5528 ; 17,2075)	1,2326	1,7529	(-0,0670; -0,5904)	0,0036
5	(4,4355 ; 16,1726)	1,2323	3,7647	(-0,1172; -0,622)	0,0039
6	(3,9941 ; 15,9383)	1,2315	0,5	(0,0068; 0,0675)	$3,93 \times 10^{-5}$

Tabela 7.4: Método Quase-Newton (BFGS) aplicado ao Exemplo 2

**EXEMPLO 2:** Agora, o Método Quase-Newton será aplicado visando otimizar a função (3.36) com  $x^0 = (10; 20)^T$ . As Tabelas 7.3 e 7.4 <sup>1</sup> mostram os cálculos para os métodos DFP e BFGS respectivamente. O critério de parada usado foi  $\|\nabla f^k\| < 0,001$ . O método *DFP* encontra  $p^* = (4,1237; 17,1489)^T$  e  $f(p^*) = 1,2318$  com 6 iterações, sendo que  $\nabla f(x^6) = (-0,1791; 0,5459) \times 10^{-3}$ , donde  $\|\nabla f(x^6)\| = 5,7453 \times 10^{-4}$ . O método *BFGS* chega no ponto  $p^* = (3,9941; 15,9383)$ , no qual a  $f$  vale 1,2315, também com 6 iterações com  $\nabla f(x^6) = (0,1980; -0,3406) \times 10^{-4}$  e  $\|\nabla f(x^6)\| = 3,9399 \times 10^{-5}$ . As Figuras 7.3 e 7.4 mostram o progresso dos Método BFGS e DFP aplicados ao Exemplo 2.

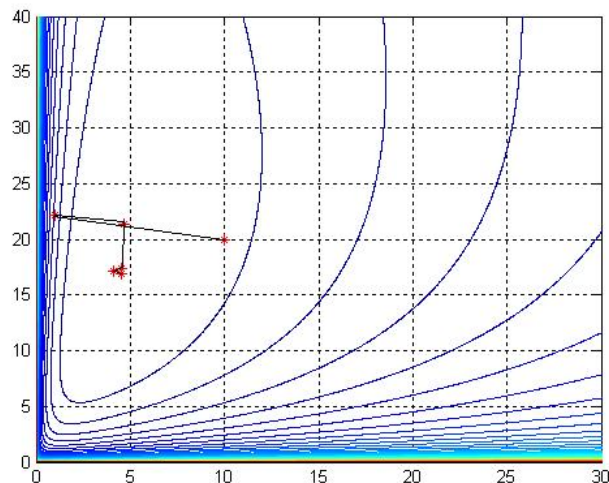


Figura 7.3: Progresso do Método Quase-Newton (DFP) aplicado ao Exemplo 2

<sup>1</sup>Adaptação para se obedecer as condições da região viável  $p_1 \geq 1$ . O valor obtido  $-7,5392$  foi substituído por 1 na segunda iteração das Tabelas 7.3 e 7.4.

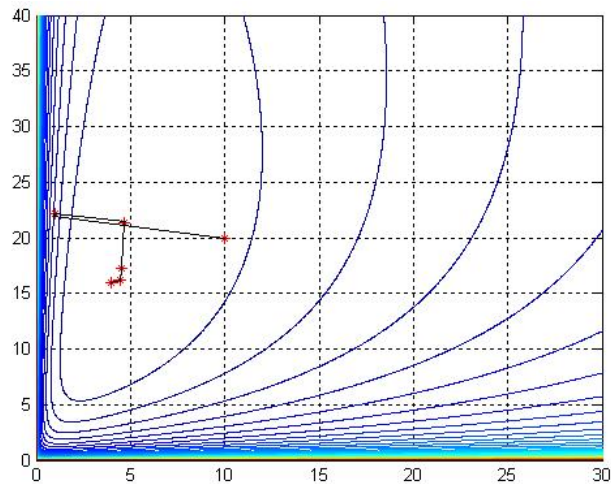


Figura 7.4: Progresso do Método Quase-Newton (BFGS) aplicado ao Exemplo 2

Pode se notar que os Métodos de Newton e Quase-Newton são parecidos, porém, uma das grandes vantagens do Método Quase-Newton sobre o Método de Newton é o fato daquele não utilizar a matriz Hessiana em seu algoritmo e sim uma aproximação desta. Esta modificação faz com que os Métodos Quase-Newton tornem-se aplicáveis a uma variedade maior de problemas, principalmente, os que possuem a matriz Hessiana  $\nabla^2 f(x)$  mal condicionada. Foi este fato que tornou possível o Método Quase-Newton encontrar as pressões intermediárias do compressor de multi-estágios do Exemplo 2. Além disso, o método convergiu, rapidamente, para o ótimo, o que não aconteceu no Método de Newton neste caso.

A semelhança dos Métodos Newton e Quase-Newton pode ser comprovada no Exemplo 1. Como o determinante da matriz Hessiana para este exemplo vale  $\det(\nabla^2 f(x)) = 96x_2^2 - 192x_2 + 96 \neq 0$  para todo valor de  $x = (x_1; x_2)^T \in \mathbb{R}^2$ , os resultados obtidos são bastante próximos. Entretanto, considerando o número de iterações, o Método Quase-Newton convergiu um pouco mais rápido.

# Capítulo 8

## Método dos Gradientes Conjugados

O método dos gradientes conjugados, proposto por Hestenes e Stiefel (1952), consiste em um algoritmo para a solução numérica de sistemas particulares de equações lineares, aqueles cuja matriz é simétrica positiva definida. O objetivo inicial deste método era resolver problemas quadráticos sem restrições, mas logo o mesmo foi estendido para casos mais gerais. Atualmente, extensões do método dos gradientes conjugados resolvem sistemas de equações não-lineares. O uso deste método para otimização sem restrição foi induzido pelo fato que minimizar uma função quadrática definida positiva é equivalente a resolver um sistema de equações lineares resultantes ao se fazer o gradiente igual a zero.

O método dos gradientes conjugados é um caso particular do método das direções conjugadas, os quais, convergem em no máximo  $n$  iterações para problemas de otimização quadráticos sem restrições em  $\mathbb{R}^n$  quando usado busca linear exata.

À seguir serão analisados dois teoremas que garantem que se a função  $f$  é quadrática as direções  $d_1, \dots, d_n$  geradas são de fato conjugadas, e portanto, o algoritmo dos gradientes conjugados produz uma solução ótima depois de uma aplicação completa do passo principal, ou seja, depois, de no máximo  $n$  buscas lineares terem sido realizadas chega-se ao ponto ótimo.

**Teorema 8.1** *Seja  $f(x) = c^T x + \frac{1}{2} x^T H x$ , sendo que  $H$  é uma matriz simétrica  $n \times n$ . Sejam  $d_1, \dots, d_n$   $H$ -conjugadas e  $x^1$  um ponto inicial arbitrário. Para  $k = 1, \dots, n$ , seja  $\alpha_k$  uma solução ótima do problema de minimizar  $f(x^k + \alpha d_k)$  sujeito à  $\alpha \in \mathbb{R}$ , e seja  $x^{k+1} = x^k + \alpha_k d_k$ . Então, para  $k = 1, \dots, n$  tem-se:*

1.  $\nabla f(x^{k+1})^T d_j = 0$  para  $j = 1, \dots, k$ .
2.  $\nabla f(x^1)^T d_k = \nabla f(x^k)^T d_k$ .
3.  $x^{k+1}$  é uma solução ótima do problema de minimizar  $f(x)$  sujeito à  $x - x^1 \in L(d_1, \dots, d_k)$ , sendo que  $L(d_1, \dots, d_k)$  é o subespaço linear gerado por  $d_1, \dots, d_k$ , isto é,  $L(d_1, \dots, d_k) = \{ \sum_{j=1}^k \mu_j d_j : \mu_j \in \mathbb{R} \text{ para cada } j \}$ . Em particular,  $x^{n+1}$  é um ponto que minimiza  $f$  sobre  $\mathbb{R}^n$ .

**Demonstração** Para mostrar a afirmativa 1, primeiro observe que  $f(x^j + \alpha d_j)$  alcança o mínimo em  $\alpha_j$  só se  $\nabla f(x^j + \alpha_j d_j)^T d_j = 0$ , donde  $\nabla f(x^{j+1})^T d_j = 0$ . Assim, a afirmativa 1 é verdadeira para  $j = k$ . Para  $j < k$ , note que

$$\nabla f(x^{k+1}) = c + H x^{k+1} = c + H x^{j+1} + H \left( \sum_{i=j+1}^k \alpha_i d_i \right) = \nabla f(x^{j+1}) + H \left( \sum_{i=j+1}^k \alpha_i d_i \right). \quad (8.1)$$

Pela conjugação,  $d_i^T H d_j = 0$  para  $i = j+1, \dots, k$ . Assim, de (8.1) segue que  $\nabla f(x^{k+1})^T d_j = 0$ , provando a primeira parte.

Trocando  $k$  por  $k - 1$  e sendo  $j = 0$  em (8.1), tem-se

$$\nabla f(x^k) = \nabla f(x^1) + H \left( \sum_{i=1}^{k-1} \alpha_i d_i \right) \quad \text{para } k \geq 2. \quad (8.2)$$

Multiplicando por  $d_k^T$  e observando que  $d_k^T H d_i = 0$  para  $i = 1, \dots, k-1$ , tem-se que  $\nabla f(x^1)^T d_k = \nabla f(x^k)^T d_k$  para  $k \geq 2$ . Para  $k = 1$  é óbvio que a afirmação 2 é verdadeira.

Para mostrarmos a afirmativa 3, já que  $d_i^T H d_j = 0$  para  $i \neq j$ , tem-se

$$\begin{aligned} f(x^{k+1}) &= f[x^1 + (x^{k+1} - x^1)] = f \left( x^1 + \sum_{j=1}^k \alpha_j d_j \right) \\ &= f(x^1) + \nabla f(x^1)^T \left( \sum_{j=1}^k \alpha_j d_j \right) + \frac{1}{2} \left( \sum_{j=1}^k \alpha_j^2 d_j^T H d_j \right). \end{aligned} \quad (8.3)$$

Agora, suponha que  $x - x^1 \in L(d_1, \dots, d_k)$ , então  $x$  pode ser escrito como  $x^1 + \sum_{j=1}^k \mu_j d_j$ . Assim como em (8.3), tem-se

$$f(x) = f(x^1) + \nabla f(x^1)^T \left( \sum_{j=1}^k \mu_j d_j \right) + \frac{1}{2} \sum_{j=1}^k \mu_j^2 d_j^T H d_j. \quad (8.4)$$

Para completar a demonstração, é necessário mostrar que  $f(x) \geq f(x^{k+1})$ . Por contradição, suponha que  $f(x) < f(x^{k+1})$ . Então, por (8.3) e (8.4) tem-se

$$\begin{aligned} \nabla f(x^1)^T \left( \sum_{j=1}^k \mu_j d_j \right) + \frac{1}{2} \sum_{j=1}^k \mu_j^2 d_j^T H d_j \\ < \nabla f(x^1)^T \left( \sum_{j=1}^k \alpha_j d_j \right) + \frac{1}{2} \sum_{j=1}^k \alpha_j^2 d_j^T H d_j. \end{aligned} \quad (8.5)$$

Pela definição de  $\alpha_j$ , observe que  $f(x^j + \alpha_j d_j) \leq f(x^j + \mu_j d_j)$  para cada  $j$ . Portanto,

$$f(x^j) + \alpha_j \nabla f(x^j)^T d_j + \frac{1}{2} \alpha_j^2 d_j^T H d_j \leq f(x^j) + \mu_j \nabla f(x^j)^T d_j + \frac{1}{2} \mu_j^2 d_j^T H d_j. \quad (8.6)$$

Da afirmativa 2,  $\nabla f(x^k)^T d_k = \nabla f(x^1)^T d_k$ . Substituindo isto na inequação acima tem-se

$$\alpha_j \nabla f(x^1)^T d_j + \frac{1}{2} \alpha_j^2 d_j^T H d_j \leq \mu_j \nabla f(x^1)^T d_j + \frac{1}{2} \mu_j^2 d_j^T H d_j. \quad (8.7)$$

Absurdo, pois (8.7) contradiz (8.5) para  $j = 1, \dots, k$ . Assim,  $x^{k+1}$  é um ponto que minimiza  $x^1 + L(d_1, \dots, d_k)$ . Em particular, já que  $d_1, \dots, d_n$  são linearmente independentes e  $L(d_1, \dots, d_n) = \mathbb{R}^n$ , então  $x^{n+1}$  é um ponto de mínimo de  $f$  sobre  $\mathbb{R}^n$ , o que completa a demonstração.  $\square$

**Teorema 8.2** *Considere o problema de minimizar  $f(x) = c^T x + \frac{1}{2} x^T H x$  sujeito à  $x \in \mathbb{R}^n$ . Suponha que o problema é resolvido iniciando com  $y^1$  e  $d_1 = -\nabla f(y^1)$ . Em particular, para  $j = 1, \dots, n$ , seja  $\alpha_j$  uma solução ótima do problema de minimizar  $f(y^j + \alpha d_j)$  sujeito à  $\alpha \geq 0$ . Sejam  $y^{j+1} = y^j + \alpha_j d_j$  e  $d_{j+1} = -\nabla f(y^{j+1}) + \beta_j d_j$ , sendo que  $\beta_j = \|\nabla f(y^{j+1})\|^2 / \|\nabla f(y^j)\|^2$ . Se  $\nabla f(y^j) \neq 0$  para  $j = 1, \dots, n$ , então as seguintes afirmações são verdadeiras:*

1.  $d_1, \dots, d_n$  são  $H$ -conjugadas.
2.  $d_1, \dots, d_n$  são direções de descida.
3.  $\beta_j = \frac{\|\nabla f(y^{j+1})\|^2}{\|\nabla f(y^j)\|^2} = \frac{d_j^T H \nabla f(y^{j+1})}{d_j^T H d_j}$  para  $j = 1, \dots, n$ .

**Demonstração:** Primeiro suponha que as afirmações 1, 2 e 3 sejam verdadeiras para  $j$ . Será mostrado que elas também valem para  $j + 1$ . Para mostrar que a afirmação 1 é verdadeira para  $j + 1$ , será mostrado que  $d_k^T H d_{j+1} = 0$  para  $k \leq j$ . Já que  $d_{j+1} = -\nabla f(y^{j+1}) + \beta_j d_j$ , observando a hipótese de indução na afirmação 3 e fazendo  $k = j$ , tem-se

$$d_j^T H d_{j+1} = d_j^T H \left[ -\nabla f(y^{j+1}) + \frac{d_j^T H \nabla f(y^{j+1})}{d_j^T H d_j} d_j \right] = 0. \quad (8.8)$$

Agora, seja  $k < j$ . Já que  $d_{j+1} = -\nabla f(y^{j+1}) + \beta_j d_j$  e  $d_k^T H d_j = 0$ , pela hipótese de indução da afirmação 3,

$$d_j^T H d_{j+1} = -d_k^T H \nabla f(y^{j+1}). \quad (8.9)$$

Ainda, como  $\nabla f(y^{k+1}) = c + H y^{k+1}$ ,  $y^{k+1} = y^k + \alpha_k d_k$  e  $d_k = -\nabla f(y^k) + \beta_{k-1} d_{k-1}$ , observe que

$$\begin{aligned} d_{k+1} &= -\nabla f(y^{k+1}) + \beta_k d_k \\ &= -[c + H y^k + \alpha_k H d_k] + \beta_k d_k \\ &= -[\nabla f(y^k) + \alpha_k H d_k] + \beta_k d_k \\ &= -[-d_k + \beta_{k-1} d_{k-1} + \alpha_k H d_k] + \beta_k d_k. \end{aligned} \quad (8.10)$$

Pela hipótese de indução da afirmação 2,  $d_k$  é uma direção de descida e portanto,  $\alpha_k > 0$ . Logo,

$$d_k^T H = \frac{1}{\alpha_k} [-d_{k+1}^T + (1 - \beta_k) d_k^T - \beta_{k-1} d_{k-1}^T]. \quad (8.11)$$

De (8.9) e (8.11) segue que

$$\begin{aligned} d_k^T H d_{j+1} &= -d_k^T H \nabla f(y^{j+1}) \\ &= \frac{1}{\alpha_k} [-d_{k+1}^T \nabla f(y^{j+1}) + (1 - \beta_k) d_k^T \nabla f(y^{j+1}) - \beta_{k-1} d_{k-1}^T \nabla f(y^{j+1})]. \end{aligned} \quad (8.12)$$

Da afirmação 1 do Teorema 8.1, já que foi assumido que  $d_1, \dots, d_j$  são conjugados, segue que  $d_{k+1}^T \nabla f(y^{j+1}) = d_k^T \nabla f(y^{j+1}) = d_{k-1}^T \nabla f(y^{j+1}) = 0$ . Assim, a equação acima implica que  $d_k^T H d_{j+1} = 0$  para  $k < j$ . Isto juntamente com (8.8), mostra que  $d_k^T H d_{j+1} = 0$  para todo  $k \leq j$ . Para mostrar que  $d_1, \dots, d_{j+1}$  são H-conjugados, é suficiente mostrar que são linearmente independentes. Suponha que

$$\sum_{i=1}^{j+1} \gamma_i d_i = 0 \quad \Rightarrow \quad \sum_{i=1}^j \gamma_i d_i + \gamma_{j+1} [-\nabla f(y^{j+1}) + \beta_j d_j] = 0. \quad (8.13)$$

Multiplicando por  $\nabla f(y^{j+1})^T$  e observando a afirmação 1 do Teorema 8.1, segue que

$$\gamma_{j+1} \|\nabla f(y^{j+1})\|^2 = 0.$$

Visto que  $\nabla f(y^{j+1}) \neq 0$  e  $\gamma_{j+1} = 0$ . Isto implica que  $\sum_{i=1}^j \gamma_i d_i = 0$ , e observando que  $d_1, \dots, d_j$  são conjugados segue que  $\gamma_1, \dots, \gamma_j = 0$ . Assim,  $d_1, \dots, d_{j+1}$  são linearmente independentes e H-conjugados, donde a afirmativa 1 é verdadeira para  $j + 1$ .

Será mostrado agora que a afirmação 2 é verdadeira para  $j + 1$ , isto é,  $d_{j+1}$  é uma direção de descida. Note que  $\nabla f(y^{j+1}) \neq 0$  por hipótese, e que  $\nabla f(y^{j+1})^T d_j = 0$  da afirmação 1 do Teorema 8.1. Então,  $\nabla f(y^{j+1})^T d_{j+1} = -\|\nabla f(y^{j+1})\|^2 + \beta_j \nabla f(y^{j+1})^T d_j = -\|\nabla f(y^{j+1})\|^2 < 0$ . Do Teorema 1.2,  $d_{j+1}$  é uma direção de descida.

Agora será mostrado que a afirmação 3 é verdadeira para  $j + 1$ . Colocando  $k = j + 1$  em (8.11) e multiplicando por  $\nabla f(y^{j+2})$ , segue que

$$\begin{aligned} \alpha_{j+1} d_{j+1}^T H \nabla f(y^{j+2}) &= [-d_{j+2}^T + (1 + \beta_{j+1}) d_{j+1}^T - \beta_j d_j^T] \nabla f(y^{j+2}) \\ &= [\nabla f(y^{j+2})^T + d_{j+1}^T - \beta_j d_j^T] \nabla f(y^{j+2}). \end{aligned} \quad (8.14)$$

Já que  $d_1, \dots, d_{j+1}$  são  $H$ -conjugados, então, da afirmação 1 do Teorema 8.1,  $d_{j+1}^T \nabla f(y^{j+2}) = d_j^T \nabla f(y^{j+2}) = 0$ . Da equação acima segue que

$$\|\nabla f(y^{j+2})\|^2 = \alpha_{j+1} d_{j+1}^T H \nabla f(y^{j+2}). \quad (8.15)$$

Multiplicando  $\nabla f(y^{j+1}) = \nabla f(y^{j+2}) - \alpha_{j+1} H d_{j+1}$  por  $\nabla f(y^{j+1})^T$  e observando que  $d_j^T H d_{j+1} = d_{j+1}^T \nabla f(y^{j+2}) = d_j^T \nabla f(y^{j+2}) = 0$ , tem-se

$$\begin{aligned} \|\nabla f(y^{j+1})\|^2 &= \nabla f(y^{j+1})^T [\nabla f(y^{j+2}) - \alpha_{j+1} H d_{j+1}] \\ &= (-d_{j+1}^T + \beta_j d_j^T) [\nabla f(y^{j+2}) - \alpha_{j+1} H d_{j+1}] \\ &= \alpha_{j+1} d_{j+1}^T H d_{j+1}. \end{aligned} \quad (8.16)$$

De (8.15) e (8.16), segue imediatamente que a afirmação 3 é verdadeira para  $j + 1$ .

Foi demonstrado que se as afirmativas 1, 2 e 3 são verdadeiras para  $j$ , então também são para  $j + 1$ . Note que as afirmativas 1 e 2 são trivialmente verdadeira para  $j = 1$ . Ainda, usando um argumento similar ao usado para mostrar que a afirmativa 3 é verdadeira para  $j + 1$ , mostra-se facilmente que a afirmativa 3 é verdadeira para  $j = 1$ . Completando assim a demonstração.  $\square$

Um esquema básico do método dos gradientes conjugados para minimizar um função diferenciável  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é gerar uma sequência de iterações  $x^j$  de acordo com

$$x^{j+1} = x^j + \alpha_j d_j, \quad (8.17)$$

sendo  $d_j$  a direção de busca e  $\alpha_j$  o tamanho do passo que minimiza  $f$  ao longo de  $d_j$  à partir do ponto  $x^j$ . Para  $j = 1$ , pode-se usar a direção de busca  $d_1 = -\nabla f(x^1)$ , e para as iterações seguintes, dado  $x^{j+1}$  com  $\nabla f(x^{j+1}) \neq 0$  para  $j \geq 1$ , usa-se

$$d_{j+1} = -\nabla f(x^{j+1}) + \beta_j d_j, \quad (8.18)$$

sendo  $\beta_j$  um parâmetro de desvio adequado que caracteriza um método particular dos gradientes conjugados. Observe que sempre que  $\beta_j \geq 0$  podemos escrever  $d_{j+1}$  em (8.18) como

$$d_{j+1} = \frac{1}{\mu} [\mu [-\nabla f(x^{j+1})] + (1 - \mu) d_j], \quad (8.19)$$

sendo que  $\mu = 1/(1 + \beta_j)$ . Assim,  $d_{j+1}$  pode ser visto como uma combinação convexa da direção de descida atual e a descida usada na última iteração.

Vários estudiosos contribuíram para os avanços deste método, e atualmente existem muitos métodos particulares dos gradientes conjugados que se diferenciam pela escolha de  $\beta_j$ . Serão considerados aqui três escolhas para  $\beta_j$ .

Sejam

$$\begin{aligned} p_j &= x^{j+1} - x^j, \\ q_j &= \nabla f(x^{j+1}) - \nabla f(x^j) = H(x^{j+1} - x^j) = H p_j. \end{aligned} \quad (8.20)$$

Suponha que  $f$  é uma função quadrática com Hessiana  $H$  positiva definida e que  $d_{j+1}$  e  $d_j$  sejam  $H$ -conjugados. De (8.17) e (8.20) segue que  $d_{j+1}^T H d_j = 0$ , donde  $0 = d_{j+1}^T H p_j = d_{j+1}^T q_j$ . Usando este fato em (8.18) encontra-se a escolha de  $\beta_j$  de Hestenes e Stiefel (1952), usado também em situações não-lineares por assumir um comportamento quadrático local, como

$$\beta_j^{HS} = \frac{\nabla f(x^{j+1})^T q_j}{d_j^T q_j} \equiv \frac{\alpha_j \nabla f(x^{j+1})^T q_j}{p_j^T q_j}. \quad (8.21)$$

Quando são realizadas buscas lineares exatas, tem-se  $d_j^T \nabla f(x^{j+1}) = 0 = d_{j-1}^T \nabla f(x^j)$ , levando a  $d_j^T q_j = -d_j^T \nabla f(x^j) = [\nabla f(x^j) - \beta_{j-1} d_{j-1}]^T \nabla f(x^j) = \|\nabla f(x^j)\|^2$ . Substituindo isto em (8.21) obtém-se a escolha de  $\beta_j$  de Polak e Ribiere (1969) como

$$\beta_j^{PR} = \frac{\nabla f(x^{j+1})^T q_j}{\|\nabla f(x^j)\|^2}. \quad (8.22)$$

Além disso, se  $f$  é quadrática e se são realizadas buscas lineares exatas, tem-se, usando (8.17) e (8.18) junto com  $\nabla f(x^{j+1})^T d_j = 0 = \nabla f(x^j)^T d_{j-1}$  como acima, que

$$\begin{aligned} \nabla f(x^{j+1})^T \nabla f(x^j) &= \nabla f(x^{j+1})^T [\beta_{j-1} d_{j-1} - d_j] \\ &= \beta_{j-1} \nabla f(x^{j+1})^T d_{j-1} = \beta_{j-1} [\nabla f(x^j) + \alpha_j H d_j]^T d_{j-1} \\ &= \beta_{j-1} \alpha_j d_j^T H d_{j-1} = 0, \end{aligned} \quad (8.23)$$

onde a última igualdade vem do fato de que  $d_j$  e  $d_{j-1}$  são  $H$ -conjugados (sendo que  $d_0 \equiv 0$ ). Portanto,

$$\nabla f(x^{j+1})^T \nabla f(x^j) = 0. \quad (8.24)$$

Substituindo (8.24) em (8.22) e usando (8.20) tem-se a escolha de  $\beta_j$  de Fletcher e Reeves (1964) como

$$\beta_j^{FR} = \frac{\|\nabla f(x^{j+1})\|^2}{\|\nabla f(x^j)\|^2}. \quad (8.25)$$

Agora, será apresentada uma análise formal do método dos gradientes conjugados usando a escolha para  $\beta_j$  de Fletcher e Reeves dada pela equação (8.25). Uma argumentação similar pode ser feita para as outras escolhas de  $\beta_j$ .

Se a função  $f$  é quadrática e são realizadas buscas lineares exatas então a escolha de  $\beta_j$ , dadas diferentemente por (8.21), (8.22) e (8.25), coincidem, e portanto o Teorema 8.2 permanece verdadeiro para a escolha de  $\beta_j$  de Hestenes e Stiefel (HS) e de Polak e Ribiere (PR). Entretanto, para funções não-quadráticas, a escolha  $\beta_j^{PR}$  parece ser empiricamente melhor do que  $\beta_j^{FR}$ . Isto é compreensível, já que a redução de (8.22) para (8.25) assume que  $f$  é quadrática.

Além disso, note que tem-se usado  $d_1 = -I \nabla f(x^1)$  nas análises precedentes. Mas no lugar de usar a matriz identidade aqui, poderia-se ter usado alguma matriz genérica pré-condicionada  $D$ , sendo que  $D$  é simétrica e positiva definida. Assim, teria-se dado  $d_1 = -D \nabla f(x^1)$  e (8.18) teria se tornado  $d_{j+1} = -D \nabla f(x^{j+1}) + \beta_j d_j$ , donde, por exemplo, no caso de (8.21), teria-se

$$\beta_j^{HS} = \frac{q_j^T D \nabla f(x^{j+1})}{q_j^T d_j}. \quad (8.26)$$

Isto corresponde essencialmente em fazer a mudança de variável  $y' = D^{-1/2} y$  e usar o algoritmo original do método dos gradientes conjugados.

## 8.1 Síntese do método dos gradientes conjugados de Fletcher e Reeves

Um resumo do método dos gradientes conjugados para minimizar uma função diferenciável genérica é dado à seguir.

**Dados iniciais:** Escolha um escalar  $\epsilon > 0$  como critério de parada e um ponto inicial  $x^1$ . Sejam  $y^1 = x^1$ ,  $d_1 = -\nabla f(y^1)$ ,  $k = j = 1$ , e siga para o passo principal.

**Passo 1:** Se  $\|\nabla f(y^j)\| < \epsilon$ , pare. Caso contrário, seja  $\alpha_j$  a solução ótima do problema de minimizar  $f(y^j + \alpha d_j)$  sujeito à  $\alpha \geq 0$ , e seja  $y^{j+1} = y^j + \alpha_j d_j$ . Se  $j < n$ , vá para o passo 2, se não vá para o passo 3.

**Passo 2:** Seja  $d_{j+1} = -\nabla f(y^{j+1}) + \beta_j d_j$ , sendo que

$$\beta_j = \frac{\|\nabla f(y^{j+1})\|^2}{\|\nabla f(y^j)\|^2}.$$

Troque  $j$  por  $j + 1$ , e vá para o passo 1.

**Passo 3:** Seja  $y^1 = x^{k+1} = y^{n+1}$ , e seja  $d_1 = -\nabla f(y^1)$ . Tome  $j = 1$ , troque  $k$  por  $k + 1$  e vá para o passo 1.

Como mostrado no Teorema 8.1, se a função é quadrática, então qualquer algoritmo de direção conjugada produz uma solução ótima num número finito de passos. Entretanto, se a função não é necessariamente quadrática, Bazaraa *et al.* (1979) mostra que se algumas propriedades são satisfeitas então o algoritmo das direções conjugadas é convergente.

**EXEMPLO 1:** Considere o problema em (3.31). Um resumo computacional da utilização do método de Fletcher e Reeves é dado na Tabela 8.1. Em cada iteração,  $d_1$  é dado por  $-\nabla f(y^1)$  e  $d_2$  é dado por  $-\nabla f(y^2) + \alpha_1 d_1$ , sendo que  $\alpha_1 = \|\nabla f(y^2)\|^2 / \|\nabla f(y^1)\|^2$ . Além disso,  $y^{j+1}$  é obtido pela otimização ao longo de  $d_j$ , partindo de  $y^j$ . Na terceira iteração, o ponto  $y^2 = (1,4198; 0,9466)^T$ , que é muito próximo do ponto ótimo  $(\frac{3}{2}; 1)^T$ , é alcançado. Já que a norma do gradiente vale  $3,3826 \times 10^{-4} < 0,001$ , o critério de parada foi satisfeito. Além disso,  $f(y^2) = 8,1485 \times 10^{-6}$ . O progresso do algoritmo é mostrado na Figura 8.1.

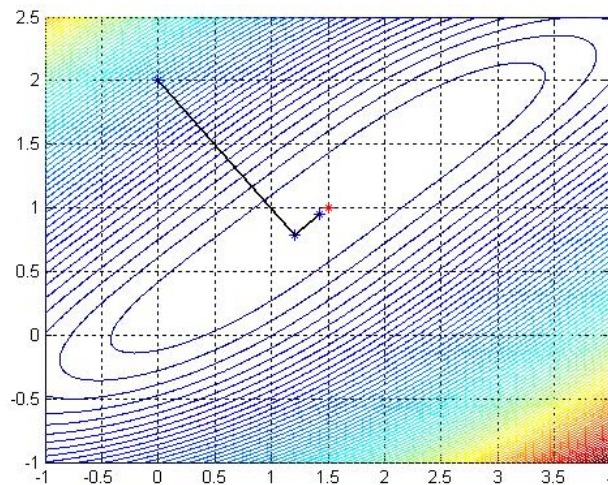


Figura 8.1: Progresso do Método dos Gradientes Conjugados aplicado ao Exemplo 1

**EXEMPLO 2:** Considere, finalmente, o problema definido em (3.36). A Tabela 8.2 <sup>1</sup> resume os cálculos para que o ponto inicial  $(10; 20)^T$  se aproxime de um valor ótimo, o que acontece já na segunda iteração, visto que  $\|\nabla f(3,9467; 16,0573)\| = 6,7796 \times 10^{-4} < 0,001$ , satisfazendo o critério de parada. Tem-se como aproximação do valor ótimo  $y^2 = (3,9467; 16,0573)^T$  e  $f(y^2) = 1,2315$ . A Figura 8.2 ilustra o progresso do método.

O Método dos Gradientes Conjugados mostrou-se, tanto para o Exemplo 1 quanto para o Exemplo 2, o método mais eficaz entre as técnicas apresentadas aqui neste estudo.

<sup>1</sup>Foi feita uma adaptação para se obedecer a região viável, onde  $p_1 \geq 1$ . O valor obtido  $-7,5392$  foi substituído por 1 na primeira iteração da Tabela 8.2.

Iteração k	$(x^k)^T$ $f(x^k)$	j	$(y^j)^T$ $f(y^j)$	$(\nabla f(y^j))^T$	$\ \nabla f(y^j)\ $	$\beta_{j-1}$	$d_j^T$	$\alpha_j$	$y^{j+1}$
1	(0; 2) 37	1	(0; 2) 37	(-24; 40)	46,6476	-	(24; -40)	0,0355	(0,8521; 0,5799)
		2	(0,8521; 0,5799) 0,0324	(-0,1418; -0,0839)	0,1648	$1,2478 \times 10^{-5}$	(0,1421; 0,0834)	2,4725	(1,2034; 0,7862)
2	(1,2034; 0,7862) 0,0044	1	(1,2034; 0,7862) 0,0044	(0,1928; -0,3283)	0,3807	-	(-0,1928; 0,3283)	0,0380	(1,1961; 0,7987)
		2	(1,1961; 0,7987) 0,0016	(-0,0156; -0,0092)	0,0181	0,0023	(0,0152; 0,0099)	14,7747	(1,4206; 0,9454)
3	(1,4206; 0,9454) $3,5666 \times 10^{-5}$	1	(1,4206; 0,9454) $3,5666 \times 10^{-5}$	(0,0207; -0,0317)	0,0379	-	(-0,0207; 0,0317)	0,0384	(1,4198; 0,9466)
		2	(1,4198; 0,9466) $8,1485 \times 10^{-6}$	(-0,2832; -0,1849) $\times 10^{-3}$	$3,3826 \times 10^{-4}$	$7,9894 \times 10^{-5}$	(0,2816; 0,1875) $\times 10^{-3}$	0	(1,4198; 0,9466)

Tabela 8.1: Método dos Gradientes Conjugados aplicado ao Exemplo 1

Iteração k	$(x^k)^T$ $f(x^k)$	j	$(y^j)^T$ $f(y^j)$	$(\nabla f(y^j))^T$	$\ \nabla f(y^j)\ $	$\beta_{j-1}$	$d_j^T$	$\alpha_j$	$y^{j+1}$
1	(10; 20) 1,2927	1	(10; 20) 1,2927	(0,0145; -0,0018)	0,0146	-	(-0,0145; 0,0018)	$1,2096 \times 10^3$	(1; 22,2040)
		2	(1; 22,2040) 1,4583	(-0,2872; 0,0096)	0,2875	$3,8896 \times 10^2$	(-5,3375; 0,6991)	-0,6911	(4,6888; 21,7208)
2	(4,6888; 21,7208) 1,2375	1	(4,6888; 21,7208) 1,2375	(0,0002; 0,0018)	0,0018	-	(-0,0002; -0,0018)	$3,1921 \times 10^3$	(3,9467; 16,0573)
		2	(3,9467; 16,0573) 1,2315	(-0,6688; 0,1113) $\times 10^{-3}$	$6,7796 \times 10^{-4}$	0,1435	(0,6354; -0,3660) $\times 10^{-3}$	0	(3,9467; 16,0573)

Tabela 8.2: Método dos Gradientes Conjugados aplicado ao Exemplo 2

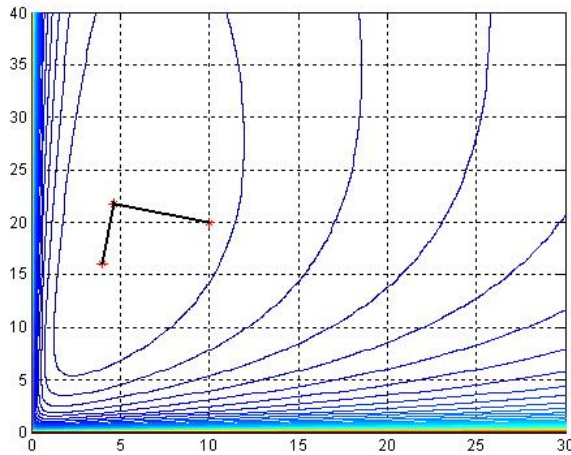


Figura 8.2: Progresso do Método dos Gradientes Conjugados aplicado ao Exemplo 2

## 8.2 Recomendações para a reinicialização do método dos gradientes conjugados

Em muitos experimentos computacionais usando diferentes técnicas de gradientes conjugados, com ou sem busca linear exata, tem-se mostrado que a performance dos métodos dos gradientes conjugados pode ser bastante reforçada por se empregar critérios de reinicialização. Em particular, o processo de reinício sugerido por Beale (1972) e argumentado por Powell (1977) tem mostrado ser muito eficiente e invariavelmente implementado, como descrito abaixo.

Considere o método dos gradientes conjugados adotando a escolha de  $\beta_j$  de Fletcher e Reeves. (Naturalmente, esta estratégia se aplica para qualquer outra escolha aceitável para  $\beta_j$ ). Em algum laço do algoritmo apresentado anteriormente, dentro da iteração  $j$  deste processo, tendo encontrado  $y^{j+1} = y^j + \alpha_j d_j$ , suponha que decida-se fazer um reinício (parar e recomeçar); esta decisão pode ser feita sempre que  $j = n$ . Seja  $R = j$  a iteração de reinicialização. Para a próxima iteração, encontra-se a direção de busca

$$d_{R+1} = -\nabla f(y^{R+1}) + \beta_R d_R, \quad (8.27)$$

como o usual. Então, no passo 3, troca-se  $y^1$  por  $y^{R+1}$ ,  $x^{k+1} \equiv y^{R+1}$ ,  $d_1 = d_{R+1}$  e retorne ao passo 1 para continuar com o próximo conjunto de laços dentro das iterações. Contudo, ao invés de computar  $d_{j+1} = -\nabla f(y^{j+1}) + \beta_j d_j$  para  $j \geq 1$ , agora usa-se

$$d_2 = -\nabla f(y^2) + \beta_2 d_2 \quad (8.28)$$

e

$$d_{j+1} = -\nabla f(y^{j+1}) + \beta_j d_j + \gamma_j d_1 \quad (8.29)$$

para  $j \geq 2$ , sendo que

$$\gamma_j = \frac{\nabla f(y^{j+1})^T q_1}{d_1^T q_1}, \quad (8.30)$$

e  $\alpha_j$  é computado como antes, dependendo do método a ser usado. Observe que (8.28) implica que o esquema usual dos gradientes conjugados produz  $d_1$  e  $d_2$  como H-conjugados quando  $f$  é quadrática. Entretanto, quando  $f$  é quadrática com Hessiana  $H$  positiva definida e  $d_1$  é escolhido arbitrariamente, então quando  $j = 2$ , por exemplo, a escolha usual de  $\beta_2$  faz com que  $d_3$  e  $d_2$  sejam H-conjugados. Mas é necessário algo adicional para que  $d_3$  e  $d_1$  sejam H-conjugados. Isto é realizado pelo termo extra  $\gamma_2 d_1$ . De fato, exigindo  $d_3^T H d_1 = 0$ , sendo que  $d_3$  é dado pela expressão (8.29), e observando que  $d_2^T H d_1 = 0$ , tem-se que

$$\gamma_2 = \frac{\nabla f(y^3)^T H d_1}{d_1^T H d_1} = \frac{\nabla f(y^3)^T q_1}{d_1^T q_1}.$$

Seguindo indutivamente com o mesmo raciocínio, segue que o termo adicional em (8.29) garante a H-conjugação de todas as direções geradas.

O esquema acima mencionado foi sugerido por Beale com a motivação de que sempre que o a reinicialização é feita usando  $d_1 = -\nabla f(y^1)$  ao invés de  $d_1 = d_{R+1}$  dado por (8.27), perde-se uma importante informação de segunda ordem inerente a  $d_R$ .

# Capítulo 9

## Discussões e Conclusões

Para facilitar a comparação entre as nove técnicas estudadas, as Tabelas 9.1 e 9.2 apresentam os resultados ótimos obtidos considerando o Exemplo 1 e o Exemplo 2, respectivamente.

Observando a Tabela 9.1 pode-se concluir que, para o funcional definido em (3.31), independente dos pontos iniciais o desempenho quanto ao número de iterações e precisão dos métodos que usam derivadas (Descida Máxima, Newton, Quase-Newton e Gradiente Conjugado) foi superior aos métodos independentes de derivada. O valor da função objetivo nos pontos ótimos encontrados com aqueles métodos está próximo a  $10^{-6}$  e, foram necessárias poucas iterações para que convergissem com precisão de  $10^{-3}$ .

O problema de otimização dado no Exemplo 2 é mais complexo por tem muitos mínimos locais, e por isto o algoritmo pode chegar a uma solução que não pertença a região viável do problema. Observando a Figura 3.15 é possível notar que o gráfico da função em torno do mínimo é uma região quase plana o que dificulta encontrar o ponto ótimo global. Analizando os valores ótimos apresentados na Tabela 9.2, verifica-se que as pressões  $p_2^*$  variam de 15,7813 à 22,9, entretanto os valores da função objetivo nestes pontos estão muito próximos, variando de 1,2315 à 1,2397 (excluindo o método de Newton que não convergiu).

Note que o menor valor da função objetivo encontrado é de 1,2315, sendo que o método dos Gradientes Conjugados apresentou o melhor resultado para este exemplo, convergindo com apenas 2 iterações.

A aplicação dos métodos de otimização irrestrita aos problemas (3.31) e (3.35) comprova a dificuldade para qualificar o quanto é bom ou ruim um método de otimização. Isto ficou claro quando foi utilizado o Método de Newton, este é um dos métodos mais famosos de otimização, mas sempre que a matriz Hessiana não for bem comportada, isto é, simétrica e positiva definida, o método encontrará dificuldades para convergir. Pode-se observar também que, quando a matriz Hessiana é mal condicionada, os métodos independentes de derivada

Método	nº Iterações	Critério de Parada	$(x^*)^T$	$f(x^*)$
Coord. Cíclicas	18	$ f(x^k)  < 0,01$	(1,975; 1,310)	0,0096
H.J Busca Linear	5	$ f(x^k)  < 0,01$	(1,913; 1,271)	0,005
H.J Passo Discreto	4	$ f(x^k)  < 0,01$	(1,2; 0,8)	0,002
Rosenbrock	6	$ f(x^k)  < 0,01$	(1,653; 1,102)	$1,8 \times 10^{-4}$
Descida Máxima	16	$\ \nabla f(x^k)\  < 0,001$	(1,3873; 0,9249)	$3,1791 \times 10^{-5}$
Newton	7	$\ \nabla f(x^k)\  < 0,001$	(1,53; 1,02)	$1,6782 \times 10^{-7}$
DFP	3	$\ \nabla f(x^k)\  < 0,001$	(1,5305; 1,0203)	$1,6995 \times 10^{-7}$
BFGS	3	$\ \nabla f(x^k)\  < 0,001$	(1,5507; 1,0338)	$1,3097 \times 10^{-6}$
G. Conjugados	3	$\ \nabla f(x^k)\  < 0,001$	(1,4198; 0,9466)	$8,11485 \times 10^{-6}$

Tabela 9.1: Resultados ótimos encontrados para o Exemplo 1, aplicando as técnicas estudadas

Método	n° Iterações	Critério de Parada	$(x^*)^T$	$f(x^*)$
Coord. Cíclicas	3	$\ f(x^{k+1}) - f(x^k)\  < 0,001$	(4,1132; 16,2249)	1,2315
H.J Busca Linear	3	$\ f(x^{k+1}) - f(x^k)\  < 0,001$	(3,4509; 14,8611)	1,2329
H.J Passo Discreto	9	$\ f(x^{k+1}) - f(x^k)\  < 0,001$	(4,8; 23)	1,2399
Rosenbrock	3	$\ f(x^{k+1}) - f(x^k)\  < 0,001$	(4,1159; 16,6040)	1,2316
Descida Máxima	5	$\ \nabla f(x^k)\  < 0,001$	(4,1185; 17,1572)	1,2318
Newton <sup>1</sup>	1000	1000 Iterações	(10,0534; 20,5196)	1,2926
DFP	6	$\ \nabla f(x^k)\  < 0,001$	(4,1237; 17,1489)	1,2318
BFGS	6	$\ \nabla f(x^k)\  < 0,001$	(3,9941; 15,9383)	1,2315
G. Conjugados	2	$\ \nabla f(x^k)\  < 0,001$	(3,9467; 16,0573)	1,2315

<sup>1</sup>O método de Newton não convergiu para o ponto ótimo

Tabela 9.2: Resultados ótimos encontrados para o Exemplo 2, aplicando as técnicas estudadas

tendem a convergir para o ótimo mais rápido do que os métodos que a utilizam. Contudo, para as funções cuja Hessiana é simétrica, positiva definida, que é o caso do Exemplo 1, os métodos que usam derivadas, além de serem mais rápidos, apresentam uma precisão superior.

## 9.1 Conclusões

Na presente dissertação foi desenvolvido o estudo de alguns métodos de otimização determinística irrestrita, aplicando seus algoritmos em dois problemas de otimização e analisando os resultados encontrados. Também foram apresentados o desenvolvimento da formulação matemática dos métodos e considerados os principais aspectos relacionados à convergência dos mesmos. Os passos de cada método foram implementados computacionalmente, à partir do qual também foram traçados os gráficos e curvas de nível das funções mencionadas. Assim, entende-se que os objetivos desta dissertação foram cumpridos.

Mas é intuitivo surgir a dúvida: por que estudar métodos irrestritos já que na prática os problemas de otimização tem algumas ou várias restrições? A justificativa vem do fato de que na otimização restrita muitos algoritmos transformam os problemas restritos em irrestritos, podendo ser resolvido por qualquer um dos métodos apresentados aqui. Como exemplo de tais algoritmos pode-se citar o Método dos Multiplicadores de Lagrange, o Método da Penalidade Quadrática e Métodos Seqüenciais. Desta forma, justifica-se o aprofundamento do estudo concernente aos métodos irrestritos, uma vez que são fundamentais para o desenvolvimento das técnicas de otimização restritas. Além disso, é natural propor o estudo dos métodos determinísticos de otimização restrita como trabalho futuro, pois representam uma continuidade desta pesquisa.

# Referências Bibliográficas

- [1] Arora, J.S., *Introduction to Optimum Design*, McGraw-Hill Book Company, Singapore, Roma, 1987.
- [2] Bazaraa, S.M., Sherali, H.D, Shetty, C.M., *Nonlinear Programming Theory and Algorithms*, John Wiley & Sons, Atlanta, 1979.
- [3] Beale, E.M.L., Nonlinear Programming Using a General Mathematical Programming System, in *Design and Implementation of Optimization Software*, H.J. Greenberg (Ed.), Sijthoff and Noordhoff, Netherlands, 1972.
- [4] Bergamaschi, P.R., Nogueira, A.C., Saramago, S.F.P., *Design and optimization of 3R manipulators using the workspace features*. Applied Mathematics and Computation, V. 172, pp. 439-463, 2006.
- [5] Bertsekas, D.P., *Nonlinear Programming*, Athena Scientific, Second Edition, Belmont, MA, 1999.
- [6] Boonnans, J.F., Gilbert, J.C., Lemaréchal, C., Sagastizábal, C., *Numerical Optimization. Theoretical and Practical Aspects*, Universitext, Springer-Verlag, Berlin, 2002.
- [7] Broydon, C.G., *The Convergence of a Class of Double Rank Minimization Algorithms*, parts I and II, J. Inst. Math. Appl., vol. 6, pp. 76-90, 222-231, 1970.
- [8] Bueno, L.F.C.R., *Medidas de Risco em Otimização de Portfólios*, UNIVERSIDADE ESTADUAL DE CAMPINAS - MATEMÁTICA APLICADA, Mestrado, 2008.
- [9] Carrera, D.H.Z., *Determinação da Trajetória de Veículos Terrestres a Alta Velocidade em Pistas Pré-definidas através de Técnicas de Otimização*, PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO - Engenharia Mecânica, Mestrado, 2006.
- [10] Davidon, W.C., *Variable metric method for minimization*, Technical Report ANL-5990 (revised), Argone National Laboratory, Argone, IL, 1959.
- [11] Davidon, W.C., *Variable metric method for minimization*, SIAM Journal on Optimization, pp. 1-7, 1 (1991).
- [12] Fletcher, R., *New Approach to Variable Metric Algorithms*, Computer J., vol. 13, pp. 317-322, 1970.
- [13] Fletcher, R., Powell, M.J.D., *A Rapidly Convergent Method for Minimization*, Computer J., vol.6, no.2, pp. 163-168, 1963.
- [14] Fletcher, R., and Reeves, C., *Function Minimization by Conjugate Gradients*, Computer Journal, 7, pp. 149-154, 1964.

- [15] Friedlander, A., *Elementos de Programação Não-Linear*, Editora da Unicamp, Campinas, SP, 1994.
- [16] Gill, P.E., Murray, W., Wright, M.H., *Practical Optimization*, Academic Press, Stanford University, California, USA, 1981.
- [17] Goldfarb, D., *A Family of Variable Metric Methods Derived by Variational Means*, Math. Comput., vol. 24, pp. 23-36, 1970.
- [18] Hestenes, M.R., and Stiefel, E., *Methods of Conjugate Gradients for Solving Linear Systems*, J. Research National Bureau of Standards, 49, pp. 409-436, 1952.
- [19] Himmelblau, D.M., *Applied Nonlinear Programming*, McGraw-Hill Book Company, University of Texas, Austin, Texas, USA, 1972.
- [20] Lasdon, L.S., *Optimization Theory of Large Systems*, Macmillan Company, New York, NY, 1970.
- [21] Lima, E.L., *Curso de Análise vol.2*, Rio de Janeiro: IMPA, 2006.
- [22] Luenberger, D.G., *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, second ed., 1984.
- [23] Liu, G.H., Jing, L.L., Han, L.X., Han, D., *A Class of Nonmonotone Conjugate Gradient Methods for Unconstrained Optimization*, Journal of Optimization Theory and Applications, Springer Netherlands, Volume 101, Number 1 / April, 1999.
- [24] Nocedal, J., Wright, S.J., *Numerical Optimization*, Springer Series in Operations Research, 2000.
- [25] Pirsig, R.M., *Zen and the Art of Motorcycle Maintenance: An Inquiry into Values*, New York: Quill. 25th Anniversary Edition, 1974.
- [26] Polak, E., and Ribiere, G. , *Note sur la Convergence de Methods de Directions Conjugues*, Revue Francaise Informat, Recherche Operationelle, 16, pp. 35-46, 1969.
- [27] Powell, M.J.D., *Restart Procedures for the Conjugate Gradient Method*, Mathematical Programming, 12, pp. 241-254, 1977.
- [28] Santos, R.R., Steffen Jr., V., Saramago, S.F.P., *Robot Path Planning in a Constrained Workspace by using Optimal Control Techniques*, Multibody System Dynamics, V. 19, pp. 159-177, 2008.
- [29] Shanno D.F., *Conditioning of Quasi-Newton Methods for Function Minimization*, Math. Comput., vol. 24, pp. 647-656, 1970.
- [30] Shi, Y., *Globally Convergent Algorithms for Unconstrained Optimization*, Journal Computational Optimization and Applications, Springer Netherlands, Volume 16, Number 3 / September, 2000.
- [31] Vanderplaats, G., *Numerical Optimization Techniques for Engineering Design*, McGraw-Hill, USA, 1984.
- [32] Varela, M.E.M., *Formulación y solución de problemas de programación no lineal en el diseño estructurado de amplificadores*, LANIA, México, 2006.