

UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE CIÊNCIAS EXATAS
DEPARTAMENTO DE MATEMÁTICA

Giovanni Escóssio Zanardo

**Uma comparação entre alguns Métodos de Minimização Irrestrita
utilizando Perfil de Desempenho**

**MANAUS – AM
2022**

Giovanni Escóssio Zanardo

**Uma comparação entre alguns Métodos de Minimização Irrestrita
utilizando Perfil de Desempenho**

Trabalho de Conclusão de Curso de Graduação em Matemática submetido ao Departamento de Matemática da Universidade Federal do Amazonas, como requisito necessário para obtenção do grau de Bacharel em Matemática.

Orientador: Prof. Dr. Sandro Dimy Barbosa Bitar

**MANAUS – AM
2022**

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

Zanardo, Giovanni Escóssio
Z27c Uma comparação entre alguns Métodos de Minimização Irrestrita
utilizando Perfil de Desempenho / Giovanni Escóssio Zanardo .
2022
65 f.: il. color; 31 cm.

Orientador: Sandro Dimy Barbosa Bitar
TCC de Graduação (Matemática) - Universidade Federal do
Amazonas.

1. Programação Não-Linear Irrestrita. 2. Método do Gradiente. 3.
Método de Newton. 4. Métodos Quase-Newton. 5. Método dos
Gradientes Conjugados . I. Bitar, Sandro Dimy Barbosa. II.
Universidade Federal do Amazonas III. Título

*Dedico à
minha família.*

AGRADECIMENTOS

Agradeço à minha avó Nazaré(in memoriam), ao meu avô Francisco(in memoriam), à minha mãe Ana, à minha tia Márcia, ao meu tio Paulo e ao meu tio Luís pelo apoio familiar e financeiro dado ao longo desses vários anos. Por terem desde o início incetivado ao estudo e mostrado a mim o valor do conhecimento como elemento primordial para a transformação do ser. Sem eles, nada teria conseguido.

Agradeço aos colegas de curso pelos muitos diálogos sobre as disciplinas e sobre a vida além da Matemática. Posso afirmar que com a companhia deles os desafios da graduação tornaram-se mais leves e suaves. Em especial, ao Vinícius Rosário das Chagas, Luís Filipe Vital França, Pedro Piovan Silva Cardoso, Samiles Lima de Sousa, Lucas Oliveira Rodrigues, Henrique Lopes Borges e Luccas Amorim Rocha.

Agradeço aos excelentes professores que tive ao longo da minha formação, em especial aos professores Sandro Bitar, Flávia Morgana e Roberto Prata que tanto investiram seus tempos, esforços e amparos no desenvolvimento do meu potencial dentro do âmbito acadêmico, além de se constituírem verdadeiros modelos para a minha formação profissional.

Agradeço aos funcionários da Universidade Federal do Amazonas e todos aqueles aos quais não citei, mas que de alguma maneira contribuíram para eu chegar até aqui.

*“Não é o conhecimento, mas o ato de aprender,
não a posse, mas o ato de chegar lá,
que concede a maior satisfação.”
(Carl Friedrich Gauss)*

RESUMO

O enfoque principal deste trabalho é a apresentação de métodos para a resolução do Problema de Programação Não-Linear Irrestrita. Diante desta perspectiva, o Método do Gradiente, o Método de Newton, o DFP, o BFGS e o Método dos Gradientes Conjugados foram analisados. Por fim, foram realizadas avaliações computacionais destes métodos através do estudo de dez problemas utilizando perfil de desempenho.

Palavras-chaves: Programação Não-Linear Irrestrita; Método do Gradiente; Método de Newton; DFP; BFGS; Método dos Gradientes Conjugados; Perfil de Desempenho.

ABSTRACT

The main focus of this work is the presentation of methods for solving the Unrestricted Nonlinear Programming Problem. In view of this perspective, the Gradient Method, the Newton Method, the DFP, the BFGS and the Conjugate Gradient Method were analyzed. Finally, computational evaluations of these methods were performed by studying ten problems using performance profile.

Keywords: Unrestricted Nonlinear Programming; Gradient Method; Newton's method; DFP; BFGS; Conjugate Gradient Method; Performance Profile.

SUMÁRIO

Introdução	10
1 Conceitos Básicos e Resultados Iniciais	12
1.1 O Problema de Programação Não-Linear	12
1.2 Condições de Otimalidade para Minimizações Sem Restrições	13
1.3 Convexidade	16
1.4 Algoritmo Geral de Busca Linear	20
1.4.1 Algoritmo de descida com backtracking	21
1.5 Ordem ou razão de convergência	24
2 Métodos Clássicos de Descida	25
2.1 Método do Gradiente	25
2.1.1 Algoritmo(Método do Gradiente)	25
2.1.2 O Caso Não-Quadrático	29
2.2 Método de Newton	30
2.2.1 O Caso Quadrático	30
2.2.2 Caso Não-Quadrático	30
2.2.3 Algoritmo(Método de Newton)	31
2.2.4 Convergência do Método de Newton	32
2.3 Métodos Quase-Newton	33
2.3.1 Algoritmo(DFP)	34
2.3.2 Análise da Obtenção do DFP	42
2.3.3 Atualizações Família Broyden e BFGS	44
2.4 Método dos Gradientes Conjugados	45
2.4.1 Abordagem Básica do Método dos Gradientes Conjugados	46
2.4.2 Algoritmo(Método dos Gradientes Conjugados)	48
2.4.3 Caso Quadrático	48
2.4.4 Métodos Quase-Newton Sem Memória	51
2.4.5 Convergência dos Algoritmos com Direções Conjugadas	52
2.4.5.1 Composição de Aplicações	52
2.4.5.2 Convergência do Método dos Gradientes Conjugados	54
2.4.5.3 Característica da Razão de Convergência do Método dos Gra- dientes Conjugados	55
3 Avaliações Computacionais	57
3.1 Perfil de Desempenho	57
3.1.1 Análise do Perfil de Desempenho	58
3.1.2 Gráficos de Perfil de Desempenho	59
3.1.3 Implementação dos Algoritmos e Especificações	60

3.1.4	Banco de Problemas	60
3.1.5	Pontos Iniciais Arbitrados	61
3.1.6	Perfil de Desempenho dos Métodos com Ponto Inicial Próximo da Solução	61
3.1.7	Perfil de Desempenho dos Métodos com Ponto Inicial Distante da Solução	62
4	Conclusão	64
	REFERÊNCIAS	65

INTRODUÇÃO

Neste trabalho registram-se assuntos relacionados ao problema de Programação Não-Linear Irrestrita. Assim sendo, primeiramente, define-se o que seria este problema e suas respectivas condições de otimalidade, onde apresentam-se lemas e teoremas como as condições necessárias de primeira e segunda ordem, e as condições suficientes de segunda ordem. O que é relevante, pois se o problema em questão cumpre as hipóteses de tais condições, já teremos resultados que nos possibilitam inferir se certos pontos constituem-se minimizadores ou, pelo menos, candidatos a minimizadores.

Além disso, certos tópicos em análise convexa são abordados, de forma que resultados relacionados a minimizadores de funções convexas também fossem devidamente descritos. De uma certa maneira, veremos que o conhecimento prévio da convexidade de dada função é relevante, pois se no problema $f : \mathbb{R}^n \rightarrow \mathbb{R}$ for convexa, temos que $x^* \in \mathbb{R}^n$ é minimizador global da f se, e somente se, $\nabla f(x^*) = 0$. Ou seja, o ponto estacionário da f é a solução do problema de Programação Não-Linear Irrestrita.

Entretanto, nem todas as funções $f : \mathbb{R}^n \rightarrow \mathbb{R}$ apresentam a força das propriedades de funções convexas. Por conseguinte, define-se uma direção de descida $d \in \mathbb{R}^n$ tal que $f(x + \alpha d) < f(x)$, $\alpha > 0$. Ademais, um fato relevante a saber é que esta direção $d \in \mathbb{R}^n$ sempre existe se certas hipóteses são satisfeitas. Com efeito, mostra-se que se $f \in C^1$ e $\nabla f(x) \neq 0$, então existe $\bar{\alpha} \in \mathbb{R}$, onde $f(x + \alpha d) < f(x)$, se $\alpha \in (0, \bar{\alpha})$. Com esta informação em mãos, é necessário também saber como se calcula $\alpha \in \mathbb{R}$, processo chamado de cálculo do tamanho do passo que propicie o decréscimo no valor da função. Nesse quesito, existem duas metodologias, ou realiza-se uma busca linear exata, como o método da Seção Áurea exposto em (MARTÍNEZ; AUGUSTA, 1998), ou utiliza-se um método de busca linear inexata, como a Condição de Armijo que será apresentada e preferencialmente usada neste trabalho.

Dessa maneira, um algoritmo com boas propriedades de convergência global, o chamado Algoritmo de Busca Linear (MARTÍNEZ; AUGUSTA, 1998) é descrito, ou seja, diante de determinadas condições impostas quanto à direção d_k em cada k -ésima iteração, garantimos que no término do procedimento, o algoritmo apresenta pontos estacionários ou gera uma sequência infinita cujo os pontos de acumulação desta são pontos estacionários. Logo, pelas condições necessárias de primeira ordem, temos que o método exhibe candidatos a minimizadores globais.

Por consequência, a bibliografia clássica descreve variadas metodologias de busca direcional que visam o decréscimo no valor da f . Para $d_k = -\nabla f(x_k)$ temos o Método do Gradiente. Por outro lado, assumindo que a hessiana da f seja invertível em $x_k \in \mathbb{R}^n$ e escolhendo $d_k = -(\nabla^2 f(x_k))^{-1} \nabla f(x_k)$, tem-se o Método de Newton que possui convergência quadrática local. Porém, o método é computacionalmente caro, pois a hessiana da função ne-

cessita ser calculada a cada passo. Logo, o ideal é uma metodologia intermediária, onde uma aproximação da hessiana da f seja usada, havendo ainda o aproveitamento desta para o cálculo do passo seguinte. Estes são os chamados Métodos Quase-Newton, sendo $d_k = -H_k \nabla f(x_k)$, H_k de ordem n , simétrica e positiva definida. No entanto, os Métodos Quase-Newton possuem limitações ao tratarem de problemas cujo o domínio seja de dimensões elevadas (BAZARAA, 2006), como $n \geq 100$, visto que o cálculo da aproximação da hessiana possui elevado custo computacional. Assim, diante deste cenário, o Método dos Gradientes Conjugados apresenta particular vantagem, pois armazena 3 vetores n -dimensionais a cada iteração, isto é, a memória envolvida no procedimento é bastante inferior se comparado aos Quase-Newton.

Ao término, após o desenvolvimento teórico, uma avaliação dos algoritmos implementados em Python foi realizada com base em dois cenários distintos, sendo estes, pontos iniciais próximos da solução e pontos distantes da mesma, visando assim medir e analisar a eficiência dos métodos quanto ao seu tempo de processamento e sua robustez por meio de ferramentas denominadas perfis de desempenho, onde estas foram geradas por meio da linguagem de programação Julia.

1. Conceitos Básicos e Resultados Iniciais

1.1 O Problema de Programação Não-Linear

Neste trabalho, investigamos o seguinte problema de Programação Não-Linear:

$$\begin{cases} f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \\ \text{minimizar } f(x) \\ \text{sujeita a } x \in S \end{cases}$$

Definição 1. Diz-se que $x^* \in S$ é minimizador global em f em S se $f(x) \geq f(x^*), \forall x \in S$.

Definição 2. Afirma-se que x^* é minimizador local de f em S se $\exists \epsilon > 0$ tal que $f(x) \geq f(x^*), \forall x \in S, \|x - x^*\| < \epsilon$. Caso $f(x) > f(x^*), \forall x \in S, x \neq x^*, \|x - x^*\| < \epsilon$, diz-se que x^* é minimizador local estrito em S .

É interessante revelar previamente a importância do primeiro teorema a seguir, visto que ele garantirá a presença de máximos e mínimos globais em certos problemas de Programação Não-Linear, havendo assim sentido em realizar buscas por minimizadores de funções definidas em conjuntos compactos. Entretanto, note que tal teorema revela condições suficientes para existência de extremantes globais em funções contínuas com domínios compactos, ou seja, embora as funções expostas neste trabalho tenham como domínio todo o \mathbb{R}^n (não compacto), isto não implica no fato de a função não possuir um valor mínimo global, como é trivialmente observado na função contínua $g : \mathbb{R} \rightarrow \mathbb{R}, g(x) = x^2$. Portanto, exemplos como este motivam a procura de métodos que envolvam minimizar funções contínuas cujos domínios sejam conjuntos que não possuam características como a compacidade.

Teorema 1.1. (Teorema de Weierstrass) Seja $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$, onde A é compacto. Se f é contínua em A , então existem $x_1, x_2 \in A$ tais que $f(x_1) \leq f(x) \leq f(x_2), \forall x \in A$.

Demonstração. Para esta prova faremos uso das notações e definições expostas em (LIMA, 2020). Primeiramente, seja $(y_k)_{k \in \mathbb{N}}$ uma sequência de pontos com $y_k \in f(A)$. Consequente-

mente, existe uma sequência de pontos $(x_k)_{k \in \mathbb{N}}$ com $x_k \in A$ tal que $y_k = f(x_k)$ para cada $k \in \mathbb{N}$. Considere $\mathbb{N}' \subset \mathbb{N}$. Como A é compacto, então existe uma subsequência $(x_k)_{k \in \mathbb{N}'}$ que converge para um elemento $a \in K$, isto é, $\lim_{k \in \mathbb{N}'} x_k = a$. Pela continuidade de f , segue que $\lim_{k \in \mathbb{N}'} f(x_k) = f(a) \implies \lim_{k \in \mathbb{N}'} y_k = \lim_{k \in \mathbb{N}'} f(x_k) = a$, logo mostrou-se que uma sequência arbitrária $(y_k)_{k \in \mathbb{N}}$ de $f(A)$ possui uma subsequência $(y_k)_{k \in \mathbb{N}'}$ que converge para um ponto de $f(A)$. Portanto, $f(A)$ é compacto. Finalmente, já que $f(A) \subset \mathbb{R}$ é compacto, então sua limitação garante a existência de supremo e ínfimo para $f(A)$. Ademais, o fato de A ser fechado permite afirmar que $\inf f(A)$ e $\sup f(A)$ pertencem ao conjunto $f(A)$, logo existem x_1 e x_2 em A tais que $f(x_1) = \inf f(A)$ e $f(x_2) = \sup f(A)$. Em uma palavra, $f(x_1) \leq f(x) \leq f(x_2), \forall x \in A$. □

1.2 Condições de Otimalidade para Minimizações Sem Restrições

Diante da perspectiva comentada na seção anterior, um problema de Programação Não-Linear cujo o objetivo é minimizar uma função de domínio $S = \mathbb{R}^n$ é chamado de **problema de minimização irrestrita**. Em suma, tem-se o seguinte:

$$\begin{cases} \text{minimizar } f(x) \\ \text{sujeita a } x \in \mathbb{R}^n \end{cases}$$

Portanto, esta seção será encarregada de apresentar resultados relevantes e definições pertinentes a este problema.

Definição 3. *Seja $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, onde S é aberto. Diz-se que f é de classe C^n se as derivadas parciais de f forem contínuas até a n -ésima ordem em S .*

Lema 1.2.1. *Se $f : \mathbb{R} \rightarrow \mathbb{R}, f \in C^1$. Se x^* é extremante local, então $f'(x^*) = 0$.*

Demonstração. Encontra-se em (FRIEDLANDER, s.d.) □

Lema 1.2.2. *Se $f : \mathbb{R} \rightarrow \mathbb{R}, f \in C^2$. Se x^* é minimizador local de f em \mathbb{R} , então $f'(x^*) = 0$ e $f''(x^*) \geq 0$.*

Demonstração. Encontra-se em (FRIEDLANDER, s.d.) □

Definição 4. *Uma matriz $A \in M_n(\mathbb{R})$ simétrica ($A = A^T$) é **semidefinida positiva** se $x^t A x \geq 0 \forall x, x \neq 0, x \in \mathbb{R}^n$. Define-se de maneira análoga, A **definida positiva** ($x^t A x > 0$), A **definida negativa** ($x^t A x < 0$) e A **semidefinida negativa** ($x^t A x \leq 0$).*

Definição 5. *(Gradiente de uma função) Seja $f \in C^1$ uma função de valor real definida em \mathbb{R}^n . O gradiente da f é dado por $\nabla f(x) = f_x(x) = \left[\frac{\partial f(x)}{\partial x_i} \right]$.*

Definição 6. (Hessiano de uma função) Seja $f \in C^2$, define-se o hessiano (matriz hessiana) de f em $x \in \mathbb{R}^n$, como uma matriz quadrada de ordem n explicitada como

$$\nabla^2 f(x) = \left[\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right] = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix} \text{ com } i = 1, \dots, n \text{ e } j = 1, \dots, n.$$

Observação 2. Ainda supondo que f seja de classe C^2 , note que o **Teorema de Schwarz** (LIMA, 2020) será válido, isto é, $\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}$. Portanto, tem-se que a **matriz hessiana é simétrica**.

Lema 1.2.3. (Fórmula de Taylor com resto de Lagrange de ordem $n + 1$ (versão em \mathbb{R})) Seja $f : I \rightarrow \mathbb{R}$ derivável até a ordem $n + 1$ no intervalo I e sejam $x \in I$, $x_0 \in I$. Então existe um $\bar{x} \in (x_0, x)$ tal que:

$$f(x) = P(x) + \frac{f^{(n+1)}(\bar{x})}{(n+1)!} (x - x_0)^{n+1}, \text{ onde}$$

$$P(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)(x - x_0)^2}{2!} + \dots + \frac{f^{(n)}(x_0)(x - x_0)^n}{n!} = \sum_{k=0}^n \frac{f^{(k)}(x_0)(x - x_0)^k}{k!}$$

Demonstração. Encontra-se em (LIMA, 2008) □

Lema 1.2.4. (Fórmula de Taylor de ordem 1 para funções de valor real com várias variáveis) Seja f de classe C^2 , $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e sejam $x_0 \in A$ e $h \neq 0$ tais que o segmento de extremidades x_0 e $x_0 + h$ esteja contido em A . Assim:

$$f(x_0 + h) = f(x_0) + h^T \nabla f(x_0) + \frac{1}{2} h^T H(x_0) h + E(h), \quad H(x) = \nabla^2 f(x)$$

Onde $E(h) = \sum_{k=1}^n \sum_{r=1}^n \frac{dx_k}{dt} \frac{\partial^2 f(\bar{x})}{\partial x_r \partial x_k}$, sendo \bar{x} do interior do segmento de extremos x_0 e $x_0 + h$.

Demonstração. Encontra-se em (LIMA, 2020) □

Teorema 1.2. (Condições necessárias de primeira ordem) Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^1$. Se x^* é minimizador local da f em \mathbb{R}^n , então $\nabla f(x^*) = 0$.

Demonstração. Defina $\phi : \mathbb{R} \rightarrow \mathbb{R}$ tal que $\phi(\lambda) = f(x^* + \lambda d)$, $d \in \mathbb{R}^n$, d arbitrário. Note que $\phi'(\lambda) = \nabla^T f(x^* + \lambda d)(x^* + \lambda d)' = \nabla^T f(x^* + \lambda d)d$. Como x^* é extremante local da f ,

e $\phi(\lambda) = f(x^* + \lambda d)$, segue que $\lambda = 0$ possibilita que $\phi(0) = f(x^* + 0d) = f(x^*)$, portanto $\lambda = 0$ é minimizador de $\phi : \mathbb{R} \rightarrow \mathbb{R}$. Como $\phi \in C^1$, decorre que $\phi'(0) = 0$. Lembre que $\phi'(\lambda) = \nabla^T f(x^* + \lambda d)d$, portanto $\phi'(\lambda) = \nabla^T f(x^* + \lambda d)d \implies \phi'(0) = \nabla^T f(x^* + 0d)d = [\nabla^T f(x^*)]d$ (pois $\lambda = 0$). Logo, $\phi'(0) = \nabla^T f(x^*)d$. Como $\phi'(0) = 0 \implies \nabla^T f(x^*)d = 0$. Visto que $d \in \mathbb{R}^n$ e arbitrário, tome $d = \nabla f(x^*) \in \mathbb{R}^n \implies \nabla^T f(x^*)\nabla f(x^*) = \|\nabla^T f(x^*)\|^2 = 0 \iff \nabla f(x^*) = 0$ (Recorde que dado um vetor $u \in \mathbb{R}^n$, tem-se que $\|u\|^2 = 0 \iff u = 0$). \square

Teorema 1.3. (Condições necessárias de segunda ordem) Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^2$. Se x^* é minimizador local da f em \mathbb{R}^n , então:

(i) $\nabla f(x^*) = 0$.

(ii) $\nabla^2 f(x^*)$ é semidefinida positiva.

Demonstração. Note que a parte (i) é o **Teorema 1.2.** já mostrado. Para a parte ii), tome $\phi : \mathbb{R} \rightarrow \mathbb{R}$ tal que $\phi(\lambda) = f(x^* + \lambda d)$, onde $d \in \mathbb{R}^n$ e $d \neq 0$ arbitrário. Ora, $\phi'(\lambda) = \nabla^T f(x^* + \lambda d) \cdot (x^* + \lambda d)' = \nabla^T f(x^* + \lambda d) \cdot d$. Consequentemente,

$$\begin{aligned} [\phi'(\lambda)]' &= [\nabla^T f(x^* + \lambda d)]' \cdot d \\ &= [d^T \nabla(\nabla f(x^* + \lambda d))] \\ &= d^T \nabla^2 f(x^* + \lambda d) d \end{aligned}$$

Como $\lambda = 0$ é minimizador da ϕ , pois $\phi(0) = f(x^*)$, e x^* é minimizador local da f , tem-se pelo **Lema 1.2.2** que $[\phi'(\lambda)]' = \phi''(\lambda) \geq 0$. Já que $\lambda = 0 \implies \phi''(\lambda) = \phi''(0) = d^T \nabla^2 f(x^* + 0d) d = d^T \nabla^2 f(x^*) d \geq 0, \forall d \neq 0, d \in \mathbb{R}^n \implies \nabla^2 f(x^*)$ é semidefinida positiva por definição. \square

Teorema 1.4. (Condições suficientes de 2ª ordem) Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$, f de classe C^2 . Se $x^* \in \mathbb{R}^n$, $\nabla f(x^*) = 0$ e $\nabla^2 f(x^*) > 0$, então tem-se que x^* é minimizador local estrito da f .

Demonstração. Inicialmente, tome $B = \{h \in \mathbb{R}^n, \|h\| = 1\}$. Defina $\Gamma : B \rightarrow \mathbb{R}$ tal que $\Gamma(h) = h^T \nabla^2 f(x^*) h$. Note que Γ é contínua num compacto B , logo pelo **Teorema de Bolzano-Weierstrass**, sabe-se que Γ tem máximos e mínimos em B . Assim, se a é o valor mínimo, então $\forall h \in B$, tem-se que $\Gamma(h) \geq a \implies h^T \nabla^2 f(x^*) h \geq a$. Por hipótese, sabe-se que $\nabla^2 f(x^*)$ é definida positiva. Logo, sendo $x^* \in \mathbb{R}^n$, $x^* \neq 0$, prossegue que $\Gamma(h) = h^T \nabla^2 f(x^*) h \geq a > 0, \forall h \in B$. Agora, considere $d \in \mathbb{R}^n, d \neq 0$, logo $\frac{d}{\|d\|} \in B$, logo $\frac{d}{\|d\|} \in B \implies \Gamma\left(\frac{d}{\|d\|}\right) \geq a > 0 \implies \frac{d^T}{\|d\|} \nabla^2 f(x^*) \frac{d}{\|d\|} \geq a > 0 \implies d^T \nabla^2 f(x^*) d \geq a \|d\|^2 > \|d\|^2 > 0$. Desenvolvendo a série de Taylor de $f : \mathbb{R}^n \rightarrow \mathbb{R}$ em torno de x^* segue que $f(x^* + d) = f(x^*) + \nabla^T f(x^*) d + \frac{1}{2} d^T \nabla^2 f(x^*) d + O(\|d\|^2)$. Recorde da hipótese $\nabla^T f(x^*) = 0$, portanto $f(x^* + d) = f(x^*) +$

$\nabla^T f(x^*)d + \frac{1}{2}d^T \nabla^2 f(x^*)d + O(\|d\|^2) = f(x^*) + 0d + \frac{1}{2}d^T \nabla^2 f(x^*)d + O(\|d\|^2) = f(x^*) + \frac{1}{2}d^T \nabla^2 f(x^*)d + O(\|d\|^2) \implies f(x^* + d) - f(x^*) = \frac{1}{2}d^T \nabla^2 f(x^*)d + O(\|d\|^2)$. Sabe-se que $d^T \nabla^2 f(x^*)d \geq a\|d\|^2$, logo $f(x^* + d) - f(x^*) = \frac{1}{2} [d^T \nabla^2 f(x^*)d] + O(\|d\|^2) \geq \frac{1}{2}[a\|d\|^2] + O(\|d\|^2)$, lembre que $a > 0$ e $d \neq 0$. Assim, para um d suficientemente pequeno e não-nulo, isto é, um elemento $d \in \mathbb{R}^n$, onde $0 < \|d\| < \epsilon$, segue que $f(x^* + d) - f(x^*) > 0$. Ou seja, para todo $x \in B(x^*; \epsilon)$, $x \neq x^*$, tem-se que $f(x) > f(x^*)$. Em uma palavra, x^* é minimizador local estrito da f . \square

1.3 Convexidade

Após a apresentação das condições de otimalidade para uma dada função com domínio em \mathbb{R}^n , o fato de saber se uma função definida em ambiente convexo é ou não convexa agrega certas vantagens na procura por minimizadores globais, pois como será visto, as condições necessárias de primeira ordem torna-se-ão suficientes para afirmar se determinado ponto é ou não minimizador global. Portanto, uma informação relevante a se saber sobre determinada função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é se ela é localmente convexa em uma bola aberta centrada na solução do problema, isto é, se existe um $\epsilon > 0$ tal que para todo elemento $y \in B(x^*; \epsilon)$, onde x^* é a solução do problema de minimização, a definição de convexidade vale. Logo, o conhecimento da presença de convexidade local em uma vizinhança de x^* permitirá inferir aspectos globais da função quanto aos seus minimizadores.

Definição 7. (Conjunto Convexo) $S \subset \mathbb{R}^n$ é dito um conjunto convexo se $\forall x, y \in S$ tem-se que:

$$\lambda x + (1 - \lambda)y \in S, \lambda \in [0, 1]$$

Em outras palavras, se os pontos x e y estiverem no conjunto S e, sendo S convexo, segue que o segmento de extremos x e y estará contido em S .

Definição 8. (Função Convexa) $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, onde Ω é um conjunto convexo, é dita convexa se, e só se, $\forall x_1, x_2 \in \Omega$ e para qualquer $\alpha \in [0, 1]$, vale a seguinte expressão:

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2)$$

Se $f(\alpha x_1 + (1 - \alpha)x_2) < \alpha f(x_1) + (1 - \alpha)f(x_2) \forall \alpha, 0 < \alpha < 1$ e $x_1 \neq x_2$, então diz-se que f é estritamente convexa.

Geometricamente, uma função $f : \Omega \rightarrow \mathbb{R}$ é convexa se ao selecionar dois pontos $f(x_1)$ e $f(x_2)$ com x_1 e x_2 em Ω , tem-se que o segmento que une os pontos $(x_1, f(x_1))$ e $(x_2, f(x_2))$ não encontra-se em nenhum local abaixo do gráfico.

Lema 1.3.1. Sejam f_1 e f_2 funções convexas em um conjunto convexo Ω , então $f_1 + f_2$ é convexa.

Demonstração. Sejam $x_1, x_2 \in \Omega$ e $0 \leq \alpha \leq 1$, decorre que $(f_1 + f_2)(\alpha x_1 + (1 - \alpha)x_2) = [f_1(\alpha x_1 + (1 - \alpha)x_2)] + [f_2(\alpha x_1 + (1 - \alpha)x_2)] \leq [\alpha f_1(x_1) + (1 - \alpha)f_1(x_2)] + [\alpha f_2(x_1) + (1 - \alpha)f_2(x_2)] = \alpha(f_1(x_1) + f_2(x_1)) + (1 - \alpha)(f_1(x_2) + f_2(x_2)) = \alpha(f_1 + f_2)(x_1) + (1 - \alpha)(f_1 + f_2)(x_2)$. Logo, $(f_1 + f_2)(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha(f_1 + f_2)(x_1) + (1 - \alpha)(f_1 + f_2)(x_2)$. Portanto, $f_1 + f_2$ é uma função convexa. \square

Lema 1.3.2. *Seja f uma função convexa sobre um conjunto convexo Ω , então af é convexa $\forall a \in \mathbb{R}, a \geq 0$.*

Demonstração. Seja f uma função convexa sobre um conjunto convexo Ω , então $f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) \forall x_1, x_2 \in \Omega, \alpha \in [0, 1]$. Assim, $f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) \implies af(\alpha x_1 + (1 - \alpha)x_2) \leq a(\alpha f(x_1) + (1 - \alpha)f(x_2)) \implies (af)(\alpha x_1 + (1 - \alpha)x_2) \leq a\alpha f(x_1) + a(1 - \alpha)f(x_2) \implies (af)(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha(af)(x_1) + (1 - \alpha)(af)(x_2)$. Consequentemente, a função af é convexa. Pelo lema anterior, perceba que a função $\sum_{i=1}^m a_i f_i$ será convexa, caso $a_i \geq 0$ e se f_i for convexa em Ω para cada i . \square

Lema 1.3.3. *Seja f uma função convexa em um conjunto convexo Ω . O conjunto $\Gamma_c = \{x \in \Omega, f(x) \leq c\}$ é convexo para todo número real c .*

Demonstração. Considere $x_1, x_2 \in \Gamma_c \in \Omega \implies f(x_1) \leq c$ e $f(x_2) \leq c$. Ora, para $\alpha \in [0, 1]$, tem-se que $f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) \leq \alpha c + (1 - \alpha)c = \alpha c - \alpha c + c = c \implies \alpha x_1 + (1 - \alpha)x_2 \in \Gamma_c$. Consequentemente, $\alpha x_1 + (1 - \alpha)x_2$ pertence ao conjunto Γ_c para qualquer $\alpha \in [0, 1]$. Assim, conclui-se que Γ_c é convexo. \square

O lema anterior possui sua importância para com a programação matemática, porque frequentemente o conjunto factível é definido por meio de restrições explicitadas em forma de desigualdades. Na programação linear, por exemplo, temos um poliedro. Tais conjuntos são definidos por meio de desigualdades lineares. Precisamente, poliedros são escritos na forma $\{x \in \mathbb{R}^n; Ax \geq b\}$, onde A é uma matriz $m \times n$ e b é um vetor $m \times 1$. Note que tais conjuntos serão convexos já que são a interseção dos semiespaços $a_i^T x \geq b_i$, sendo a_i a i -ésima linha da matriz A .

Lema 1.3.4. *Seja S_i um conjunto convexo para cada i ($i = 1, \dots, n$). Então $\bigcap_{i=1}^n S_i$ é um conjunto convexo.*

Demonstração. $x_1, x_2 \in \bigcap S_i, \alpha \in [0, 1] \implies \alpha x_1 + (1 - \alpha)x_2 \in S_1, \alpha x_1 + (1 - \alpha)x_2 \in S_2, \dots, \alpha x_1 + (1 - \alpha)x_2 \in S_n \implies \alpha x_1 + (1 - \alpha)x_2 \in \bigcap S_i$. Portanto, $\bigcap S_i$ é um conjunto convexo. \square

Como foi mostrado que $\bigcap_{i=1}^n \Gamma_{c_i}$, sendo $\Gamma_{c_i} = \{x \in \Omega, f_i(x) \leq c_i\}$, e f_i é convexa em Ω , segue que $\bigcap_{i=1}^n \Gamma_{c_i}$ é um conjunto convexo.

Definição 9. (Derivada direcional) Considere $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$, sendo U aberto do \mathbb{R}^n , $a \in U$ e $v \in \mathbb{R}^n$. A derivada direcional do ponto $a \in U$ segundo o vetor $v \in \mathbb{R}^n$ é

$$\frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t}$$

caso o limite acima exista.

Teorema 1.5. Seja f diferenciável em $a \in U$, logo f é contínua em a e f admite derivadas parciais em $a \in U$. Assim, f admite derivada direcional em qualquer $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ e vale:

$$\frac{\partial f}{\partial v}(a) = \frac{\partial f(a)}{\partial x_1} \alpha_1 + \frac{\partial f(a)}{\partial x_2} \alpha_2 + \dots + \frac{\partial f(a)}{\partial x_n} \alpha_n = \sum_{i=1}^n \frac{\partial f(a)}{\partial x_i} \alpha_i = \nabla^T f(a)v$$

Demonstração. Encontra-se em (LIMA, 2020) □

Teorema 1.6. Seja f de classe C^1 . Afirma-se que a função f é convexa em um conjunto convexo $\Omega \subset \mathbb{R}^n$ se, e só se, $f(y) \geq f(x) + \nabla^T f(x)(y - x), \forall x, y \in \Omega$.

Demonstração. Suponha que f seja convexa, logo $\forall \alpha \in \mathbb{R}$ e $0 \leq \alpha \leq 1$:

$$f(\alpha y + (1 - \alpha)x) \leq \alpha f(y) + (1 - \alpha)f(x), \text{ sendo } x, y \in \Omega$$

Portanto, como a relação é válida para $0 \leq \alpha \leq 1$, então ela será ela será válida para $0 < \alpha \leq 1$.

$$\begin{aligned} \text{Logo, } f(\alpha y + (1 - \alpha)x) \leq \alpha f(y) + (1 - \alpha)f(x) &\implies \frac{f(\alpha y + (1 - \alpha)x)}{\alpha} - \frac{f(x)}{\alpha} \leq f(y) - f(x) \\ &\implies \frac{f(x + \alpha(y - x)) - f(x)}{\alpha} \leq f(y) - f(x) \implies \nabla^T f(x)(y - x) \leq f(y) - f(x), \end{aligned}$$

devido ao lema anterior. Assim, confirma-se a primeira parte.

Por outro lado, assuma que $f(y) \geq f(x) + \nabla^T f(x)(y - x), \forall x, y \in \Omega$. Se $x_1, x_2 \in \Omega$, como Ω é convexo, segue que $x = \alpha x_1 + (1 - \alpha)x_2$. Consequentemente, as relações abaixo são válidas:

$$\begin{cases} (1) f(x_1) \geq f(x) + \nabla^T f(x)(x_1 - x) \\ (2) f(x_2) \geq f(x) + \nabla^T f(x)(x_2 - x) \end{cases}$$

Multiplique (1) por α e (2) por $(1 - \alpha)$, ou seja:

$$\begin{cases} (3) \alpha f(x_1) \geq \alpha f(x) + \alpha \nabla^T f(x)(x_1 - x) \\ (4) (1 - \alpha) f(x_2) \geq (1 - \alpha) f(x) + (1 - \alpha) \nabla^T f(x)(x_2 - x) \end{cases}$$

Somando (3) e (4):

$$\alpha f(x_1) + (1 - \alpha)f(x_2) \geq [\alpha f(x) + \alpha \nabla^T f(x)(x_1 - x)] + [(1 - \alpha)f(x) + (1 - \alpha)\nabla^T f(x)(x_2 - x)]$$

Manipulando a expressão anterior, decorre que $\alpha f(x_1) + (1 - \alpha)f(x_2) \geq \alpha \nabla^T f(x)x_1 - \alpha \nabla^T f(x)x + \nabla^T f(x)x_2 - \alpha \nabla^T f(x)x_2 + \alpha \nabla^T f(x)x - \nabla^T f(x)x \implies \alpha f(x_1) + (1 - \alpha)f(x_2) \geq \alpha \nabla^T f(x)x_1 - \alpha \nabla^T f(x)x + \nabla^T f(x)x_2 - \alpha \nabla^T f(x)x_2 + \alpha \nabla^T f(x)x - \nabla^T f(x)x \implies (1 - \alpha)f(x_2) + \alpha f(x_1) \geq f(x) + \nabla^T f(x)[[\alpha x_1 + x_2(1 - \alpha)] - x] \implies (1 - \alpha)f(x_2) + \alpha f(x_1) \geq f(x) + \nabla^T f(x)([x] - x)$, lembre que $x = \alpha x_1 + x_2(1 - \alpha) \implies (1 - \alpha)f(x_2) + \alpha f(x_1) \geq f(x) + \nabla^T f(x)(0) \implies (1 - \alpha)f(x_2) + \alpha f(x_1) \geq f(x) \implies f(x) \leq (1 - \alpha)f(x_2) + \alpha f(x_1) \implies f$ é convexa. \square

Portanto, há aqui uma maneira alternativa de definir a convexidade de uma função f com domínio convexo Ω . Além disso, tal resultado afirma que a aproximação linear baseada na derivada local subestima a o valor da função no ponto, isto é, dados $x, y \in \Omega$, tem-se que $f(y) \geq f(x) + \nabla^T f(x)(y - x)$. Note que $f(x) + \nabla^T f(x)(y - x)$ é o polinômio de Taylor de ordem 1.

Teorema 1.7. *Seja $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, onde f é convexa em um conjunto convexo Ω . Logo, o conjunto Γ de pontos de mínimo é convexo, e qualquer ponto de mínimo de f é mínimo global.*

Demonstração. Considere $c_0 \in \mathbb{R}$ o valor mínimo global de f . Ademais, tome $\Gamma = \{x \in \Omega; f(x) \leq c_0\}$, perceba que Γ é o conjunto dos mínimos da f . Pelo **Lema 1.3.4** sabe-se que Γ é convexo. Agora, assuma que x^* é ponto de mínimo local, então $x^* \in \Gamma \subset \Omega$ e $f(x^*) \leq c_0$. Ora, suponha que exista $y \in \Omega$ tal que $f(y) < f(x^*) \leq c_0$. Como $f(y) < c_0$, então é fato que $y \in \Gamma$. Seja $0 < \alpha < 1$, portanto $f(y) < f(x^*) \implies \alpha f(y) < \alpha f(x^*) \implies \alpha f(y) + f(x^*) < \alpha f(x^*) + f(x^*) \implies \alpha f(y) + f(x^*) - \alpha f(x^*) < f(x^*) \implies \alpha f(y) + (1 - \alpha)f(x^*) < f(x^*)$ (lembre que f é convexa) $\implies f(\alpha y + (1 - \alpha)x^*) \leq \alpha f(y) + (1 - \alpha)f(x^*) < f(x^*) \implies f(\alpha y + (1 - \alpha)x^*) < f(x^*)$. O que contradiz o fato de x^* ser o mínimo local do segmento $\alpha y + (1 - \alpha)x^*$. Pelo fato de Γ ser convexo, note que o segmento $\alpha y_1 + (1 - \alpha)x^*$ sempre estará presente para qualquer $y_1 \in \Gamma$. Decorre que ao supor x^* como mínimo local em relação à linha $\alpha y_1 + (1 - \alpha)x^*$, temos que necessariamente x^* será mínimo global devido às condições acima. \square

Teorema 1.8. *Seja $f \in C^1$ em um conjunto $\Omega \subset \mathbb{R}^n$. Se existe $x^* \in \Omega$ tal que $\forall y \in \Omega$, $\nabla^T f(x^*)(y - x^*) \geq 0$, então x^* é mínimo global da f .*

Demonstração. Suponha que existe x^* tal que $\nabla^T f(x^*)(y - x^*) \geq 0, \forall y \in \Omega$. $\nabla^T f(x^*)(y - x^*) \geq 0 \implies f(x^*) + \nabla^T f(x^*)(y - x^*) \geq f(x^*) \implies f(y) \geq f(x^*) + \nabla^T f(x^*)(y - x^*) \geq f(x^*) \implies f(y) \geq f(x^*), \forall y \in \Omega \implies x^*$ é mínimo global em Ω . \square

Teorema 1.9. *Seja $f : \Omega \rightarrow \mathbb{R}$, onde Ω é convexo. Assim, x^* é minimizador global da f se, e somente se, $\nabla f(x^*) = 0$.*

Demonstração. As condições necessárias de primeira ordem satisfazem a ida, mesmo que a função não seja convexa (**Teorema 1.2**). Suponha que $\nabla f(x^*) = 0$. Ora, como f é convexa, note que pela **proposição 4**, tem-se que $f(y) \geq f(x) + \nabla^T f(x)(y - x), \forall x, y \in \Omega$. Em particular, imponha $x = x^*$, logo $f(y) \geq \nabla^T f(x^*)(y - x^*) = f(x^*) + 0^T(y - x^*) = f(x^*)$. Visto que y é um elemento genérico de Ω , conclui-se que x^* é mínimo global da f . \square

Teorema 1.10. *Considere uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}, f \in C^2$. Tem-se que f é convexa se, e só se, $\nabla^2 f(x) \geq 0$.*

Demonstração. Primeiramente, a volta é analisada. Pela fórmula de Taylor de ordem 1, existe $0 \leq \alpha \leq 1$ tal que $f(x_2) = f(x_1) + \nabla^T f(x_1)(x_2 - x_1) + \frac{1}{2}(x_1 - x_2)^T \nabla^2 f(\alpha x_1 + (1 - \alpha)x_2)(x_1 - x_2) \geq f(x_1) + \nabla^T f(x_1)(x_2 - x_1) + 0 = f(x_1) + \nabla^T f(x_1)(x_2 - x_1)$. Assim, $f(x_2) \geq f(x_1) + \nabla^T f(x_1)(x_2 - x_1)$. Consequentemente, pelo **Teorema 1.6**, f é convexa em \mathbb{R}^n . Pensando agora na ida, defina $g : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que $g(y) := f(y) - \nabla^T f(x)(y - x)$, onde $x \in \mathbb{R}^n$. Claramente, g é convexa, pois $-\nabla^T f(x)(y - x)$ é convexa (intuitivamente lembre de uma função afim, e de que f é uma função convexa), além disso, o **Lema 1.3.1** informa que qualquer combinação linear de convexas é convexa. Logo, $\nabla g(y) = \nabla f(y) - \nabla f(x) \implies \nabla^2 g(y) = \nabla^2 f(y), \forall y \in \mathbb{R}^n$. Note que $\nabla g(y) = \nabla f(y) - \nabla f(x) \implies \nabla g(x) = 0$. Portanto, pela **proposição 5** x é minimizador global. Logo, pelas condições necessárias de segunda ordem, decorre que $\nabla^2 g(x) = \nabla^2 f(x) \geq 0$. Como $x \in \mathbb{R}^n$ é arbitrário, segue que $\nabla^2 f(x) \geq 0, \forall x \in \mathbb{R}^n$, isto é, a matriz hessiana da f é semidefinida positiva. \square

1.4 Algoritmo Geral de Busca Linear

Embora o uso da convexidade seja uma característica interessante na busca por minimizadores, é notável que nem sempre uma função será localmente convexa em uma vizinhança da solução para o problema de minimização. Portanto, embora o desejo seja exibir explicitamente os extremantes de uma dada função, o objetivo desta seção visa apresentar um algoritmo que forneça candidatos a extremantes, isto é, pontos que satisfaçam as condições de primeira ordem (nulidade do gradiente da função no ponto) já comentadas acima. Ademais, exige-se que a função seja de classe C^1 . Assim sendo, considere o problema de minimização sem restrições, onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $f \in C^1$:

$$\begin{cases} \text{Minimizar } f(x) \\ x \in \mathbb{R}^n \end{cases}$$

Tal problema é o enfoque, entretanto, antes do algoritmo principal ser apresentado, certos resultados necessitam ser inicialmente expostos.

Lema 1.4.1. (*Direção de descida*) *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}, f \in C^1, x \in \mathbb{R}^n$ tal que $\nabla f(x) \neq 0, d \in \mathbb{R}^n$ e $\nabla^T f(x)d < 0$. Então existe $\bar{\alpha}$ tal que $f(x + \alpha d) < f(x), \forall \alpha \in (0, \bar{\alpha})$.*

Demonstração. Defina $\phi(\alpha) := f(x+\alpha d)$. Logo, $\phi'(\alpha) = \nabla f(x+\alpha d)(x+\alpha d)' = \nabla f(x+\alpha d)d$. Como $\alpha \neq 0$, tem-se que $\lim_{\alpha \rightarrow 0} \frac{\phi(\alpha) - \phi(0)}{\alpha - 0} = \phi'(0) = \nabla^T f(x)d < 0$. Consequentemente, deduz-se que é possível escolher um $\bar{\alpha} > 0$ suficientemente pequeno tal que para $\alpha \in (0, \bar{\alpha})$ tem-se que $\frac{\phi(\alpha) - \phi(0)}{\alpha - 0} < 0$. Portanto, existe um $\bar{\alpha} > 0$ onde $\forall \alpha \in (0, \bar{\alpha}) : \frac{\phi(\alpha) - \phi(0)}{\alpha - 0} < 0 \implies \phi(\alpha) < \phi(0) \implies f(x + \alpha d) < f(x)$. \square

Teorema 1.11. (Condição de Armijo) Sejam $x, d \in \mathbb{R}^n$ tal que $\nabla^T f(x)d < 0$ e $\alpha \in (0, 1)$. Existe $\epsilon > 0$ tal que $\forall t \in (0, \epsilon]$, tem-se que:

$$f(x + td) \leq f(x) + \alpha t \nabla^T f(x)d$$

Demonstração. Set $t > 0$, $\lim_{t \rightarrow 0^+} \left[\frac{f(x + td) - f(x)}{t} - \alpha \nabla^T f(x)d \right] = \lim_{t \rightarrow 0^+} \left[\frac{f(x + td) - f(x)}{t} \right] - \alpha \nabla^T f(x)d = \nabla^T f(x)d - \alpha \nabla^T f(x)d = (1 - \alpha) \nabla^T f(x)d < 0$, pois $0 < \alpha < 1 \implies 1 - \alpha > 0$ e $\nabla^T f(x)d < 0$. Assim, tem-se que para um t suficientemente pequeno ($t \rightarrow 0^+$) obteve-se $\frac{f(x + td) - f(x)}{t} - \alpha \nabla^T f(x)d < 0$. Portanto, é possível escolher um $\epsilon > 0$ suficientemente pequeno tal que $\forall t \in (0, \epsilon]$, $\frac{f(x + td) - f(x)}{t} - \alpha \nabla^T f(x)d < 0$. Isto é, $\frac{f(x + td) - f(x)}{t} - \alpha \nabla^T f(x)d < 0 \implies \frac{f(x + td) - f(x)}{t} < \alpha \nabla^T f(x)d \implies f(x + td) < f(x) + t\alpha \nabla^T f(x)d \implies f(x + td) \leq f(x) + t\alpha \nabla^T f(x)d$. \square

O algoritmo a seguir satisfaz os objetivos desejados no início da seção.

1.4.1 Algoritmo de descida com backtracking

Seja $x_k \in \mathbb{R}^n$, $\alpha \in (0, 1)$, $\beta > 0$, $\theta \in (0, 1)$.

Se $\nabla f(x_k) \neq 0$, os passos para determinar x_{k+1} são:

Passo 1: Escolher $d_k \in \mathbb{R}^n$, tal que:

- (i) $\|d_k\| \geq \beta \|\nabla f(x_k)\|$
- (ii) $\nabla^T f(x_k)d_k \leq -\theta \|\nabla f(x_k)\| \|d_k\|$

Passo 2(Busca Linear):

- (i) $t = 1$
- (ii) Se $f(x_k + td_k) < f(x_k) + \alpha t \nabla^T f(x_k)d_k$, ir a (iv)
- (iii) Escolher $\bar{t} \in [0.1t, 0.9t]$. Fazer $t = \bar{t}$ e ir a (ii)
- (iv) Fazer $t_k = t$ e $x_{k+1} = x_k + t_k d_k$

Lema 1.4.2. O algoritmo de descida com backtracking está bem definido. (É possível completar a busca linear com um número finito de tentativas para t).

Demonstração. Suponha por absurdo que não seja possível realizar a busca linear para um número finito de tentativas para t em relação a um ponto x_k . Logo, forma-se a seguinte sequência

$(t_i)_{i \in \mathbb{N}}$ e $t_i \in [(0.1)^i, (0.9)^i]$. Se $n \rightarrow \infty$, é nítido que $(0.1)^n \rightarrow 0$ e $(0.9)^n \rightarrow 0$ e, consequentemente, $t_{n_1} \rightarrow 0$. Portanto, para um $n_1 \in \mathbb{N}$, segue que:

$$\begin{aligned}
f(x_k + t_{n_1}d_k) > f(x_k) + \alpha t_{n_1} \nabla^T f(x_k) d_k &\implies f(x_k + t_{n_1}d_k) - f(x_k) > \alpha t_{n_1} \nabla^T f(x_k) d_k \\
&\implies \frac{f(x_k + t_{n_1}d_k) - f(x_k)}{t_{n_1}} > \frac{\alpha t_{n_1} \nabla^T f(x_k) d_k}{t_{n_1}} \\
&\implies \frac{f(x_k + t_{n_1}d_k) - f(x_k)}{t_{n_1}} > \alpha \nabla^T f(x_k) d_k \\
&\implies \nabla^T f(x_k) d_k > \alpha \nabla^T f(x_k) d_k (t_{n_1} \rightarrow 0) \\
&\implies (1 - \alpha) \nabla^T f(x_k) d_k > 0, (1 - \alpha) > 0 \\
&\implies \nabla^T f(x_k) d_k > 0.
\end{aligned}$$

O que contradiz o critério inicial $\nabla^T f(x_k) d_k \leq -\theta \|\nabla f(x_k)\| \|d_k\|$.

□

Teorema 1.12. (Convergência Global) *O algoritmo de descida com backtracking para em algum valor k tal que $\nabla f(x^k) = 0$, ou gera uma sequência infinita $\{x_k\}$ tal que qualquer ponto de acumulação dela é um ponto estacionário de f .*

Demonstração. Por uma questão de praticidade, a demonstração será dividida em passos.

I) Denote $s_k = x_{k+1} - x_k = td_k, \forall k \in N$, sendo N um conjunto de índices. Considere $K_1 \subset_{\infty} N$, tal que $\lim_{k \in K_1} x_k = x_*$. Sendo que o símbolo \subset_{∞} denota um subconjunto infinito de N .

II) São analisados dois casos:

a) $\lim_{k \in K_1} \|s_k\| = 0$ (Para um k suficientemente grande, temos que o tamanho do passo é nulo).

b) Existe um conjunto $K_2 \subset_{\infty} K_1$ e $\epsilon > 0$ tal que $\|s_k\| \geq \epsilon, \forall k \in K_2$ (Existe uma sequência de pontos com índices em K_1 tal que a distância entre eles formará uma sequência cuja distância será maior ou igual a um ϵ dado).

Suponha a validade do caso a).

$a_1)$ Suponha uma sequência de pontos com índices em um conjunto $K_3 \subset_{\infty} K_1$ tal que para todo $k \in K_3$ tem-se que $s_k = d_k$, isto é, $t = 1$ satisfaz as condições de descida $\forall k \in K_3$, logo $\|\nabla f(x^*)\| = \lim_{k \in K_3} \|\nabla f(x_k)\|, \lim_{k \in K_1} x_k = x_* \leq \lim_{k \in K_3} \frac{\|d_k\|}{\beta}, \frac{\|d_k\|}{\beta} \geq \|f(x_k)\| = \frac{1}{\beta} \lim_{k \in K_3} \|s_k\| = 0 \implies \|\nabla f(x_*)\| = 0 \implies \nabla f(x_*) = 0$. Portanto, a sequência de pontos com índices no conjunto K_3 , que por hipótese iria convergir para x_* , mostra que diante de sua caracterização específica, revela que x_* é ponto estacionário.

$a_2)$ Entretanto, é possível que para uma certa sequência de pontos com índices em K_1 , a condição de Armijo falhe para $t = 1$. Logo, se para todo $k \in K_1, k \geq k_0$ tem-se que $t <$

1. Consequentemente, existe um \bar{s}_k múltiplo de s_k tal que Armijo falha para $t = 1$, ou seja, $f(x_k + \bar{s}_k) > f(x_k) + \alpha \nabla^T f(x_k) \bar{s}_k$. Ora, note que $d_k \in \mathbb{R}^n$ foi escolhido inicialmente. Além disso, sendo $\bar{s}_k = \bar{t} d_k$, então $\bar{t} \in [1(0.1)^n, 1(0.9)^n] \xrightarrow{n \geq 1} \bar{t} \leq 1 \implies \|d_k\| \bar{t} \leq \|d_k\| t \implies \|d_k \bar{t}\| \leq \|d_k t\| \implies \|\bar{s}_k\| \leq \|s_k\| \implies \lim_{k \in K_1} \|\bar{s}_k\| \leq \lim_{k \in K_1} \|s_k\| = 0 \implies \lim_{k \in K_1} \|\bar{s}_k\| \leq 0 \xrightarrow{\|\bar{s}_k\| \geq 0} \lim_{k \in K_1} \|\bar{s}_k\| = 0$.

Lembre que $\nabla^T f(x_k) \bar{s}_k \leq -\theta \|\nabla f(x_k)\| \|\bar{s}_k\|$ (i), visto que é uma das condições de escolha para $d_k \in \mathbb{R}^n$. Considere agora que v é ponto de acumulação de $\frac{\bar{s}_k}{\|\bar{s}_k\|}$, isto é, existe uma sequência de pontos com índices em um subconjunto $K_4 \subset K_1$ tal que $\lim_{k \in K_4} \frac{\bar{s}_k}{\|\bar{s}_k\|} = v$. Onde claramente $\|v\| = 1$. Logo: $\nabla^T f(x_*) v = \lim_{k \in K_4} \nabla^T f(x_k) v = \lim_{k \in K_4} \nabla^T f(x_k) \frac{\bar{s}_k}{\|\bar{s}_k\|}$ (ii).

Assim, de (i): $\nabla^T f(x_k) \bar{s}_k \leq -\theta \|\nabla f(x_k)\| \|\bar{s}_k\| \implies \nabla^T f(x_k) \frac{\bar{s}_k}{\|\bar{s}_k\|} \leq -\theta \|\nabla f(x_k)\| \implies \lim_{k \in K_4} \nabla^T f(x_k) \frac{\bar{s}_k}{\|\bar{s}_k\|} \leq \lim_{k \in K_4} -\theta \|\nabla f(x_k)\| \xrightarrow{(ii)} \nabla^T f(x_*) v \leq -\theta \lim_{k \in K_4} \|\nabla f(x_k)\|$ (iii).

Seja $\phi : [0, 1] \rightarrow \mathbb{R}$, $\phi(\epsilon_k) = f(x_k + \epsilon_k \bar{s}_k) \implies \phi'(\epsilon_k) = \nabla f(x_k + \epsilon_k \bar{s}_k)^T \bar{s}_k$, $k \in K_4$. Devido ao fato $f \in C^1$, tem-se que ϕ é diferenciável em $(0, 1)$ e contínua em $[0, 1]$, então pelo Teorema do Valor Médio (LIMA, 2008), existe um $\epsilon \in (0, 1)$ tal que: $\phi'(\epsilon) = \frac{\phi(1) - \phi(0)}{1 - 0} = \phi(1) - \phi(0) \implies \nabla f(x_k + \epsilon \bar{s}_k)^T \bar{s}_k = f(x_k + \bar{s}_k) - f(x_k)$.

Pelo fracasso de Armijo exibida no início do item a_2), ou seja, $\forall k \in k_1, k \geq k_0$, tem-se que $f(x_k + \bar{s}_k) - f(x_k) > \alpha \nabla^T f(x_k) \bar{s}_k$, logo $f(x_k + \epsilon \bar{s}_k)^T \bar{s}_k = f(x_k + \bar{s}_k) - f(x_k) > \alpha \nabla^T f(x_k) \bar{s}_k \implies f(x_k + \epsilon \bar{s}_k)^T \bar{s}_k > \alpha \nabla^T f(x_k) \bar{s}_k, \forall k \in K_4 \implies \frac{f(x_k + \epsilon \bar{s}_k)^T \bar{s}_k}{\|\bar{s}_k\|} > \frac{\alpha \nabla^T f(x_k) \bar{s}_k}{\|\bar{s}_k\|} \implies \lim_{k \in K_4} \frac{f(x_k + \epsilon \bar{s}_k)^T \bar{s}_k}{\|\bar{s}_k\|} \geq \lim_{k \in K_4} \frac{\alpha \nabla^T f(x_k) \bar{s}_k}{\|\bar{s}_k\|} \implies \nabla^T f(x_*) v \geq \alpha \nabla^T f(x_*) v$, pois $\lim_{k \in K_1} x_k = x_*$.

Portanto, $\nabla^T f(x_*) v \geq \alpha \nabla^T f(x_*) v \implies (1 - \alpha) \nabla^T f(x_*) v \geq 0 \implies \nabla^T f(x_*) v \geq 0$. De (iii), temos que $\nabla^T f(x_*) v \leq -\theta \lim_{k \in K_4} \|\nabla f(x_k)\| \implies \nabla^T f(x_*) v \leq 0$. Assim, $\nabla^T f(x_*) v \geq 0$ e $\nabla^T f(x_*) v \leq 0 \implies \nabla^T f(x_*) v = 0$.

Dessa maneira, suponha que $\nabla f(x_*) \neq 0$, logo para um $k \in K_4$ suficientemente grande tem-se por (iii) que: $\nabla^T f(x_*) v \leq -\theta \lim_{k \in K_4} \|\nabla f(x_k)\| < 0 \implies \nabla^T f(x_*) v < 0$ e $\nabla^T f(x_*) v = 0$. O que é absurdo, logo $\nabla f(x_*) = 0$.

Suponha agora a validade do caso b). Logo, temos que $\|s_k\| \geq \epsilon, \forall k \in K_2$. Pelo sucesso na condição de Armijo, temos que $f(x_k + s_k) \leq f(x_k) + \alpha [\nabla^T f(x_k) s_k] \leq f(x_k) - \alpha \theta \|\nabla f(x_k)\| \|s_k\| \leq f(x_k) - \alpha \theta \|\nabla f(x_k)\| \epsilon$. Assim, $f(x_k + s_k) \leq f(x_k) - \alpha \theta \epsilon \|\nabla f(x_k)\| \implies \frac{f(x_k) - f(x_{k+1})}{\alpha \theta \epsilon} \geq \|\nabla f(x_k)\| \implies \lim_{k \in K_2} \frac{f(x_k) - f(x_{k+1})}{\alpha \theta \epsilon} \geq \lim_{k \in K_2} \|\nabla f(x_k)\|$.

Como $\lim_{k \in K_1} x_k = x_*$ e $f \in C^1$: $\frac{f(x_*) - f(x_*)}{\alpha \theta \epsilon} \geq \|\nabla f(x_*)\| \implies \|\nabla f(x_*)\| \leq 0 \implies \|\nabla f(x_*)\| = 0 \implies \nabla f(x_*) = 0$. É importante ressaltar que o algoritmo está bem definido,

isto é, temos que para algum $t \in (0, 1]$ a condição de Armijo funcionará para uma sequência de pontos com índices em K_1 , pois caso não, chegaríamos em uma contradição relacionada à boa definição do algoritmo explicitada no lema anterior.

□

1.5 Ordem ou razão de convergência

Se $\lim_{k \rightarrow \infty} x_k = x^* \implies \lim_{k \rightarrow \infty} \|x_k - x^*\| = 0$, onde $\|x_k - x^*\|$ é o erro cometido na aproximação de x_k . Ademais, quanto mais rápido o erro se aproximar de zero, melhor. Defina $e_{k+1} = \|x_{k+1} - x^*\|$ e $e_k = \|x_k - x^*\|$.

Definição 10. (Convergência Linear) Se $e_{k+1} \leq r e_k$ para algum $r \in (0, 1)$, diz-se que $\{x_k\}$ converge com ordem linear e taxa não superior a r .

Definição 11. (Convergência superlinear) Se $\lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k} = 0$, então a convergência é dita superlinear.

Definição 12. Se $e_{k+1} \leq a(e_k)^p$, $a > 0$, $p > 1$, diz-se que a sequência $\{x_k\}$ converge a x^* com ordem não inferior a p . Se $p = 2$, então a convergência é dita quadrática.

Lema 1.5.1. Se a convergência é superlinear, então a convergência é linear.

Demonstração. Com efeito, se $\lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k} = 0 \implies \forall \epsilon > 0, \exists k_0 \in \mathbb{N}; k > k_0 \implies \left\| \frac{e_{k+1}}{e_k} - 0 \right\| < \epsilon$. Ora, tome $\epsilon = r$ para $r \in (0, 1)$. Como $e_{k+1} \geq 0$, $e_k \geq 0$, decorre $\frac{e_{k+1}}{e_k} < r \implies e_{k+1} \leq r e_k$. □

Lema 1.5.2. A convergência quadrática implica em superlinear.

Demonstração. Temos $e_{k+1} \leq a(e_k)^2$, $a > 0$. Precisamos mostrar que $\lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k} = 0$. De fato, $e_{k+1} \leq a(e_k)^2 \implies \frac{e_{k+1}}{e_k} \leq a e_k \implies \lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k} \leq \lim_{k \rightarrow \infty} a e_k$. Já que $k \rightarrow \infty \implies x_k \rightarrow x^* \implies \lim_{k \rightarrow \infty} e_k = \lim_{k \rightarrow \infty} \|x_k - x^*\| = 0$. Assim, $\lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k} \leq 0$. Como $e_{k+1} = \|x_{k+1} - x^*\| > 0$ e $e_k = \|x_k - x^*\| > 0$, segue que $\frac{e_{k+1}}{e_k} > 0 \implies \lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k} \geq 0$. Logo, $\lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k} = 0$. Por definição, temos a convergência superlinear. □

2. Métodos Clássicos de Descida

Diante da perspectiva do algoritmo de descida com backtracking, vimos que este possui propriedades de convergência global, isto é, selecionando os parâmetros do problema nos intervalos explicitados, bem como escolhendo uma direção d_k que satisfaça dois critérios iniciais, temos que a busca linear retornará um ponto estacionário, ou irá gerar uma sequência infinita de pontos, tal que qualquer ponto de acumulação dessa sequência é um ponto estacionário. Dessa forma, a maneira de se escolher a direção d_k é o que constitui a diferença entre os métodos analisados neste capítulo, sendo estes, o Método do Gradiente, o Método de Newton, o DFP, o BFGS e o Método dos Gradientes Conjugados.

2.1 Método do Gradiente

Caracteriza-se pela escolha de d_k na direção de $-\nabla f(x_k)$. Perceba que se $d_k = -\nabla f(x_k)$ tem-se satisfeitas as duas condições iniciais na escolha da direção do algoritmo de descida com backtracking. Com efeito, veja que $\|d_k\| = \|\nabla f(x_k)\| \geq \beta \|\nabla f(x_k)\| \iff \beta \leq 1$, o que vale, pois exige-se apenas $\beta > 0$. Além disso, $\nabla^T f(x_k)(-\nabla f(x_k)) \leq -\theta \|\nabla f(x_k)\| \|\nabla f(x_k)\| \iff -\|f(x_k)\|^2 \leq -\theta \|\nabla f(x_k)\|^2 \iff \theta \leq 1$, o que é válido, pois pelas condições $\theta \in (0, 1)$.

2.1.1 Algoritmo(Método do Gradiente)

Se $x_k \in \mathbb{R}^n$ é tal que $\nabla f(x_k) \neq 0$, os passos para determinar x_{k+1} são:

Passo 1: Calcular $d_k = -\nabla f(x_k)$

Passo 2: Realize uma busca linear exata, isto é, determine α_k sujeita a $\alpha \geq 0$ que minimize a expressão $f(x_k + \alpha d_k)$. Ou realize uma busca linear inexata utilizando as condições de Armijo.

Passo 3: Fazer $x_{k+1} = x_k + \alpha_k d_k$

Lema 2.1.1. *Seja $A \in M_n(\mathbb{R})$. Se $A = A^t$, tem-se os seguintes resultados:*

- i) *A possui apenas autovalores reais.*
- ii) *A é diagonalizável.*
- iii) *Autovetores de A associados a distintos autovalores são ortogonais.*

Demonstração. Encontra-se em (HOFFMAN; KUNZE, 1971) □

Lema 2.1.2. *Seja $A \in M_n(\mathbb{R})$. A é definida positiva. Então temos que as seguintes condições são equivalentes:*

- i) *Os autovalores de A são estritamente positivos.*
- ii) *Os menores principais de A são positivos.*

Demonstração. Encontra-se em (HOFFMAN; KUNZE, 1971). □

Lema 2.1.3. *Seja $A \in M_n(\mathbb{R})$. Se A é definida positiva, então A é invertível.*

Demonstração. Encontra-se em (HOFFMAN; KUNZE, 1971). □

Lema 2.1.4. *Seja $A \in M_n(\mathbb{R})$. Se A é invertível e λ é autovalor de A, então $\frac{1}{\lambda}$ é autovalor de A^{-1} .*

Demonstração. Encontra-se em (HOFFMAN; KUNZE, 1971). □

Lema 2.1.5. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) = \frac{1}{2}x^t Qx - x^t b + c$, sendo Q simétrica e definida positiva. Se x^* é minimizador local, então x^* é minimizador global. Além disso, $Qx^* = b$.*

Demonstração. Ora, sabe-se que $\nabla f^2(x) = \frac{Q + Q^t}{2}$. Como $Q = Q^t$ e $x^t Qx > 0, \forall x \in \mathbb{R}^n, x \neq 0 \implies \nabla^2 f(x) = Q > 0$. Pelo Teorema 1.10, do fato de que $f \in C^2$ e $\nabla^2 f(x) > 0$, segue que f é convexa. Da hipótese de que x^* é mínimo local, do Teorema 1.7 garantimos que x^* é mínimo global. Ademais, das condições necessárias de primeira ordem, se x^* é mínimo global da f , então $\nabla f(x^*) = 0 \implies \nabla f(x^*) = \left(\frac{Q^t + Q}{2} \right) x^* - b = 0 \implies Qx^* - b = 0 \implies Qx^* = b$. □

Ainda no cenário do lema anterior, introduza a função $E : \mathbb{R}^n \rightarrow \mathbb{R}$, $E(x) = \frac{1}{2}(x - x^*)^T Q(x - x^*)$. Após algumas operações, obtemos $E(x) = f(x) + \frac{1}{2}x^{*T} Qx^*$. Note que o termo $\frac{1}{2}x^{*T} Qx^*$ é uma constante real. Portanto, as funções E e f diferem apenas por uma constante. Logo, é importante ressaltar este fato, pois será conveniente minimizar E ao invés da f diretamente. Perceba também que $\nabla E(x) = Qx - b = \nabla f(x)$ e denote $g_k = \nabla f(x_k) = Qx_k - b$.

O interessante do caso quadrático é a possibilidade de se exibir α_k de maneira explícita. Ou seja, na k -ésima iteração, sempre podemos encontrar α_k que minimiza a expressão $f(x_k -$

$\alpha_k g_k$). Com efeito, $\nabla^T f(x_k - \alpha g_k) = 0$, pela condições necessárias de primeira ordem. Assim, $\nabla^T f(x_k - \alpha g_k)(-g_k) = 0$. Utilizando as definições anteriores, a simetria de Q (que permitirá afirmar $x_k^T Q g_k = g_k^T Q x_k$) e expandindo as contas, temos que $\alpha = \frac{g_k^T g_k}{g_k^T Q g_k}$. Dessa forma, o método do gradiente para o caso quadrático possui a seguinte estratégia para a determinação do passo seguinte: $x_{k+1} = x_k - \alpha g_k = x_k - \left(\frac{g_k^T g_k}{g_k^T Q g_k} \right) g_k$, onde $g_k = \nabla f(x_k) = Q x_k - b$.

Lema 2.1.6. *O processo iterativo satisfaz:*

$$E(x_{k+1}) = \left[1 - \frac{(g_k^T g_k)^2}{(g_k^T Q g_k)(g_k^T Q^{-1} g_k)} \right] E(x_k)$$

Demonstração. Defina $y_k = x_k - x^*$. Lembre que $E(x) = f(x) + \frac{1}{2} x^{*T} Q x^*$. Ora, estude o quociente $\frac{E(x_k) - E(x_{k+1})}{E(x_k)}$. Obtemos, $\frac{E(x_k) - E(x_{k+1})}{E(x_k)} = \frac{2\alpha_k g_k^T (Q y_k) - \alpha_k^2 g_k^T Q g_k}{y_k^T Q y_k} (*)$. Como $y_k = x_k - x^* \implies Q y_k = Q(x_k - x^*) = Q x_k - Q x^* = Q x_k - b = g_k$, pois recorde que $Q x^* = b$ conforme vimos acima. Ademais, por meio dessa última equação obtida, temos que $y_k^T = g_k^T Q^{-1}$, visto que Q é invertível pelo lema 2.1.2 acima. Por fim, $\alpha_k = \frac{g_k^T g_k}{g_k^T Q g_k}$ segundo a busca linear exata. Logo, aplicando todas estas observações em $(*)$, segue a tese. \square

Lema 2.1.7. *(Lema de Kantarovich) Considere Q positiva definida de ordem n . Para qualquer $x \in \mathbb{R}^n$:*

$$\frac{x^T x}{(x^T Q x)(x^T Q^{-1} x)} \geq \frac{4aA}{(a + A)^2}$$

Onde A e a são, respectivamente, o maior e o menor autovalor da matriz Q .

Demonstração. Seja $Q > 0$ e de ordem n . Desenvolvendo seu polinômio característico, percebe-se que o grau do polinômio é n , logo temos n autovalores associados à Q . Além disso, todos os autovalores de Q são estritamente positivos graças ao Lema 2.1.2. Assim, considere tais autovalores como $0 < a = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n = A$. Ora, seja β a base de \mathbb{R}^n formada pelos n distintos autovetores de Q . Reescreva-a conforme à base β , isto é, $[Q]_\beta = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. Ademais, pelo Lema 2.1.4, temos que $[Q^{-1}]_\beta = \text{diag}\left(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_n}\right)$. Consequentemente, temos a seguinte expressão, onde $x \in \mathbb{R}^n$:

$$\frac{(x^T x)^2}{(x^T Q x)(x^T Q^{-1} x)} = \frac{\left(\sum_{i=1}^n x_i^2 \right)^2}{\left(\sum_{i=1}^n \lambda_i x_i^2 \right) \left(\sum_{i=1}^n \frac{x_i^2}{\lambda_i} \right)}$$

Desejamos exprimir a igualdade acima como um quociente de combinações convexas envolvendo os autovalores de Q de maneira a facilitar a visualização de um limitante inferior

para esta expressão. Defina $\epsilon_i := \frac{x_i^2}{\sum_{i=1}^n x_i^2}$ e $\frac{\phi(\epsilon)}{\psi(\epsilon)} := \frac{\frac{1}{\left(\sum_{i=1}^n \epsilon_i \lambda_i\right)}}{\left(\sum_{i=1}^n \frac{\epsilon_i}{\lambda_i}\right)}$. Ou seja, $\phi(\epsilon)$ é combinação

convexa de λ_i e $\psi(\epsilon)$ é combinação convexa de $1/\lambda_i$ com $i = 1, \dots, n$. Também é importante perceber que $\phi(\epsilon)$ é um ponto de uma curva $1/\lambda$ com $\lambda_1 \leq \lambda \leq \lambda_n$. Analogamente, $\psi(\epsilon)$ é um ponto situado abaixo dessa mesma curva $1/\lambda$.

$$\frac{\phi(\epsilon)}{\psi(\epsilon)} = \frac{\frac{1}{\left(\sum_{i=1}^n \epsilon_i \lambda_i\right)}}{\left(\sum_{i=1}^n \frac{\epsilon_i}{\lambda_i}\right)} = \frac{\left(\sum_{i=1}^n x_i^2\right)^2}{\left(\sum_{i=1}^n \lambda_i x_i^2\right) \left(\sum_{i=1}^n \frac{x_i^2}{\lambda_i}\right)} = \frac{(x^T x)^2}{(x^T Q x)(x^T Q^{-1} x)}$$

Ora, para um mesmo vetor ϵ temos que $\phi(\epsilon)$ e $\psi(\epsilon)$ estão numa mesma reta vertical. Assim, decorre que $\frac{\phi(\epsilon)}{\psi(\epsilon)}$ atinge valor mínimo para certo λ entre λ_1 e λ_n . Logo, $\lambda = \alpha_1 \lambda_1 + \alpha_n \lambda_n$

com $\alpha_1 + \alpha_n = 1$. Note que $\frac{\alpha_1}{\lambda_1} + \frac{\alpha_n}{\lambda_n} = \frac{\lambda_1 + \lambda_n - \lambda}{\lambda_1 \lambda_n}$.

Portanto, um valor limitante para o quociente $\frac{\phi(\epsilon)}{\psi(\epsilon)}$ é:

$$\frac{\phi(\epsilon)}{\psi(\epsilon)} \geq \lim_{\lambda_1 \leq \lambda \leq \lambda_n} \frac{1/\lambda}{\frac{\lambda_1 + \lambda_n - \lambda}{\lambda_1 \lambda_n}} \quad (*)$$

Ora, $\frac{1/\lambda}{\frac{\lambda_1 + \lambda_n - \lambda}{\lambda_1 \lambda_n}} = \frac{\lambda_1 \lambda_n}{-\lambda^2 + (\lambda_1 + \lambda_n)\lambda}$. Para encontrar o valor mínimo desta expres-

são, basta saber o máximo de $-\lambda^2 + (\lambda_1 + \lambda_n)\lambda$, cujo o valor é $\lambda = \frac{\lambda_1 + \lambda_n}{2}$. Substituindo este valor de λ em $(*)$, temos que:

$$\frac{(x^T x)^2}{(x^T Q x)(x^T Q^{-1} x)} = \frac{\phi(\epsilon)}{\psi(\epsilon)} \geq \frac{4aA}{(a+A)^2}$$

□

Teorema 2.1. (*Método do Gradiente - Caso Quadrático*) Para qualquer $x_0 \in \mathbb{R}^n$ o método do gradiente converge para um único ponto mínimo x^* de f . Além disso, se $E(x) = \frac{1}{2}(x - x^*)^T Q(x - x^*)$, temos que $E(x_{k+1}) \leq \left(\frac{A-a}{A+a}\right)^2 E(x_k)$.

Demonstração. Pelo Lema 2.1.6 e pelo Lema de Kantarovich, temos que:

$$E(x_{k+1}) = \left[1 - \frac{(g_k^T g_k)^2}{(g_k^T Q g_k)(g_k^T Q^{-1} g_k)} \right] E(x_k) \leq \left[1 - \frac{4aA}{(a+A)^2} \right] E(x_k) = \left[\frac{A-a}{A+a} \right]^2 E(x_k).$$

Como

$$E(x_k) = \frac{1}{2}(x - x^*)^T Q (x - x^*), \text{ se } E(x_k) \rightarrow 0 \text{ e } Q > 0, \text{ então } x_k \rightarrow x^*.$$

□

Algumas observações necessitam ser feitas. Perceba que as curvas de nível da forma quadrática são elipsóides n-dimensionais. A relevância do teorema anterior constitui-se do fato de que a razão de convergência do Método do Gradiente é diminuída quando tais curvas vão aumentando sua excentricidade, visto que os eixos das curvas de nível são as direções dos n autovetores mutuamente ortogonais de Q , onde o eixo correspondendo ao i -ésimo autovetor possui tamanho proporcional à $1/\lambda_i$. Note que se $A = a$, então a convergência do Método do Gradiente ocorre em um único passo. Porém, se $n - 1$ autovalores são iguais, restando apenas 1 autovalor distinto λ_n , onde $|\lambda_i - \lambda_n| \gg \epsilon$, para $\epsilon > 0$, isto é, quanto maior a distância entre tais autovalores, mais lenta será a convergência. Portanto, podemos dizer que anormalidade relativa à distância existente entre um único autovalor em relação aos demais autovalores de Q , destrói a eficiência do método.

Além disso, já que $E(x_{k+1}) \leq \left(\frac{A-a}{A+a} \right)^2 E(x_k)$, note que $\left[\frac{A-a}{A+a} \right]^2 \in [0, 1)$, pois $0 \leq a \leq A$. Assim, veja que temos convergência linear com taxa não superior a $\left[\frac{A-a}{A+a} \right]^2$. Com isso, demonstramos o seguinte teorema:

Teorema 2.2. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função quadrática com hessiana definida positiva. Seja x^* o minimizador global da f . Dado $x_0 \in \mathbb{R}^n$ arbitrário, tem-se que o algoritmo referente ao Método do Gradiente gera uma sequência $\{x_k\}$ com as propriedades abaixo:*

$$i) \lim_{k \rightarrow \infty} x_k = x^*$$

$$ii) \lim_{k \rightarrow \infty} f(x_k) = f(x^*) \text{ e } f(x_{k+1}) - f(x^*) \leq \left(\frac{A-a}{A+a} \right)^2 (f(x_k) - f(x^*)),$$

onde A e a são, respectivamente, o maior e o menor autovalor de Q .

Demonstração. Analise os lemas 2.1.5, 2.1.6, 2.1.7 e o Teorema 2.1.

□

2.1.2 O Caso Não-Quadrático

O teorema a seguir não garante convergência para estes casos, porém relata a ordem quando ela ocorre.

Teorema 2.3. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^2$. Seja x^* um minimizador local de f , tal que a matriz $\nabla^2 f(x^*)$ é definida positiva. Se o algoritmo referente ao Método do Gradiente está bem definido para todo $k \in \mathbb{N}$ e a sequência $\{x_k\}$ gerada por ele converge a x^* , então a sequência $\{f(x_k)\}$ converge linearmente a $f(x^*)$ com taxa não superior a $\left(\frac{A-a}{A+a}\right)^2$, onde A e a são o maior e o menor autovalor de $\nabla^2 f(x^*)$, respectivamente.*

Demonstração. Encontra-se em (LUENBERGER; YE, 2010). □

2.2 Método de Newton

2.2.1 O Caso Quadrático

Lema 2.2.1. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) = \frac{1}{2}x^T Gx + b^T x + c$ com $G > 0$. Dado $x^0 \in \mathbb{R}^n$ arbitrário, a direção $d \in \mathbb{R}^n$ dada por $d = -G^{-1}(Gx^0 + b)$ confirma que $x^* \equiv x^0 + d$ é minimizador global de f em \mathbb{R}^n .*

Demonstração. Primeiramente, como $G > 0$ e $G = \nabla^2 f(x)$, então f é convexa, logo se $\nabla f(x^*) = 0$, temos que x^* é mínimo global. Ora, $\nabla f(x^*) = \nabla f(x^0 + d) = G(x^0 + d) + b = Gx^0 + Gd + b = Gx^0 + G(-G^{-1}(Gx^0 + b)) + b = Gx^0 + (-I(Gx^0 + b)) + b = Gx^0 - Gx^0 - b + b = 0 \implies \nabla f(x^*) = 0 \implies x^* = x^0 + d$ é mínimo global da f . □

Note que $d = -G^{-1}(Gx^0 + b) = -G^{-1}(\nabla f(x_0)) = G^{-1}(\nabla f(x_0)) \iff Gd = -\nabla f(x_0)$. Assim, minimizar uma função quadrática $f(x) = \frac{1}{2}x^T Gx + b^T x + c$ com $G > 0$ é o mesmo que achar a solução d do sistema linear $Gd = -\nabla f(x_0)$, onde x_0 é um ponto inicialmente arbitrado e $G \in M_n(\mathbb{R})$, $G > 0$.

2.2.2 Caso Não-Quadrático

Caso a função seja não-quadrática e temos uma aproximação x_k da solução de

$$\begin{cases} \text{minimizar } f(x) \\ x \in \mathbb{R}^n \end{cases}$$

Onde a estratégia é utilizar a fórmula de Taylor em torno de x_k .

Assim, $f(x_k + d) = f(x_k) + \nabla^T f(x_k)d + \frac{1}{2}d^T \nabla^2 f(x_k)d + O(\|d\|^2) \implies q(d) = f(x_k + d) - O(\|d\|^2) = f(x_k) + \nabla^T f(x_k)d + \frac{1}{2}d^T \nabla^2 f(x_k)d$. Dessa forma, temos:

$$q(d) = f(x_k) + \nabla^T f(x_k)d + \frac{1}{2}d^T \nabla^2 f(x_k)d$$

Suponha que $\nabla^2 f(x_k) > 0$. Logo, podemos calcular o minimizador da quadrática $q(d)$ a partir de um valor $d_0 = 0$. Ora, sabemos a resposta deste problema graças à análise do caso quadrático. Dessa forma, a direção de descida é $d^* = (-\nabla^2 f(x_k))^{-1}(\nabla^2 f(x_k)d_0 + b) = (-\nabla^2 f(x_k))^{-1}(\nabla f(x_k))$.

Entretanto, $d_k = -(\nabla^2 f(x_k))^{-1}\nabla f(x_k)$ é sempre direção de descida? Não, pois pode ser que $\nabla^2 f(x_k)$ não seja definida positiva. Um exemplo é $f(x, y) = \frac{1}{2}x^2 - \frac{1}{2}y^2$, onde arbitrou-se que $x_0 = (0, 1)^T$. Veja que $\nabla f(x, y) = (x, -y) \implies \nabla f(0, 1) = (0, -1)$. Além disso, temos que $\nabla^2 f(x, y) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \implies \nabla^2 f(0, 1) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Conseqüentemente, $\nabla^2 f(x)$ não é definida positiva, pois seus autovalores não são todos estritamente positivos. Agora, compute a direção de Newton:

$$d_0 = -(\nabla^2 f(x_0))^{-1}(\nabla f(x_0)) = - \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

Conseqüentemente,

$$\nabla^T f(x_0)d_0 = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \end{bmatrix} = 1 > 0$$

Logo, d_0 não é direção de descida por causa do Lema 1.4.1.

Outro questionamento possível é se não bastaria escolher $\tilde{\lambda} = -d_0$ para obter direção de descida. Entretanto, aqui vai mais um contraexemplo:

$$f(x, y) = x^4 + xy + (1+y)^2, \text{ onde tomamos } x_0 = (0, 0)^T. \text{ Ora, } \nabla f(x, y) = \begin{bmatrix} 4x^3 + y \\ x + 2y + 2 \end{bmatrix}$$

e $\nabla^2 f(x, y) = \begin{bmatrix} 12x^2 & 1 \\ 1 & 2 \end{bmatrix} \implies \nabla^2 f(x_0) = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$.

Perceba que $\begin{bmatrix} -2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = -2 < 0 \implies \nabla^2 f(x, y)$ não é definida positiva. Ademais, temos que $d_0 = -(\nabla^2 f(x_0))^{-1}\nabla f(x_0) = \begin{bmatrix} -2 & 0 \end{bmatrix} \implies \nabla^T f(x_0)d_0 = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \begin{bmatrix} -2 & 0 \end{bmatrix} = 0$. Dessa maneira, é importante ressaltar que se d_k é direção de descida, a verificação do passo 1 no algoritmo exibido em 1.4.1 depende das características da função objetivo. No caso, uma rápida maneira de verificar tal situação é analisar se os autovalores da hessiana da f estão contidos em um intervalo $(a, b) \subset \mathbb{R}$, $a > 0$.

2.2.3 Algoritmo(Método de Newton)

Dados $\alpha \in (0, 1), \beta > 0, \theta \in (0, 1), x_k \in \mathbb{R}^n$. Se x_k é tal que $g(x_k) = \nabla f(x_k) \neq 0$, os passos a determinar são:

Passo 1: Tentar fatoração de Cholesky, $\nabla^2 f(x_k) = LDL^T$

Passo 2: Se houver sucesso no passo 1, obtenha d_k :

$$Lz = -g(x_k) \text{ e } DL^T d_k = z$$

Passo 3: Se o passo 1 fracassou, defina a matriz $B_k = \nabla^2 f(x_k) + \mu I$, $\mu > 0$ de maneira que $B_k > 0$. Obter a fatoração de Cholesky: $B_k = \overline{L}\overline{D}\overline{L}^T$ e calcular d_k resolvendo

$$\overline{L}z = g(x_k) \text{ e } \overline{D}\overline{L}^T d_k = z$$

Passo 4: Se $g(x_k)^t d_k > -\theta \|g(x_k)\| \|d_k\|$, fazer $\mu \leftarrow \max\{2\mu, 10\}$ e repetir o Passo 3 como se tivesse havido fracasso na fatoração de Cholesky.

Passo 5: $\|d_k\| < \beta \|g(x_k)\|$, corrigir

$$d_k \leftarrow \beta \frac{\|g(x_k)\|}{\|d_k\|} d_k$$

Passo 6: Obter t por backtracking(algoritmo 1.4.1) de modo a satisfazer:

$$f(x_k + td_k) \leq f(x_k) + \alpha t g(x_k)^t d_k,$$

definir $x_{k+1} = x_k + td_k$ e voltar para o passo 1 se $g(x_k) \neq 0$.

Note que tivemos de utilizar a fatoração de Cholesky como método de correção da hessiana da f visto que ela não é necessariamente positiva definida. Outro detalhe é que se $\nabla^2 f(x_k) > 0$, temos que as condições do Algoritmo Geral de Busca Linear são automaticamente satisfeitas, portanto poderíamos ir diretamente ao passo 5 nessa k -ésima iteração, bastando selecionar β e θ apropriados.

2.2.4 Convergência do Método de Newton

O método de Newton possui propriedades bastante desejáveis caso o $x_0 \in \mathbb{R}^n$ arbitrado comece suficientemente próximo da solução do problema.

Teorema 2.4. *Seja $f \in C^3$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Assuma que em um ponto mínimo local x^* temos que $\nabla^2 f(x^*) > 0$. Logo, se x_0 for suficientemente próximo de x^* ($\exists \epsilon > 0, x_0 \in B(x^*, \epsilon)$), então a sequência de pontos gerada pelo Método de Newton converge para x^* . A ordem de convergência do Método de Newton é pelo menos quadrática.*

Demonstração. Por questão de praticidade, denote $F(x_k) = \nabla^2 f(x_k)$. Ora, para $|x - x^*| < p$, temos que $|F^{-1}(x)| < \beta_1$ e $\frac{|\nabla^T f(x^*) - \nabla^T f(x) - F(x)(x^* - x)|}{|x - x^*|^2} \leq \beta_2 \implies |\nabla^T f(x^*) -$

$|\nabla^T f(x) - F(x)(x^* - x)| \leq |x - x^*|\beta_2$. Lembre que x_k é um ponto arbitrário do \mathbb{R}^n , portanto escolha-o de tal sorte que $\beta_1\beta_2|x_k - x^*| < 1$ e $|x_k - x^*| < p$. Dessa forma:

$$\begin{aligned} |x_{k+1} - x^*| &= |(x_k - F(x_k)^{-1}\nabla^T f(x_k)) - x^*| \\ &= | - F(x_k)^{-1}\nabla^T f(x_k) + (x_k - x^*)| \\ &= |F(x_k)^{-1}[-\nabla^T f(x_k) + F(x_k)(x_k - x^*)]| \\ &= |F(x_k)^{-1}||-\nabla^T f(x_k) + F(x_k)(x_k - x^*)| \\ &\leq \beta_1(|x - x^*|\beta_2) = \beta_1\beta_2|x - x^*|^2 \end{aligned}$$

Como $\beta_1\beta_2|x_k - x^*| < 1 \iff \beta_1\beta_2|x_k - x^*|^2 < |x_k - x^*|$. Portanto, $|x_{k+1} - x^*| \leq \beta_1\beta_2|x - x^*|^2 < |x_k - x^*|$, isto é, mostrou-se pela primeira desigualdade que a convergência do método é quadrática. Além disso, $|x_{k+1} - x^*| < |x_k - x^*|$ confirma que a sequência $\{x_k\}$ converge para x^* .

□

Um questionamento feito sobre o ponto $x_0 \in \mathbb{R}^n$ é se este encontra-se suficientemente próximo de x^* . Em geral, não sabemos dessa informação. Porém, pelo resultado acima, se $x_0 \in B(x^*; \epsilon)$, podemos afirmar que quanto menor o $\epsilon > 0$, maior será a eficiência do método. Ademais, o Método de Newton é computacionalmente caro, visto que a cada k -ésima iteração a matriz hessiana da função necessita ser calculada e invertida, tornando assim o processo inviável em diversas situações.

2.3 Métodos Quase-Newton

Após o estudo dos dois métodos acima, é notável que ambos possuem suas limitações. O método do Gradiente possui ordem de convergência linear, onde esta é dependente da distância dos autovalores da hessiana da f e, conforme o problema em questão, não há sequer convergência em tempo viável. Por sua vez, o Método de Newton possui melhores propriedades de convergência(ordem 2), porém sua eficiência é extremamente dependente do ponto inicial tomado, além do elevado custo computacional do cálculo da hessiana a cada iteração. Sendo assim, a filosofia dos métodos Quase-Newton consiste na procura de uma metodologia intermediária, isto é, que a sequência $\{x_k\}$ gerada por ele, quando converge, tenha ordem de convergência pelos menos superlinear e que seu trabalho computacional seja reduzido em comparação ao de Newton.

Lema 2.3.1. *Seja $d_k = -H_k\nabla f(x_k)$ com $H_k > 0$ e $\nabla f(x_k) \neq 0$, então d_k é direção de descida.*

Demonstração. $\nabla^T f(x_k)d_k = \nabla^T f(x_k)(-H_k\nabla f(x_k)) = (-1)[\nabla^T f(x_k)H_k\nabla f(x_k)] < 0$, pois $H_k > 0$ e $\nabla f(x_k) \neq 0$. □

Analisando mais uma vez o caso quadrático, considere $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $G > 0$, $f(x) = \frac{1}{2}x^T Gx + b^T x + c \implies \nabla f(x) = Gx + b$. Dessa forma, $\nabla f(x+d) - \nabla f(x) = G(x+d) + b - [Gx + b] = Gd, \forall d \in \mathbb{R}^n$. Logo, temos $\nabla f(x+d) - \nabla f(x) = Gd$ ou $G^{-1}(\nabla f(x+d) - \nabla f(x)) = d$. Portanto, note que obtemos informações sobre G e G^{-1} utilizando a diferença do valor do gradiente aplicado em dois pontos distintos. Assim, isto induz ao fato de que dado n pares de pontos $\{x_i, x_i + d_i\}$ de modo que $\{d_1, d_2, \dots, d_n\}$ seja linearmente independente, tem-se que as n diferenças $\nabla f(x_i + d_i) - \nabla f(x_i)$ determinam completamente G e G^{-1} .

Estas observações sugerem o seguinte algoritmo:

Sejam $x_0 \in \mathbb{R}^n$ arbitrário, $H_0 \in M_n(\mathbb{R})$, $H_0 > 0$, $H_0^T = H_0$. Se $\nabla f(x_k) \neq 0$, os passos a obter x_{k+1} são:

Passo 1: Calcular $d_k = -H_k \nabla f(x_k)$

Passo 2: Determinar λ_k através de uma busca linear e definir $x_{k+1} = x_k + \lambda_k d_k$

Passo 3: Determinar H_{k+1} simétrica e definida positiva tal que

$$H_{k+1}(\nabla f(x_{j+1}) - \nabla f(x_j)) = x_{j+1} - x_j, \forall j \leq k.$$

Se a função objetivo é quadrática e o conjunto $\{d_0, d_1, \dots, d_{n-1}\}$ é linearmente independente, pelas observações anteriores teremos que

$$H_n = G^{-1}.$$

Portanto,

$$d_n = -G^{-1}(\nabla f(x_n)) \quad \text{e} \quad x_{n+1} = x^*.$$

A pergunta a ser feita é: Há a possibilidade de construir um algoritmo com tais propriedades? A resposta é afirmativa. O primeiro método deste tipo foi proposto por Davidon, Fletcher e Powell e consiste no seguinte:

2.3.1 Algoritmo(DFP)

Inicialização do algoritmo: Seja $\epsilon > 0$. Escolha um ponto inicial x_1 e uma matriz simétrica definida positiva D_1 . Considere $y_1 = x_1$, $k = j = 1$ e vá para a Etapa Principal.

Etapa Principal:

Passo 1: Se $\|\nabla f(y_j)\| < \epsilon$, pare. Caso contrário, seja $d_j = -D_j \nabla f(y_j)$ e considere λ_j uma solução ótima para minimizar $f(y_j + \lambda d_j)$ sujeito a $\lambda \geq 0$. Seja $y_{j+1} = y_j + \lambda_j d_j$. Se $j < n$, vá para o Passo 2. Se $j = n$, seja $y_1 = x_{k+1} = y_{n+1}$, substitua k por $k + 1$, tome $j = 1$ e repita o passo 1 novamente.

Passo 2: Construa D_{j+1} segundo o processo abaixo:

$$D_{j+1} = D_j + \frac{p_j p_j^T}{p_j^T q_j} - \frac{D_j q_j q_j^T D_j}{q_j^T D_j q_j}$$

onde

$$p_j \equiv \lambda_j d_j = y_{j+1} - y_j$$

e

$$q_j \equiv \nabla f(y_{j+1}) - \nabla f(y_j)$$

Substitua j por $j + 1$ e vá ao Passo 1.

Enunciado a estrutura do DFP, iremos agora mostrar algumas propriedades relativas ao método.

Teorema 2.5. *Seja $y_1 \in \mathbb{R}^n$. Considere $D_1 > 0$ e $D_1 = D_1^T$. Para $j = 1, \dots, n$ considere $y_{j+1} = y_j + \lambda_j d_j$, em que $d_j = -D_j \nabla f(y_j)$, onde λ_j resolve o problema de minimizar $f(y_j + \lambda d_j)$ sujeita a $\lambda \geq 0$. Além disso, para $j = 1, \dots, n - 1$, considere D_{j+1} , p_j , e q_j da mesma maneira como foi explicitado no algoritmo DFP. Se $\nabla f(y_j) \neq 0$ para $j = 1, \dots, n$ com D_1, \dots, D_n positivas definidas e simétricas, então d_1, \dots, d_n são direções de descida.*

Demonstração. A demonstração será feita por indução em j . Para $j = 1$, temos que D_1 é simétrica e definida positiva por hipótese. Ora,

$$\nabla^T f(y_1) d_1 = \nabla^T f(y_1) (-D_1 \nabla f(y_1)) = -[\nabla^T f(y_1) D_1 \nabla f(y_1)] < 0$$

pois $D_1 > 0$. Portanto, d_1 é direção de descida.

Agora, suponha validade para $j \leq n - 1$. Mostraremos que a propriedade é válida para $j + 1$. De fato, considere $x \in \mathbb{R}^n$, $x \neq 0$. Tome a expressão para D_{j+1} , logo:

$$x^T (D_{j+1}) x = x^T D_j x + \frac{(x^T p_j)^2}{p_j^T q_j} - \frac{(x^T D_j q_j)^2}{q_j^T D_j q_j} (*)$$

Como D_j é simétrica e $D_j > 0$, então existe $D_j^{1/2}$ (veja o apêndice A de (LUENBERGER; YE, 2010)) tal que $D_j = D_j^{1/2} D_j^{1/2}$ e $D_j^{1/2} = (D_j^{1/2})^T$. Tome $a = D_j^{1/2} x$ e $b = D_j^{1/2} q_j \implies a^T a = (D_j^{1/2} x)^T (D_j^{1/2} x) = x^T (D_j^{1/2})^T D_j^{1/2} x = x^T D_j x$, $b^T b = q_j^T (D_j^{1/2})^T D_j^{1/2} q_j = q_j^T D_j q_j$. Além de termos que $a^t b = (D_j^{1/2} x)^T (D_j^{1/2} q_j) = x^T (D_j^{1/2})^T D_j^{1/2} q_j = x^T D_j q_j$. Em suma,

$$a^T a = x^T D_j x, \quad b^T b = q_j^T D_j q_j \quad \text{e} \quad a^T b = x^T D_j q_j (**)$$

Aplicando (**) em (*), segue que:

$$\begin{aligned} x^T(D_{j+1})x &= x^T D_j x + \frac{(x^T p_j)^2}{p_j^T q_j} - \frac{(x^T D_j q_j)^2}{q_j^T D_j q_j} \\ &= a^T a + \frac{(x^T p_j)^2}{p_j^T q_j} - \frac{(a^T b)^2}{(b^T b)} \\ &= \frac{(a^T a)(b^T b) - (a^T b)^2}{(b^T b)} + \frac{(x^T p_j)^2}{p_j^T q_j} \end{aligned}$$

Pela desigualdade de Cauchy-Schwartz (LIMA, 2020), temos que $(a^T a)(b^T b) - (a^T b)^2 \geq 0$. Para mostrar que $D_{j+1} > 0$, basta verificar o sinal das expressões $p_j^T q_j$ e $b^T b$.

Veja que $p_j = \lambda_j d_j = y_{j+1} - y_j$ e $q_j = \nabla f(y_{j+1}) - \nabla f(y_j) \implies p_j^T q_j = \lambda_j d_j^T [\nabla f(y_{j+1}) - \nabla f(y_j)] = \lambda_j (d_j^T \nabla f(y_{j+1})) - \lambda_j (d_j^T \nabla f(y_j))$. Lembre que λ_j resolve o problema de minimização de $\gamma(\lambda) = f(x_j + \lambda d_j)$ com $\lambda \geq 0$, ou seja, pelas condições necessárias de primeira ordem temos que $\gamma'(\lambda_j) = 0$. Assim, $\gamma'(\lambda) = \nabla^T f(y_j + \lambda d_j)(y_j + \lambda d_j)' = \nabla^T f(y_j + \lambda d_j)(d_j) \implies \gamma'(\lambda_j) = 0 = \nabla^T f(y_j + \lambda_j d_j)d_j = 0 \implies d_j^T \nabla f(y_{j+1}) = 0$, pois $y_{j+1} = y_j + \lambda_j d_j$.

Aplicando este último em $p_j^T q_j = \lambda_j (d_j^T \nabla f(y_{j+1})) - \lambda_j (d_j^T \nabla f(y_j)) = 0 - \lambda_j (d_j^T \nabla f(y_j)) = -\lambda_j (d_j^T \nabla f(y_j)) \implies p_j^T q_j = -\lambda_j (d_j^T \nabla f(y_j))$. Do fato de que $d_j = -D_j \nabla f(y_j) \implies p_j^T q_j = -\lambda_j (d_j^T \nabla f(y_j)) = -\lambda_j (-D_j \nabla f(y_j))^T \nabla f(y_j) = +\lambda_j \nabla^T f(y_j) D_j \nabla f(y_j) > 0$, visto que $\lambda_j > 0$, $D_j > 0$ e $\nabla f(y_j) \neq 0$, decorre que $p_j^T q_j > 0$. Como $q_j = \nabla f(y_{j+1}) - \nabla f(y_j) \neq 0$, pois $\nabla f(y_{j+1}) \neq \nabla f(y_j)$, logo $b^T b = q_j^T D_j q_j > 0$, devido à $D_j > 0$. Portanto, concluímos que $x^T D_{j+1} x \geq 0$, $\forall x \in \mathbb{R}^n$, $x \neq 0$.

Suponha por absurdo $x^T D_{j+1} x = 0$. Sabe-se que $D_{j+1} = \frac{(a^T a)(b^T b) - (a^T b)^2}{(b^T b)} + \frac{(x^T p_j)^2}{p_j^T q_j} \implies \frac{(a^T a)(b^T b) - (a^T b)^2}{(b^T b)} = \frac{-(x^T p_j)^2}{p_j^T q_j}$. Já que $p_j^T q_j > 0$, $b^T b > 0$, $(a^T a)(b^T b) - (a^T b)^2 \geq 0$ e $-(x^T p_j)^2 \leq 0$, segue que $(a^T a)(b^T b) - (a^T b)^2 = 0$ e $x^T p_j = 0$. Pela igualdade em Cauchy-Schwartz, temos que existe $\lambda \in \mathbb{R}$ tal que $a = \lambda b$. Do fato de que $x \neq 0 \implies \lambda \neq 0$ (recorde que $q_j = \nabla f(y_{j+1}) - \nabla f(y_j) \neq 0$). Ora, $x^T p_j = p_j^T x = p_j^T (\lambda q_j) = \lambda (p_j^T q_j)$. Da hipótese de absurdo, segue que $x^T p_j = 0 = \lambda (p_j^T q_j)$. No entanto, $\lambda \neq 0$ e $p_j^T q_j > 0$, o que confirma o absurdo. Logo, $x^T D_{j+1} x > 0$, $\forall x \neq 0$, o que quer dizer que D_{j+1} é definida positiva. Por fim, como $\nabla f(y_{j+1}) \neq 0$:

$$\begin{aligned} \nabla^T f(y_{j+1})(D_{j+1})\nabla f(y_{j+1}) > 0 &\implies -(\nabla^T f(y_{j+1})D_{j+1}\nabla f(y_{j+1})) < 0 \\ &\implies \nabla^T f(y_{j+1})(-D_{j+1}\nabla f(y_{j+1})) < 0 \\ &\implies \nabla^T f(y_{j+1})d_{j+1} < 0 \\ &\implies d_{j+1} \text{ é direção de descida.} \end{aligned}$$

□

Antes de enunciarmos o teorema principal referente ao método DFP para funções quadráticas, um resultado envolvendo vetores H -conjugados necessita ser apresentado.

Definição 13. (*Vetores H -conjugados*) Seja H uma matriz de ordem n e simétrica. Os vetores d_1, d_2, \dots, d_n são ditos H -conjugados ou conjugados se eles são linearmente independentes e $d_i^T H d_j = 0$ para $i \neq j$.

Note que se $f(x) = c^T x + \frac{1}{2} x^T H x$, H é de ordem n , $H = H^T$ e $\{d_1, d_2, \dots, d_n\}$ é um conjunto de vetores linearmente independentes, então tem-se que dado um $x \in \mathbb{R}^n$, é possível escrevê-lo de maneira única como $x = x_1 + \sum_{j=1}^n \lambda_j d_j$. Substituindo na expressão da quadrática, temos que:

$$\begin{aligned} f(x) &= c^T x + \frac{1}{2} x^T H x \\ &= c^T \left[x_1 + \sum_{j=1}^n \lambda_j d_j \right] + \frac{1}{2} \left[x_1 + \sum_{j=1}^n \lambda_j d_j \right]^T H \left[x_1 + \sum_{j=1}^n \lambda_j d_j \right] \\ &= \sum_{j=1}^n \left[c^T (x_1 + \lambda_j d_j) + \frac{1}{2} (x_1 + \lambda_j d_j)^T H (x_1 + \lambda_j d_j) \right] \end{aligned}$$

Perceba que f é separável conforme o índice $j = 1, 2, \dots, n$, de maneira que é possível minimizar cada termo $(x_1 + \lambda_j d_j)$. Ademais, minimizar $[\bullet]$, ou seja, cada parcela do somatório referente à $f(x)$, é o mesmo que minimizar f partindo do ponto x_1 e caminhando ao longo da direção d_j . Se $H > 0$, é fácil exibir explicitamente λ_j^* que minimiza $f(x_1 + \lambda_j d_j)$. Com efeito,

$$\begin{aligned} \nabla f \left(x_1 + \sum_{j=1}^n \lambda_j^* d_j \right) &= \frac{1}{2} d_j^T H x_1 + \frac{1}{2} x_1^T H d_j + \lambda_j^* d_j^T H d_j + d_j^T c \\ &\implies -\lambda_j^* d_j^T H d_j = d_j^T H x_1 + c^T d_j \\ &\implies \lambda_j^* = - \left[\frac{d_j^T H x_1 + c^T d_j}{d_j^T H d_j} \right] \end{aligned}$$

Eis agora o principal resultado relativo à vetores H -conjugados.

Teorema 2.6. *Seja $f(x) = c^T x + \frac{1}{2} x^T H x$, $H = H^T$ e H de ordem n . Considere d_1, d_2, \dots, d_n H -conjugados e x_1 um ponto de partida arbitrário. Para $k = 1, \dots, n$, seja λ_k uma solução ótima para o problema de minimizar $f(x_k + \lambda d_k)$ com $\lambda \in \mathbb{R}$ e $x_{k+1} = x_k + \lambda_k d_k$. Portanto, tem-se que:*

- i) $\nabla^T f(x_{k+1}) d_j = 0$ para $j = 1, 2, \dots, k$
- ii) $\nabla^T f(x_1) d_k = \nabla^T f(x_k) d_k$

iii) x_{k+1} é uma solução ótima para minimizar $f(x)$ sujeita a $x - x_1 \in \text{span}\{d_1, d_2, \dots, d_k\}$, isto é, $x - x_1$ é um vetor do espaço gerado pelos vetores H -conjugados.

Demonstração. Para i), $f(x_j + \lambda_j d_j)$ atinge mínimo em λ_j , logo $\nabla f(x_j + \lambda_j d_j) = 0 \implies \nabla^T f(x_j + \lambda_j d_j) d_j = 0 \implies \nabla^T f(x_{j+1}) d_j = 0$, confirmando para $j = k$. Se $j < k$, $f(x) = c^T x + \frac{1}{2} x^T H x \implies \nabla f(x) = c + H x \implies \nabla f(x_{k+1}) = c + H x_{k+1}$. Ora,

$$\begin{aligned} \nabla f(x_{k+1}) - \nabla f(x_{j+1}) &= [H x_{k+1} + c] - [H x_{j+1} + c] \\ &= H [x_{k+1} - x_{j+1}] \\ &= H \left[x_1 + \sum_{i=1}^k \lambda_i d_i - \left(x_1 + \sum_{g=1}^j \lambda_g d_g \right) \right] \\ &= H \left[\sum_{i=j+1}^k \lambda_i d_i \right] \\ &\implies \nabla f(x_{k+1}) = \nabla f(x_{j+1}) + H \left[\sum_{i=j+1}^k \lambda_i d_i \right] (*) \end{aligned}$$

Multiplicando por d_j com $j < k$, $d_j^T \nabla f(x_{k+1}) = d_j^T \nabla f(x_{j+1}) + d_j^T H \left[\sum_{i=j+1}^k \lambda_i d_i \right] = 0 + 0 = 0$ por $\nabla^T f(x_{j+1}) d_j = 0$ e pelo fato de que d_1, \dots, d_n são H -conjugados, isto é, $d_i^T H d_j = 0$ para $i \neq j$.

Para ii), tome novamente a expressão (*). Coloque $j = 0$, o que é possível, pois $j < k$ e troque k por $k - 1$. Segue que $\nabla f(x_{(k-1)+1}) = \nabla f(x_1) + H \left(\sum_{i=1}^{k-1} \lambda_i d_i \right) \implies \nabla f(x_k) = \nabla f(x_1) + H \left(\sum_{i=1}^{k-1} \lambda_i d_i \right) \implies d_k^T \nabla f(x_k) = d_k^T \nabla f(x_1) + d_k^T H \left(\sum_{i=1}^{k-1} \lambda_i d_i \right) = d_k^T \nabla f(x_1)$. Assim, $d_k^T \nabla f(x_k) = d_k^T \nabla f(x_1) \implies \nabla^T f(x_1) d_k = \nabla^T f(x_k) d_k$.

Para iii),

$$\begin{aligned} f(x_{k+1}) &= f((x_{k+1} - x_1) + x_1) \\ &= f \left(\left(x_1 + \sum_{j=1}^n \lambda_j d_j - x_1 \right) + x_1 \right) \\ &= f \left(x_1 + \sum_{j=1}^n \lambda_j d_j \right) \\ &= c^T \left[x_1 + \sum_{j=1}^n \lambda_j d_j \right] + \frac{1}{2} \left[x_1 + \sum_{j=1}^n \lambda_j d_j \right]^T H \left[x_1 + \sum_{j=1}^n \lambda_j d_j \right] \end{aligned}$$

Realizando algumas manipulações algébricas e usando a definição de H -conjugação, obtemos $f(x_{k+1}) = f(x_1) + \nabla^T f(x_1) \left(\sum_{j=1}^n \lambda_j d_j \right) + \frac{1}{2} \left(\sum_{j=1}^n \lambda_j^2 d_j^T H d_j \right)$. Da hipótese da parte

iii), $x - x_1 \in \text{span}\{d_1, \dots, d_k\} \implies x = x_1 + \sum_{j=1}^k \mu_j d_j$. Utilizando tal argumento na expressão

de $f(x_{k+1})$, temos que $f(x_{k+1}) = f(x_1) + \nabla^T f(x_1) \left(\sum_{j=1}^n \lambda_j d_j \right) + \frac{1}{2} \left(\sum_{j=1}^n \lambda_j^2 d_j^T H d_j \right) \implies$
 $f(x) = f(x_1) + \nabla^T f(x_1) \left(\sum_{j=1}^n \mu_j d_j \right) + \frac{1}{2} \left(\sum_{j=1}^n \mu_j^2 d_j^T H d_j \right)$.

Visando o término da prova, basta mostrar que $f(x) \geq f(x_{k+1})$. De fato, suponha por absurdo que $f(x) < f(x_{k+1})$, logo:

$$\nabla^T f(x_1) \left(\sum_{j=1}^n \mu_j d_j \right) + \frac{1}{2} \left(\sum_{j=1}^n \mu_j^2 d_j^T H d_j \right) < \nabla^T f(x_1) \left(\sum_{j=1}^n \lambda_j d_j \right) + \frac{1}{2} \left(\sum_{j=1}^n \lambda_j^2 d_j^T H d_j \right) (**)$$

Recorde que λ_j resolve o problema de minimizar $f(x_j + \lambda d_j)$.

Portanto, $f(x_j + \lambda_j d_j) \leq f(x_j + \mu_j d_j)$ para $j = 1, \dots, k$. Fazendo uso da Fórmula de Taylor, decorre:

$$f(x_j) + \nabla^T f(x_j)(\lambda_j d_j) + \frac{1}{2}(\lambda_j d_j)^T H(\lambda_j d_j) \leq f(x_j) + \nabla^T f(x_j)(\mu_j d_j) + \frac{1}{2}(\mu_j d_j)^T H(\mu_j d_j)$$

Ou seja,

$$\nabla^T f(x_j)(\lambda_j d_j) + \frac{1}{2}(\lambda_j d_j)^T H(\lambda_j d_j) \leq \nabla^T f(x_j)(\mu_j d_j) + \frac{1}{2}(\mu_j d_j)^T H(\mu_j d_j)$$

Além disso, utilizando o resultado da parte ii) ($\nabla^T f(x_j) d_j = \nabla^T f(x_1) d_j$), escrevendo as k inequações e somando-as membro a membro, obtemos:

$$\nabla^T f(x_1) \left(\sum_{j=1}^k \lambda_j d_j \right) + \frac{1}{2} \sum_{j=1}^k \lambda_j^2 d_j^T H d_j \leq \nabla^T f(x_1) \left(\sum_{j=1}^k \mu_j d_j \right) + \frac{1}{2} \sum_{j=1}^k \mu_j^2 d_j^T H d_j (***)$$

Onde (***) contradiz (**). Logo, $f(x) \geq f(x_{k+1})$. Isto é, x_{k+1} é ponto de mínimo pertencente à variedade $x_1 + \text{span}\{d_1, \dots, d_k\}$.

□

Eis agora um importante resultado para o DFP envolvendo funções quadráticas.

Teorema 2.7. (DFP - Caso Quadrático) *Seja H uma matriz de ordem n simétrica e definida positiva. Considere o problema de minimizar $f(x) = c^T x + \frac{1}{2}x^T Hx$ com $x \in \mathbb{R}^n$. Suponha que o problema seja resolvido pelo DFP, começando pelo ponto inicial y_1 e matriz simétrica positiva D_1 . Em particular, para $j = 1, \dots, n$, seja λ_j uma solução ótima para o problema de minimização de $f(y_j + \lambda d_j)$ sujeita a $\lambda \geq 0$ e seja $y_{j+1} = y_j + \lambda_j d_j$, onde $d_j = -D_j \nabla f(y_j)$ e D_j, p_j e q_j são determinados conforme exposto no algoritmo do DFP acima. Se $\nabla f(y_j) \neq 0$ para cada j , então as direções d_1, d_2, \dots, d_n são H -conjugadas e $D_{n+1} = H^{-1}$. Além disso, y_{n+1} é solução ótima para o problema.*

Demonstração. O esquema principal da prova consiste em realizar indução em j para mostrar as três afirmações abaixo:

- 1) d_1, \dots, d_n são linearmente independentes.
- 2) $d_i^T H d_k = 0, \forall i \neq k$ com $i, k \leq j$.
- 3) $D_{j+1} H p_k = p_k$ ou, equivalentemente, $D_{j+1} H d_k = d_k$ para $1 \leq k \leq j$, onde $p_k = \lambda_k d_k$.

Para $j = 1$, temos que as afirmações 1) e 2) são óbvias. Quanto à 3), inicialmente, note que para $1 \leq k \leq j$, tem-se que $H p_k = H(\lambda_k d_k) = H(y_{k+1} - y_k) = H y_{k+1} - H y_k = (H y_{k+1} + c) - (H y_k + c) = \nabla f(y_{k+1}) - \nabla f(y_k) = q_k \implies H p_k = q_k$. Como $1 \leq k \leq j$, tome $k = 1 \implies H p_1 = q_1$. Logo,

$$D_2(H p_1) = (D_{1+1} H p_1) = \left(D_1 + \frac{p_1 p_1^T}{p_1^T q_1} - \frac{D_1 q_1 q_1^T D_1}{q_1^T D_1 q_1} \right) q_1 = D_1 q_1 + \frac{p_1 p_1^T}{p_1^T q_1} q_1 - \frac{D_1 q_1 q_1^T D_1}{q_1^T D_1 q_1} q_1 = D_1 q_1 + \frac{p_1 p_1^T}{p_1^T q_1} q_1 - D_1 q_1 = p_1 \implies D_2 H p_1 = p_1, \text{ portanto a afirmativa 3) é válida para } j = 1.$$

Agora, suponha que as partes 1), 2) e 3) sejam simultaneamente válidas para $j \leq n - 1$. Queremos mostrar que as propriedades são válidas para $j + 1$.

Com efeito, recorde do resultado do teorema anterior, temos que $d_i^T \nabla f(y_{j+1}) = 0$ para $i \leq j$. A hipótese de indução da parte 3) diz que $D_{j+1} H d_i = d_i$ para $i \leq j$. Assim,

$$\begin{aligned} d_i &= D_{j+1} H d_i \implies d_i^T = d_i^T H^T D_{j+1}^T = d_i^T H D_{j+1} \implies d_i^T = d_i^T H D_{j+1} \implies \\ d_i^T \nabla f(y_{j+1}) &= d_i^T H D_{j+1} \nabla f(y_{j+1}) \implies d_i^T \nabla f(y_{j+1}) = d_i^T H D_{j+1} \nabla f(y_{j+1}) = 0 \implies \\ d_i^T \nabla f(y_{j+1}) + d_i^T H D_{j+1} \nabla f(y_{j+1}) &= 0. \text{ Do enunciado, } d_{j+1} = -D_{j+1} \nabla f(y_{j+1}) \implies 0 = \\ d_i^T H (D_{j+1} \nabla f(y_{j+1})) &= -d_i^T H d_{j+1}. \text{ Portanto, } d_i^T \nabla f(y_{j+1}) + d_i^T H D_{j+1} \nabla f(y_{j+1}) = -d_i^T H d_{j+1} = \\ 0 &\implies d_i^T H d_{j+1} = 0. \text{ Dessa maneira, a parte 2) é válida para } j + 1. \end{aligned}$$

Considerando $k \leq j + 1$, tome a expressão referente à D_{j+1} , decorre:

$$D_{j+2} = D_{(j+1)+1} = D_{j+1} + \frac{p_{j+1} p_{j+1}^T}{p_{j+1}^T q_{j+1}} - \frac{D_{j+1} q_{j+1} q_{j+1}^T D_{j+1}}{q_{j+1}^T D_{j+1} q_{j+1}}(i)$$

Realizando procedimento análogo ao do parágrafo anterior, obtemos que $D_{j+2} H p_{j+1} =$

p_{j+1} . Seja $k \leq j$. Como a parte 2) já foi provada, segue que $d_i^T H d_k = 0$ para $i \neq k$ e $i, k \leq j$. Assim, $d_{j+1}^T H d_j = 0 \implies (\lambda_{j+1} p_{j+1})^T H (\lambda_j p_j) = 0$ ou $(d_{j+1})^T H (d_j) = (\lambda_{j+1} p_{j+1})^T H (\lambda_j p_j) = (\lambda_{j+1} \lambda_j) p_{j+1}^T H p_j = 0$ (ii).

Tome a hipótese de indução para parte 3), o fato de que $H p_k = q_k$ e o resultado obtido na parte 2) para $j + 1$, segue que:

$$q_{j+1}^T (D_{j+1} H p_k) = q_{j+1}^T p_k = (H p_{j+1})^T p_k = p_{j+1}^T H^T p_k = p_{j+1}^T H p_k = (\lambda_{j+1} d_{j+1})^T H (\lambda_k d_k) = (\lambda_{j+1} \lambda_k) (d_{j+1})^T H d_k = 0 \implies q_{j+1}^T (D_{j+1} H p_k) = 0 \text{ (iii)}.$$

Aplicando (iii) e (ii) em (i):

$$\begin{aligned} D_{j+2} H p_k &= \left(D_{j+1} + \frac{p_{j+1} p_{j+1}^T}{p_{j+1}^T q_{j+1}} - \frac{D_{j+1} q_{j+1} q_{j+1}^T D_{j+1}}{q_{j+1}^T D_{j+1} q_{j+1}} \right) H p_k \\ &= D_{j+1} H p_k + \frac{p_{j+1} p_{j+1}^T}{p_{j+1}^T q_{j+1}} H p_k - \frac{D_{j+1} q_{j+1} (q_{j+1}^T D_{j+1} H p_k)}{q_{j+1}^T D_{j+1} q_{j+1}} \\ &= D_{j+1} H p_k + 0 + 0 \\ &= D_{j+1} H p_k \\ &= p_k \implies D_{j+2} H p_k = p_k, \text{ mostrando a validade da parte 3) para } j+1. \end{aligned}$$

Mostraremos a validade da parte 1) para $j + 1$. De fato, suponha que $\sum_{i=1}^{j+1} \alpha_i d_i = 0 \implies (d_{j+1}^T H) \sum_{i=1}^{j+1} \alpha_i d_i = (d_{j+1}^T H) 0 = 0 \implies \sum_{i=1}^{j+1} \alpha_i d_{j+1}^T H d_i = 0$. Da validade da parte 2) para $j + 1 \implies \alpha_1 d_{j+1}^T H d_1 + \alpha_2 d_{j+1}^T H d_2 + \alpha_3 d_{j+1}^T H d_3 + \dots + \alpha_{j+1} d_{j+1}^T H d_{j+1} = 0 \implies \alpha_{j+1} d_{j+1}^T H d_{j+1} = 0$. Ora, da hipótese, temos que $\nabla f(y_{j+1}) \neq 0$. Já que $D_{j+1} > 0$, o Teorema 2.5 garante que $D_{j+1} > 0 \implies \nabla^T f(y_{j+1}) D_{j+1} \nabla f(y_{j+1}) > 0$, portanto, se $-D_{j+1} \nabla f(y_{j+1}) = 0$, a última desigualdade não seria verdadeira, logo $d_{j+1} = -D_{j+1} \nabla f(y_{j+1}) \neq 0$. Como $H > 0 \implies d_{j+1}^T H d_{j+1} > 0$. Portanto, se $\alpha_{j+1} (d_{j+1}^T H d_{j+1}) = 0 \implies \alpha_{j+1} = 0$. Assim, $\sum_{i=1}^{j+1} \alpha_i d_i = 0 \implies \sum_{i=1}^j \alpha_i d_i + \alpha_{j+1} d_{j+1} = 0 \implies \sum_{i=1}^j \alpha_i d_i + 0 = 0 \implies \alpha_1 = \alpha_2 = \dots = \alpha_j = \alpha_{j+1} = 0$ (fazendo, claramente, uso da hipótese de indução da parte 1). Ou seja, concluímos que $\{d_1, d_2, \dots, d_j, d_{j+1}\}$ é um conjunto linearmente independente.

Por fim, considere $j = n$ na parte 3), isto é, $D_{n+1} H p_k = p_k \iff D_{n+1} H d_k = d_k$ para $1 \leq k \leq n$, $p_k = \lambda_k d_k$.

Seja D uma matriz onde cada coluna é d_k para $k = 1, 2, \dots, n$. Pela equivalência anterior, é nítido que $D_{n+1} H D = D$. Consequentemente, $D X = 0 \implies \begin{bmatrix} d_1 & d_2 & \dots & d_n \end{bmatrix} X = 0 \implies \sum_{i=1}^n d_i x_i = 0 \implies x_i = 0$ para $i = 1, \dots, n$ graças à parte 3). Dessa maneira, o sistema linear $D X = 0$ apenas aceita solução trivial, logo D é invertível (olhe (HOFFMAN; KUNZE, 1971)).

Logo, como $D_{n+1}HD = D \implies D_{n+1}HDD^{-1} = DD^{-1} = I \implies D_{n+1}H = I$. Pelo fato de H ser simétrica e definida positiva, segue que $D_{n+1}H = I \implies D_{n+1}HH^{-1} = H^{-1} \implies D_{n+1} = H^{-1}$. Pelo Teorema 2.6, concluímos que y_{n+1} é solução ótima. \square

2.3.2 Análise da Obtenção do DFP

A metodologia do DFP consiste no uso da direção $d_j = -D_j \nabla f(y_j)$, onde D_j é uma aproximação da hessiana, simétrica e positiva definida, sendo o objetivo desviar o simétrico do gradiente no ponto y_j utilizando D_j , assim como se era feito no Método de Newton, com a mudança de que quem realizava esse desvio era a própria hessiana a cada j -ésimo passo, pois a direção de Newton é $d_j = -(\nabla^2 f(y_j))^{-1} \nabla f(y_j)$.

No DFP, visando se obter D_{j+1} , perceba que temos a adição de D_j à duas matrizes de posto 1, visto que elas são do formato $\alpha(uu^T)$, $u \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$. Com efeito, seja $u =$

$$\begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{bmatrix} \implies uu^T = \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{bmatrix}^T = \begin{bmatrix} u_1u_1 & u_1u_2 & \cdot & \cdot & \cdot & u_1u_n \\ u_2u_1 & u_2u_2 & \cdot & \cdot & \cdot & u_2u_n \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ u_nu_1 & u_nu_2 & \cdot & \cdot & \cdot & u_nu_n \end{bmatrix}. \text{ É fácil ver que a } i\text{-ésima}$$

linha é o termo $u_i u^T$ ou, equivalentemente, que cada j -ésima coluna é o termo $u_j u$, isto é, todas as linhas são geradas apenas pelo vetor u^T e todas as colunas por u , logo $\text{posto}(uu^T) = 1$. Posteriormente, é feito um processo de busca linear e baseado no resultado da solução y_{j+1} e de $\nabla f(y_{j+1})$, obtemos a aproximação D_{j+1} conforme já exposto.

Como visto no teorema anterior, se $f(x) = c^T x + \frac{1}{2} x^T H x$, $x \in \mathbb{R}^n$, $H = H^T$, $H > 0$ e $\nabla f(y_j) \neq 0$ para $j = 1, 2, \dots, n$, então $D_{n+1} = H^{-1}$. Lembre que concluímos que os vetores $p_j = \lambda_j d_j$ são linearmente independentes (Parte 1). Além disso, a Parte 3) afirma que $D_{j+1} H p_k = p_k \iff D_{j+1} H d_k = d_k$ para $1 \leq k \leq j$. Ou seja, p_1, p_2, \dots, p_k são autovetores associados à matriz $D_{j+1} H$, cujos seus respectivos autovalores são iguais a 1. Por consequência, perceba que a cada passo do método, a aproximação da hessiana acumula um autovetor p_k linearmente independente dos anteriores p_1, p_2, \dots, p_{k-1} , até chegar ao n -ésimo passo, onde tem-se que $D_{n+1} H$ pode ser expressa na forma $D_{n+1} H P = P$, em que P é a matriz de autovetores (invertível) de $D_{n+1} H$.

Dessa forma, baseado nas observações acima, vamos alterar o DFP para outros esquemas de atualização de D_{j+1} . Até o fim, suponha $D_j = D_j^T$, $D_j > 0$ e D_j sendo a aproximação da inversa da hessiana de f , onde p_1, p_2, \dots, p_{j-1} são autovetores de $D_j H$ com autovalores iguais a 1. Assim, dado o ponto y_j (o ponto atual), faça uma busca linear ao longo da direção $d_j = -D_j \nabla f(y_j)$ para obter y_{j+1} conforme a seguinte definição:

$$\begin{cases} p_j = y_{j+1} - y_j \\ q_j = \nabla f(y_{j+1}) - \nabla f(y_j) = H(y_{j+1} - y_j) = Hp_j \end{cases}$$

Adotando o mesmo esquema indutivo do Teorema 2.7, obteremos o resultado de que p_1, \dots, p_{j-1} ($p_k = \lambda_k d_k$, $1 \leq k \leq j$) são linearmente independentes e H-conjugados. Agora, deseja-se construir a seguinte atualização,

$$D_{j+1} = D_j + C_j \quad \text{(i)}$$

sendo D_j a aproximação obtida no passo anterior e C_j uma matriz simétrica de correção com a propriedade de possuir p_1, p_2, \dots, p_j como autovetores de $D_{j+1}H$ com respectivos autovalores iguais a 1. Logo, $D_{j+1}Hp_k = p_k$, ou do fato de que $q_j = \nabla f(y_{j+1}) - \nabla f(y_j) = Hp_j$, tem-se que $D_{j+1}Hp_k = D_{j+1}q_k = p_k$ para $1 \leq k \leq j$.

$$\text{Portanto, } p_k = D_{j+1}q_k = (D_j + C_j)q_k = D_jq_k + C_jq_k \implies p_k = p_k + C_jq_k \implies$$

$$C_jq_k = 0, \quad 1 \leq k \leq j, \quad k \in \{1, 2, \dots, j-1\} \quad \text{(ii)}$$

Quando $k = j$, a equação

$$D_{j+1}q_j = p_j \quad \text{(iii)}$$

é chamada de **equação secante** ou **condição Quase-Newton**. Dessa maneira, retornando à equação (i), temos que $D_{j+1} = D_j + C_j \implies D_{j+1}q_j = D_jq_j + C_jq_j \implies p_j = D_jq_j + C_jq_j \implies C_jq_j = p_j - D_jq_j$.

Ora, se $C_j = \frac{p_j p_j^T}{p_j^T q_j} - \frac{(D_j q_j)(D_j q_j)^T}{(D_j q_j)^T q_j}$, então $C_j \equiv C_j^{DFP}$, ou seja, temos a atualização do método DFP. Primeiramente, veja que C_j^{DFP} atende as condições da equação secante, pois conforme já mostrado no Teorema 2.7, temos que D_{j+1} será simétrica e positiva definida. Além disso, $C_j q_k = 0$ para $1 \leq k < j$ também é satisfeita ao utilizarmos C_j^{DFP} .

$$\begin{aligned} \text{Com efeito, } C_j q_k &= C_j^{DFP}(Hp_k) = \left[\frac{p_j p_j^T}{p_j^T q_j} - \frac{(D_j q_j)(D_j q_j)^T}{(D_j q_j)^T q_j} \right] (Hp_k) = \frac{p_j p_j^T Hp_k}{p_j^T q_j} - \\ &= \frac{(D_j q_j)(D_j q_j)^T Hp_k}{(D_j q_j)^T q_j} = \frac{p_j (p_j^T Hp_k)}{p_j^T q_j} - \frac{(D_j q_j)(q_j^T D_j) Hp_k}{(D_j q_j)^T q_j} = \frac{p_j (p_j^T Hp_k)}{p_j^T q_j} - \frac{(D_j q_j)(q_j^T D_j) Hp_k}{(D_j q_j)^T q_j} = \\ &= \frac{p_j (p_j^T Hp_k)}{p_j^T q_j} - \frac{(D_j q_j)((Hp_j)^T D_j) Hp_k}{(D_j q_j)^T q_j} = \frac{p_j (p_j^T Hp_k)}{p_j^T q_j} - \frac{(D_j q_j)(p_j^T H(D_j Hp_k))}{(D_j q_j)^T q_j} = \frac{p_j (p_j^T Hp_k)}{p_j^T q_j} - \\ &= \frac{(D_j q_j)(p_j^T H(p_k))}{(D_j q_j)^T q_j} = \frac{p_j (p_j^T Hp_k)}{p_j^T q_j} - \frac{(D_j q_j)(p_j^T Hp_k)}{(D_j q_j)^T q_j} = 0, \text{ pois } p_j^T Hp_k = 0, \text{ visto que } k < j \text{ (H-} \\ &\text{conjugação) e } p_k \text{ é autovetor de } D_j H \text{ com autovalor igual a 1. Portanto, seguindo esta sequência} \\ &\text{de correções, obtemos o resultado do último teorema, isto é, } H^{-1} = D_{n+1}. \end{aligned}$$

2.3.3 Atualizações Família Broyden e BFGS

A partir da seção anterior, percebe-se que há flexibilidade na descrição da correção C_j para satisfazer a equação secante, isto é, $D_{j+1}q_j = p_j$. Entretanto, há a necessidade de manter $C_jq_j = p_j - D_jq_j$, $C_jq_k = 0$, $D_{j+1} = D_{j+1}^T$, $D_{j+1} > 0$ com $1 \leq k < j$. Por conseguinte, tendo em vista essas condições, Broyden sugere uma correção $C_j = C_j^B$ dada pela família parametrizada por ϕ :

$$C_j^B \equiv C_j^{DFP} + \frac{\phi\tau_j v_j v_j^T}{p_j^T q_j}$$

Onde $v_j \equiv p_j - \left(\frac{1}{\tau_j}\right) D_j q_j$, sendo τ_j escolhido para que a condição $C_j q_j = p_j - D_j q_j$ seja válida caso $v_j^T q_j = 0$. Se $v_j^T q_j = 0 \implies \left[\frac{\phi\tau_j v_j v_j^T}{p_j^T q_j}\right]^T q_j = 0 q_j \implies p_j^T q_j - \frac{[D_j q_j]^T q_j}{\tau_j} = 0 \implies \tau_j = \frac{[D_j q_j]^T q_j}{p_j^T q_j} = \frac{q_j^T D_j q_j}{p_j^T q_j}$. Logo, para $1 \leq k < j$, temos que $v_j^T q_k = \left[p_j^T - \frac{q_j^T D_j}{\tau_j}\right] q_k = p_j^T q_k - \frac{q_j^T D_j q_k}{\tau_j} = p_j^T (H p_k) - \frac{(H p_j)^T (D_j H p_k)}{\tau_j} = 0 - \frac{p_j^T H p_k}{\tau_j} = 0 - 0 = 0 \implies v_j^T q_k = 0$. Portanto, $C_j^B q_k = C_j^{DFP} q_k + \frac{\phi\tau_j v_j (v_j^T q_k)}{p_j^T q_j} = 0 + 0 = 0 \implies C_j^B q_k = 0$. Ademais, é claro que $D_{j+1} = D_j + C_j^B$ mantém simetria e para $\phi \leq 0$ tem-se $D_{j+1} > 0$. Por sua vez, note que a atualização de Broyden satisfaz as hipóteses do Teorema 2.7.

Analisando $\phi = 1$, tem-se que a família Broyden coincide com o BFGS (Broyden-Fletcher-Goldfarb-Shanno), em que tal método apresenta maior eficiência em diversos estudos computacionais em comparação a outros esquemas de atualização. Por exemplo, em comparação ao DFP, uma problemática exibida por este é o fato de se produzir aproximações de hessianas quase singulares a cada iteração, isto é, em determinados problemas, o DFP faz com que a atualização de D_j tenda a uma matriz não-invertível (BAZARAA, 2006).

No intuito de obter a correção relativa ao BFGS, substitua $\tau_j = \frac{q_j^T D_j q_j}{p_j^T q_j}$ em $C_j^B = C_j^{DFP} + \frac{\phi\tau_j v_j v_j^T}{p_j^T q_j}$. Com efeito, $C_j^{BFGS} \equiv C_j^B(\phi = 1) = C_j^{DFP} + \frac{q_j^T D_j q_j}{p_j^T q_j} \frac{v_j v_j^T}{p_j^T q_j}$ (I). Recorde que $v_j = p_j - \left(\frac{1}{\tau_j}\right) D_j q_j \implies v_j = p_j - \left(\frac{p_j^T q_j}{q_j^T D_j q_j}\right) D_j q_j \implies v_j^T = p_j^T - \left[\left(\frac{p_j^T q_j}{q_j^T D_j q_j}\right) D_j q_j\right]^T$. Efetuando algumas operações, obtemos que:

$$v_j v_j^T = \left[p_j p_j^T - \frac{p_j (q_j^T D_j) (q_j^T p_j)}{q_j^T D_j q_j} - \left(\frac{p_j^T q_j}{q_j^T D_j q_j}\right) D_j q_j p_j^T + \frac{p_j^T q_j D_j q_j}{q_j^T D_j q_j} \left(\frac{q_j^T D_j q_j p_j^T}{q_j^T D_j q_j}\right) \right] \quad (II)$$

Aplicando (II) em (I) e realizando simplificações, decorre que:

$$C_j^B(\phi = 1) = C_j^{BFGS} = \frac{p_j p_j^T}{p_j^T q_j} \left(1 + \frac{q_j^T D_j q_j}{p_j^T q_j} \right) - \left(\frac{D_j q_j p_j^T + p_j q_j^T D_j}{p_j^T q_j} \right)$$

Na equação da família Broyden, veja que para $\phi = 0$, temos $C_j^B = C_j^{DFP}$. Assim,

$$C_j^B = (1 - \phi)C_j^{DFP} + \phi C_j^{BFGS}$$

Se ϕ for constante, diz-se que é um método de Broyden puro. Porém, para a obtenção de resultados analíticos não há necessidade de fixar este parâmetro. Ou seja, é possível escolher ϕ_{j+1} no passo $j + 1$ distinto de ϕ_j caso seja relevante na perspectiva do problema trabalhado. Entretanto, qualquer parâmetro ϕ satisfaz as condições acima? A resposta é negativa. Por exemplo, a escolha de ϕ_j precisa ser devidamente analisada visto que para $\phi = \frac{-[\nabla^T f(y_j) D_j \nabla f(y_j)] q_j^T D_j q_j}{\nabla^T f(y_{j+1}) D_j \nabla f(y_{j+1})}$ temos que o algoritmo estagnará (BAZARAA, 2006), onde D_{j+1} se torna singular e perde a característica de ser positiva definida, logo tal valor de ϕ é dito degenerado e precisa ser evitado. Dessa forma, em geral, ϕ será não-negativo, porém em alguns momentos, o valor $\phi < 0$ apresenta-se eficiente em certas perspectivas computacionais (BAZARAA, 2006).

Por fim, um resultado relevante é que para funções diferenciáveis, se buscas lineares exatas são feitas, então pode ser mostrado que a sequência de pontos gerada pela família Broyden é invariante com respeito à escolha do parâmetro ϕ . Todavia, a seleção de um ϕ específico é essencial quando buscas lineares inexatas (como as de Armijo) são realizadas, sendo a principal preocupação em questão, a manutenção da positividade definida da aproximação da hessiana.

2.4 Método dos Gradientes Conjugados

Historicamente, o método foi proposto inicialmente por Hestenes e Stiefel em 1952 (BAZARAA, 2006) visando resolver sistemas de equações lineares. Em 1964, o método foi estendido por Fletcher e Reeves, onde o objetivo principal era a resolução de equações não-lineares e problemas de minimização irrestrita. Apesar de esses métodos serem tipicamente menos eficientes e menos robustos que os métodos Quase-Newton, eles requisitam baixo armazenamento de memória, havendo necessidade do uso de apenas 3 vetores n -dimensionais a cada iteração. Sendo assim, o Método de Fletcher e Reeves é indispensável para problemas envolvendo dimensões $n \geq 100$, visto que Métodos Quase-Newton caracterizam-se impraticáveis por conta do tamanho da matriz hessiana. Dessa maneira, entre as vantagens do método, temos a simplicidade deste, além do desvio do simétrico do gradiente usando a direção anterior. Em comparação à perspectiva dos Métodos Quase-Newton, tal desvio pode ser interpretado alternativamente como a atualização de uma matriz definida positiva, fixa, simétrica, em que a matriz identidade

é geralmente utilizada. Por esta razão, em alguns momentos o Método dos Gradientes Conjugados é dito de métrica fixa, em contraste aos métodos de métrica variável como os Quase-Newton. Ademais, como as direções do método são conjugadas, e lembrando do resultado referente a este aspecto na seção anterior, temos que o método irá convergir em no máximo n iterações quando uma busca linear exata é feita.

2.4.1 Abordagem Básica do Método dos Gradientes Conjugados

Visando minimizar $f : \mathbb{R}^n \rightarrow \mathbb{R}$, f diferenciável, tem-se que uma sequência $\{y_j\}$ é gerada de acordo com $y_{j+1} = y_j + \lambda_j d_j$, sendo d_j a direção de descida e λ_j o tamanho do passo que minimiza f ao longo da direção d_j partindo do ponto y_j . Para $j = 1$, utiliza-se $d_1 = -\nabla f(y_1)$ (a direção do gradiente) e para as iterações subsequentes, se $\nabla f(y_{j+1}) \neq 0$, é usado $d_{j+1} = -\nabla f(y_{j+1}) + \alpha_j d_j$, $j \geq 1$. Sendo α_j o parâmetro responsável pelo desvio que caracteriza um particular método do gradiente conjugado.

Outra forma de escrever a expressão para d_{j+1} sempre que $\alpha_j \geq 0$ segue do seguinte raciocínio. Defina:

$$\mu = \frac{1}{1 + \alpha_j} \iff 1 + \alpha_j = \frac{1}{\mu} \iff \alpha_j = \frac{1}{\mu} - 1$$

Logo, $d_{j+1} = -\nabla f(y_{j+1}) + \alpha_j d_j = -\nabla f(y_{j+1}) + \left(\frac{1}{\mu} - 1\right) d_j \implies d_{j+1} = -\frac{\mu}{\mu} \nabla f(y_{j+1}) + \frac{1}{\mu} (1 - \mu) d_j \implies d_{j+1} = \frac{1}{\mu} [\mu(-\nabla f(y_{j+1})) + (1 - \mu) d_j]$. Dessa forma, d_{j+1} pode ser visualizada como uma combinação convexa entre a direção do simétrico do gradiente no ponto e a direção usada na última iteração.

Agora, suponha que f seja uma função quadrática possuindo hessiana positiva definida H e colocando como hipótese que d_{j+1} e d_j sejam H -conjugadas. Além disso, considere

$$p_j = y_{j+1} - y_j$$

e

$$q_j = \nabla f(y_{j+1}) - \nabla f(y_j) = H(y_{j+1} - y_j) = H p_j$$

Ora, $d_{j+1}^T H d_j = 0 \implies \lambda_j (d_{j+1}^T H d_j) = \lambda_j \cdot 0 \implies d_{j+1}^T H (\lambda_j d_j) = d_{j+1}^T H (y_{j+1} - y_j) = d_{j+1}^T H p_j = d_{j+1}^T q_j = 0 \implies d_{j+1}^T q_j = 0$. Consequentemente,

$$\begin{aligned}
d_{j+1} = -\nabla f(y_{j+1}) + \alpha_j d_j &\implies d_{j+1}^T = -\nabla^T f(y_{j+1}) + \alpha_j d_j^T \\
&\implies d_{j+1}^T q_j = [-\nabla^T f(y_{j+1}) + \alpha_j d_j^T] q_j \\
&\implies 0 = -\nabla^T f(y_{j+1}) q_j + \alpha_j d_j^T q_j \\
&\implies \alpha_j = \frac{\nabla^T f(y_{j+1}) q_j}{d_j^T q_j} = \frac{\lambda_j \nabla^T f(y_{j+1}) q_j}{\lambda_j d_j^T q_j} \\
&\implies \alpha_j^{HS} = \frac{\lambda_j \nabla^T f(y_{j+1}) q_j}{p_j^T q_j}
\end{aligned}$$

Sendo α_j^{HS} a caracterização para α_j de Hestenes e Stiefel(1952), sendo esta utilizada mesmo se f for não-quadrática, isto é, quando f assume um comportamento quadrático local.

Caso buscas lineares exatas sejam realizadas, temos que $\nabla^T f(y_{j+1}) d_j = \nabla^T f(y_j) d_{j-1} = 0$. Logo, $d_j^T q_j = d_j^T [\nabla f(y_{j+1}) - \nabla f(y_j)] = d_j^T \nabla f(y_{j+1}) - d_j^T \nabla f(y_j) = -d_j^T \nabla f(y_j) = -[-\nabla f(y_j) + \alpha_{j-1} d_{j-1}]^T \nabla f(y_j) = \nabla^T f(y_j) \nabla f(y_j) - \alpha_{j-1} d_{j-1}^T \nabla f(y_j) = \|\nabla f(y_j)\|^2 - \alpha_{j-1}(0) = \|\nabla f(y_j)\|^2 \implies d_j^T q_j = \|\nabla f(y_j)\|^2$. Aplicando esta última igualdade em $\alpha_j^{HS} = \frac{\lambda_j \nabla^T f(y_{j+1}) q_j}{p_j^T q_j}$, segue que $\alpha_j = \frac{\nabla^T f(y_{j+1}) q_j}{d_j^T q_j} = \frac{\nabla^T f(y_{j+1}) q_j}{\|\nabla f(y_j)\|^2}$. Portanto, a caracterização

$$\alpha^{PR} = \frac{\nabla^T f(y_{j+1}) q_j}{\|\nabla f(y_j)\|^2}$$

é a escolha de Polak e Ribiere(1969) para o parâmetro α_j .

Ademais, com a hipótese de f de ser quadrática e que buscas lineares são feitas, temos que $\nabla^T f(y_{j+1}) \nabla f(y_j) = \nabla^T f(y_{j+1}) [\alpha_{j-1} d_{j-1} - d_j] = \nabla^T f(y_{j+1}) \alpha_{j-1} d_{j-1} - \nabla^T f(y_{j+1}) d_j = \alpha_{j-1} \nabla^T f(y_{j+1}) d_{j-1} - 0 = \alpha_{j-1} [\nabla f(y_j) + \lambda_j H d_j]^T d_{j-1} = \alpha_{j-1} \nabla^T f(y_j) d_{j-1} + \alpha_{j-1} \lambda_j d_j^T H d_{j-1} = \alpha_{j-1}(0) + \alpha_{j-1} \lambda_j(0) = 0 + 0 = 0 \implies \nabla^T f(y_{j+1}) \nabla f(y_j) = 0$, visto que d_j e d_{j-1} são H -conjugadas e $\nabla^T f(y_{j+1}) d_j = \nabla^T f(y_j) d_{j-1} = 0$ pelo Teorema 2.6 novamente. Dessa forma, aplique na caracterização $\alpha_j = \alpha^{PR}$ as expressões de p_j e q_j dadas acima e o fato de que $\nabla^T f(y_{j+1}) d_j = \nabla^T f(y_j) d_{j-1} = 0$. Logo, $\alpha_j = \frac{\nabla^T f(y_{j+1}) q_j}{\|\nabla f(y_j)\|^2} = \frac{\nabla^T f(y_{j+1}) [\nabla f(y_{j+1}) - \nabla f(y_j)]}{\|\nabla f(y_j)\|^2} = \frac{\nabla^T f(y_{j+1}) \nabla f(y_{j+1}) - \nabla^T f(y_{j+1}) \nabla f(y_j)}{\|\nabla f(y_j)\|^2} = \frac{\|\nabla f(y_{j+1})\|^2}{\|\nabla f(y_j)\|^2}$
 $\implies \alpha_j^{FR} = \frac{\|\nabla f(y_{j+1})\|^2}{\|\nabla f(y_j)\|^2}$ é a escolha do parâmetro α_j feita por Fletcher e Reeves(1964).

Após a exposição a respeito da caracterização do parâmetro α_j , uma análise do Método dos Gradientes Conjugados usando a escolha de Fletcher e Reeves será feita. Porém, é importante ressaltar que semelhante estudo poderia ser feito para as outras escolhas.

2.4.2 Algoritmo(Método dos Gradientes Conjugados)

Inicialização do algoritmo: Seja $\epsilon > 0$. Escolha um ponto inicial x_1 . Considere $y_1 = x_1$, $d_1 = -\nabla f(y_1)$ $k = j = 1$ e vá para a Etapa Principal.

Etapa Principal:

Passo 1: Se $\|\nabla f(y_j)\| < \epsilon$, pare. Caso contrário, seja $d_j = -\nabla f(y_j)$ e considere λ_j uma solução ótima para minimizar $f(y_j + \lambda d_j)$ sujeito a $\lambda \geq 0$. Seja $y_{j+1} = y_j + \lambda_j d_j$. Se $j < n$, vá para o Passo 2. Se $j = n$, então vá ao passo 2. Caso contrário, vá ao passo 3.

Passo 2: Considere $d_{j+1} = -\nabla f(y_{j+1}) + \alpha d_j$, onde $\alpha_j = \frac{\|\nabla f(y_{j+1})\|^2}{\|\nabla f(y_j)\|^2}$. Substitua j por $j + 1$ e vá ao passo 1.

Passo 3: Tome $y_1 = x_{k+1} = y_{n+1}$ e $d_1 = -\nabla f(y_1)$. Considere $j = 1$, substitua k por $k + 1$ e vá ao passo 1.

2.4.3 Caso Quadrático

Se f é quadrática, o teorema a seguir mostra que as direções geradas pelo Método dos Gradientes Conjugados são H -conjugadas. Por consequência, com o uso do Teorema 2.6, é possível afirmar que o método produz uma solução ótima após uma completa realização do Passo Principal, ou seja, após no máximo n buscas lineares serem feitas.

Teorema 2.8. Considere o problema de minimizar $f(x) = c^T x + (\frac{1}{2})x^T H x$ sujeita a $x \in \mathbb{R}^n$ e H simétrica. Suponha que o problema seja resolvido pelo método dos gradientes conjugados, começando com y_1 e tomando $d_1 = -\nabla f(y_1)$. Em particular, para $j = 1, \dots, n$, seja λ_j uma solução ótima para minimizar $f(y_j + \lambda_j d_j)$ sujeita a $\lambda \geq 0$. Seja $y_{j+1} = y_j + \lambda_j d_j$

e $d_{j+1} = -\nabla f(y_{j+1}) + \alpha_j d_j$, onde $\alpha_j = \frac{\|\nabla f(y_{j+1})\|^2}{\|\nabla f(y_j)\|^2}$. Se $\nabla f(y_j) \neq 0$, então as seguintes sentenças são válidas:

i) d_1, d_2, \dots, d_n são H -conjugados.

ii) d_1, d_2, \dots, d_n são direções de descida.

$$\text{iii) } \alpha_j = \frac{\|\nabla f(y_{j+1})\|^2}{\|\nabla f(y_j)\|^2} = \frac{d_j^T H \nabla f(y_{j+1})}{d_j^T H d_j}$$

Demonstração. Será feita um processo de indução em j . Para $j = 1$, temos que d_1 é de descida, pois $d_1 = -\nabla f(y_1)$, confirmando assim ii). Assumindo a parte iii) como válida para $j = 1$, temos que d_1 e d_2 serão H -conjugados, bastando calcular $d_1^T H d_2$ utilizando $d_2 = -\nabla f(y_2) + \alpha_1 d_1$ e verificar que o valor é nulo, confirmando a validade da parte i). Por outro lado, assumindo a parte i) como válida para $j = 1$, verifica-se por meio de metodologia análoga

a que será exposta abaixo que a parte *iii*) também é verdadeira para $j = 1$. Agora, suponha que a validade das propriedades *i*), *ii*) e *iii*) para j . Mostraremos que elas valem para $j + 1$.

Primeiramente, para mostrar a validade da parte *i*) para $j + 1$, iremos demonstrar que $d_k^T H d_{j+1} = 0$ para $k \leq j$. Com efeito, $d_{j+1} = -\nabla f(y_{j+1}) + \alpha_j d_j$. Ora, tome a hipótese de indução da parte *iii*) e aplique em $d_j^T H d_{j+1}$. Logo,

$$\begin{aligned} d_j^T H d_{j+1} &= d_j^T H [-\nabla f(y_{j+1}) + \alpha_j d_j] \\ &= d_j^T H \left[-\nabla f(y_{j+1}) + \frac{d_j^T H \nabla f(y_{j+1})}{d_j^T H d_j} d_j \right] \\ &= -d_j^T H \nabla f(y_{j+1}) + d_j^T H \nabla f(y_{j+1}) = 0 \end{aligned}$$

Afirmando que a parte *i*) é válida para $k = j$. Seja $k < j$, recorde que a hipótese de indução da parte *i*) garante que $d_k^T H d_j = 0$. Assim:

$$\begin{aligned} d_k^T H d_{j+1} &= d_k^T H [-\nabla f(y_{j+1}) + \alpha_j d_j] \\ &= -d_k^T H \nabla f(y_{j+1}) + \alpha_j d_k^T H d_j \\ &= -d_k^T H \nabla f(y_{j+1}) \\ &\implies d_k^T H d_{j+1} = -d_k^T H \nabla f(y_{j+1}) \end{aligned}$$

Já que $f(x) = c^T x + (\frac{1}{2})x^T H x \implies \nabla f(x) = Hx + c \implies \nabla f(y_{k+1}) - \nabla f(y_k) = H y_{k+1} + c - H y_k - c = H(y_{k+1} - y_k) = \lambda_k H d_k \implies \nabla f(y_{k+1}) = \nabla f(y_k) + \lambda_k H d_k$, pois $y_{k+1} = y_k + \lambda_k d_k$. Logo, $d_{k+1} = -\nabla f(y_{k+1}) + \alpha_k d_k = -[\nabla f(y_k) + \lambda_k H d_k] + \alpha_k d_k = -[-d_k + \alpha_{k-1} d_{k-1} + \lambda_k H d_k] + \alpha_k d_k$, visto que $\nabla f(y_k) = -d_k + \alpha_{k-1} d_{k-1}$. Dessa maneira,

$$d_{k+1} = -[-d_k + \alpha_{k-1} d_{k-1} + \lambda_k H d_k] + \alpha_k d_k$$

Pela hipótese de indução da parte *ii*), temos que d_k é direção de descida, então $\lambda_k > 0$. Portanto, podemos isolar o termo $H d_k$ da expressão anterior.

$$\begin{aligned} H d_k &= \frac{1}{\lambda_k} [-d_{k+1} + (1 + \alpha_k) d_k - \alpha_{k-1} d_{k-1}] \\ \implies (H d_k)^T &= \frac{1}{\lambda_k} [-d_{k+1}^T + (1 + \alpha_k) d_k^T - \alpha_{k-1} d_{k-1}^T]^T \\ \implies d_k^T H &= \frac{1}{\lambda_k} [-d_{k+1}^T + (1 + \alpha_k) d_k^T - \alpha_{k-1} d_{k-1}^T] (*) \end{aligned}$$

Utilizando em $d_k^T H d_{j+1} = -d_k^T H \nabla f(y_{j+1})$, segue que:

$$\begin{aligned}
d_k^T H d_{j+1} &= -d_k^T H \nabla f(y_{j+1}) \\
&= -(d_k^T H) \nabla f(y_{j+1}) \\
&= -\frac{1}{\lambda_k} [-d_{k+1}^T \nabla f(y_{j+1}) + (1 + \alpha_k) d_k^T \nabla f(y_{j+1}) - \alpha_{k-1} d_{k-1}^T \nabla f(y_{j+1})]
\end{aligned}$$

Da hipótese de indução, temos que d_1, \dots, d_j são H -conjugados. Dessa forma, usando a parte *i*) do Teorema 2.6, obtemos que para $j < n$ e $k < j$, então $\nabla^T f(y_{j+1}) d_k = 0$ com $k = 1, \dots, j$. Consequentemente, $d_k^T \nabla f(y_{j+1}) = d_{k-1}^T \nabla f(y_{j+1}) = d_{k+1}^T \nabla f(y_{j+1}) = 0$, visto que $k + 1 \leq j \iff k \leq j - 1$, o que é verdade, pois $k < j$. Portanto, analisando as parcelas de $d_k^T H d_{j+1}$, decorre que $d_k^T H d_{j+1} = 0$.

Por conseguinte, para concluir que d_1, d_2, \dots, d_{j+1} é suficiente mostrar que estes vetores formam um conjunto linearmente independente. De fato, suponha $\sum_{i=1}^{j+1} \gamma_i d_i = 0 \implies \sum_{i=1}^j \gamma_i d_i + \gamma_{j+1} [-\nabla f(y_{j+1}) + \alpha_j d_j] = 0 \implies \nabla^T f(y_{j+1}) \left[\sum_{i=1}^j \gamma_i d_i + \gamma_{j+1} [-\nabla f(y_{j+1}) + \alpha_j d_j] \right] = 0 \implies \sum_{i=1}^j \nabla^T f(y_{j+1}) \gamma_i d_i - \gamma_{j+1} \nabla^T f(y_{j+1}) \nabla f(y_{j+1}) + \alpha_j \nabla^T f(y_{j+1}) d_j = 0$. Recorde que $\nabla^T f(y_{j+1}) d_k = 0$ para $k = 1, \dots, j$ pela parte *i*) do Teorema 2.6, logo $0 - \gamma_{j+1} \|\nabla f(y_{j+1})\|^2 + 0 = 0 \implies \gamma_{j+1} \|\nabla f(y_{j+1})\|^2 = 0 \implies \gamma_{j+1} = 0$, visto que $\|\nabla f(y_{j+1})\| \neq 0$ por hipótese. Lembre que d_1, d_2, \dots, d_j são linearmente independentes (hipótese de indução). Sendo assim, $\sum_{i=1}^{j+1} \gamma_i d_i = 0 \implies \gamma_1 = \gamma_2 = \dots = \gamma_j = \gamma_{j+1} = 0$, logo $d_1, d_2, \dots, d_j, d_{j+1}$ são linearmente independentes.

Para a parte *ii*), perceba inicialmente que $\nabla f(y_{j+1}) \neq 0 \implies \nabla^T f(y_{j+1}) d_j = 0$ para $k = 1, \dots, j$. Decorre que $\nabla^T f(y_{j+1}) d_{j+1} = \nabla^T f(y_{j+1}) [-\nabla f(y_{j+1}) + \alpha_j d_j] = -\|\nabla^T f(y_{j+1})\|^2 + \alpha_j [\nabla^T f(y_{j+1}) d_j] = -\|\nabla^T f(y_{j+1})\|^2 < 0 \implies d_{j+1}$ é direção de descida.

Para a parte *iii*), veja que da hipótese de indução $\alpha_j = \frac{\|\nabla f(y_{j+1})\|^2}{\|\nabla f(y_j)\|^2} = \frac{d_j^T H \nabla f(y_{j+1})}{d_j^T H d_j}$. Na equação (*), ponha $k = j + 1$, logo:

$$\begin{aligned}
d_{j+1} H &= \frac{1}{\lambda_{j+1}} [-d_{j+2} + (1 + \alpha_{j+1}) d_{j+1}^T - \alpha_j d_j^T] \\
\implies \lambda_{j+1} d_{j+1} H &= [-d_{j+2} + (1 + \alpha_{j+1}) d_{j+1}^T - \alpha_j d_j^T] \\
\implies [\lambda_{j+1} d_{j+1} H] \nabla f(y_{j+2}) &= [-d_{j+2} + (1 + \alpha_{j+1}) d_{j+1}^T - \alpha_j d_j^T] \nabla f(y_{j+2}) \\
\implies [\lambda_{j+1} d_{j+1} H] \nabla f(y_{j+2}) &= [(-d_{j+2}^T + \alpha_{j+1} d_{j+1}^T) + d_{j+1}^T - \alpha_j d_j^T] \nabla f(y_{j+2}) \\
\implies [\lambda_{j+1} d_{j+1} H] \nabla f(y_{j+2}) &= [\nabla^T f(y_{j+2}) + d_{j+1}^T - \alpha_j d_j^T] \nabla f(y_{j+2}) \\
\implies \lambda_{j+1} d_{j+1} H \nabla f(y_{j+2}) &= \|\nabla f(y_{j+2})\|^2 + d_{j+1}^T \nabla f(y_{j+2}) - \alpha_j d_j^T \nabla f(y_{j+2})
\end{aligned}$$

Pela parte *i*) do Teorema 2.6, tem-se que $\alpha_j d_j^T \nabla f(y_{j+2}) = d_{j+1}^T \nabla f(y_{j+2}) = 0$, ou seja:

$$\lambda_{j+1} d_{j+1} H \nabla f(y_{j+2}) = \|\nabla f(y_{j+2})\|^2 (**)$$

Das deduções anteriores, vimos que $\nabla f(y_{j+1}) = \nabla f(y_{j+2}) - \lambda_{j+1} H d_{j+1}$, então:

$$\begin{aligned} \nabla^T f(y_{j+1}) \nabla f(y_{j+1}) &= \nabla^T f(y_{j+1}) [\nabla f(y_{j+2}) - \lambda_{j+1} H d_{j+1}] \\ &= -d_{j+1}^T \nabla f(y_{j+2}) + \lambda_{j+1} d_{j+1}^T H d_{j+1} + \alpha_j d_j^T \nabla f(y_{j+2}) - \alpha_j \lambda_{j+1} d_j^T H d_{j+1} \\ &= 0 + \lambda_{j+1} d_{j+1}^T H d_{j+1} + 0 - 0 \\ &= \lambda_{j+1} d_{j+1}^T H d_{j+1} \\ &\implies \|\nabla f(y_{j+1})\|^2 = \lambda_{j+1} d_{j+1}^T H d_{j+1} (***) \end{aligned}$$

Portanto,

$$\frac{(**)}{(***)} = \frac{\|\nabla f(y_{j+2})\|^2}{\|\nabla f(y_{j+1})\|^2} = \frac{\lambda_{j+1} d_{j+1} H \nabla f(y_{j+2})}{\lambda_{j+1} d_{j+1}^T H d_{j+1}} = \frac{d_{j+1} H \nabla f(y_{j+2})}{d_{j+1}^T H d_{j+1}}$$

Confirmando a parte *iii*) para $j + 1$.

□

É importante ressaltar que embora o uso de $\alpha_j = \alpha_j^{FR}$ tenha sido feito na demonstração do teorema anterior, note que como f é quadrática e buscas lineares exatas estão sendo realizadas, tem-se que a utilização de α_j^{HS} ou α_j^{PR} também seria válida. Porém, a escolha α_j^{PR} mostrou-se empiricamente superior à α_j^{FR} (BAZARAA, 2006). O que é plausível, pois para a obtenção de α_j^{FR} foi imposta a condição de f quadrática, enquanto α_j^{PR} não. Na mesma linha de raciocínio, temos que α_j^{HS} é preferível quando buscas lineares inexatas são feitas, visto que a H -conjugação apenas será mantida entre direções consecutivas, entretanto, é notável que o resultado anterior apoiou-se demasiadamente no Teorema 2.6, onde a hipótese inicial da propriedade é haver sempre a presença de um passo ótimo a cada iteração.

2.4.4 Métodos Quase-Newton Sem Memória

Há uma relação interessante entre o Método dos Gradientes Conjugados e uma versão simplificada do BFGS. De fato, suponha que a aproximação da inversa da hessiana esteja de acordo com $D_{j+1} = D_j + C_j^{BFGS}$. Ademais, tome $D_j \equiv \mathcal{I}$ e use a correção do BFGS, isto é,

$$C_j^{BFGS} = \frac{p_j p_j^T}{p_j^T q_j} \left(1 + \frac{q_j^T q_j}{p_j^T q_j} - \frac{q_j p_j^T + p_j q_j^T}{p_j^T q_j} \right). \text{ Portanto:}$$

$$D_{j+1} = \mathcal{I} + \frac{p_j p_j^T}{p_j^T q_j} \left(1 + \frac{q_j^T q_j}{p_j^T q_j} \right) - \frac{q_j p_j^T + p_j q_j^T}{p_j^T q_j}$$

Agora, mova ao longo da direção $d_{j+1} = -D_{j+1}\nabla f(y_{j+1})$. De uma certa maneira, há um “esquecimento” ou “desuso” da aproximação D_j e, ao invés dela, atualiza-se a matriz identidade da mesma maneira como foi feito na primeira iteração. Por essa razão, denomina-se esta metodologia como um “Método Quase-Newton Sem Memória”.

Observe que os requisitos de armazenagem são semelhantes aos do Método dos Gradientes Conjugados (3 vetores n -dimensionais). Além disso, buscas lineares inexatas podem ser feitas contanto que $p_j^T q_j = \lambda_j d_j^T [\nabla f(y_{j+1}) - \nabla f(y_j)]$ permaneça positivo e d_{j+1} seja direção de descida. Além disso, a preocupação quanto à perda de positividade definida para D_j não é mais necessária, visto que D_j é a própria identidade. Segundo (BAZARAA, 2006), tal esquema é bastante eficiente com buscas lineares inexatas.

Caso buscas lineares exatas sejam feitas, recordando que $p_j = \lambda_j d_j = y_{j+1} - y_j \implies p_j^T \nabla f(y_{j+1}) = \lambda_j (d_j^T \nabla f(y_{j+1})) = 0$. Aplicando em $d_{j+1} = -D_{j+1}\nabla f(y_{j+1})$, segue que:

$$\begin{aligned} d_{j+1} &= -D_{j+1}\nabla f(y_{j+1}) \\ &= -\left[\mathcal{I} + \frac{p_j p_j^T}{p_j^T q_j} \left(1 + \frac{q_j^T q_j}{p_j^t q_j} \right) - \frac{q_j p_j^T + p_j q_j^T}{p_j^T q_j} \right] \nabla f(y_{j+1}) \\ &= -\nabla f(y_{j+1}) - \frac{p_j p_j^T \nabla f(y_{j+1})}{p_j^T q_j} - \frac{p_j p_j^T}{p_j^T q_j} \left(\frac{q_j^T q_j}{p_j^T q_j} \right) \nabla f(y_{j+1}) + \frac{q_j p_j^T \nabla f(y_{j+1}) + p_j q_j^T \nabla f(y_{j+1})}{p_j^t q_j} \\ &= -\nabla f(y_{j+1}) + \left(\frac{q_j^T \nabla f(y_{j+1})}{p_j^T q_j} \right) p_j = -\nabla f(y_{j+1}) + \left(\frac{q_j^T \nabla f(y_{j+1})}{p_j^T q_j} \right) (\lambda_j d_j) \\ &\implies d_{j+1} = -\nabla f(y_{j+1}) + \alpha_j^{HS} d_j \end{aligned}$$

Pois lembre que $\alpha_j^{HS} = \frac{\lambda_j \nabla^T f(y_{j+1}) q_j}{p_j^t q_j}$. Logo, o esquema de atualização do BFGS sem memória é equivalente ao Método dos Gradientes Conjugados de Hestenes e Stiefel quando buscas lineares são exercidas. Por último, apesar de o esquema sem memória ser realizado com qualquer outro membro da Família Broyden, a equivalência com o Método dos Gradientes Conjugados ocorrerá apenas para $\phi = 1$ (BFGS).

2.4.5 Convergência dos Algoritmos com Direções Conjugadas

Como mostra o Teorema 2.6, se a função considerada for quadrática, então qualquer direção conjugada produz solução ótima num número finito de passos. Nesta seção, discutiremos esses métodos quando a função não for quadrática.

2.4.5.1 Composição de Aplicações

Em uma grande parte dos problemas de Programação Não-Linear, a função que representa a execução do algoritmo é resultado de uma composição de funções. Por exemplo, alguns

algoritmos primeiramente encontram uma direção d_k e calculam o tamanho do passo λ_k resolvendo um problema de minimizar $\alpha(x_k + \lambda_k d_k)$. Dessa forma, uma aplicação composta A é dada por $A = MD = M \circ D$, onde D busca a direção d_k e M encontra o passo ótimo λ_k . Entretanto, tais aplicações serão funções ponto-conjunto. Formalmente, temos:

Definição 14. (*Aplicação Ponto-Conjunto*) Sejam X e Y conjuntos não-vazios. $B : X \rightarrow Y$ é dita aplicação ponto-conjunto se associa cada ponto $x \in X$ a um subconjunto $B(x) \subset Y$.

Definição 15. (*Aplicação Composta*) Sejam X, Y e Z conjuntos não-vazios com $X \in \mathbb{R}^n$, $Y \subset \mathbb{R}^p$ e $Z \subset \mathbb{R}^q$. Considere $B : X \rightarrow Y$ e $C : Y \rightarrow Z$ aplicações ponto-conjunto. A composição $A = CB = C \circ B$ é dita aplicação composta $A : X \rightarrow Z$ com $A(x) = \bigcup \{C(y); y \in B(x)\}$.

Teorema 2.9. Sejam $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^p$, $Z \subset \mathbb{R}^q$ conjuntos não-vazios e fechados em \mathbb{R}^n . Considere $B : X \rightarrow Y$ e $C : Y \rightarrow Z$ aplicações ponto-conjunto e $A = CB$ a aplicação composta. Suponha $B(x)$ e $C(B(X))$ conjuntos fechados. Suponha a validade da seguinte propriedade: se $x_k \rightarrow x$, $y_k \in B(x_k)$, então existe uma subsequência convergente $\{y_k\}$. Nessas condições, $A(x)$ é fechado.

Demonstração. Seja $x_k \rightarrow x$, $z_k \in A(x_k)$ e $z_k \rightarrow z$. É preciso mostrar que $z \in A(x)$ para que tenhamos $A(x)$ fechado. Com efeito, pela definição de $A(A(x) = \bigcup \{C(y); y \in B(x)\})$, para cada k existe $y_k \in B(x_k)$ tal que $z_k \in C(y_k)$. Ora, da hipótese, existe uma subsequência $\{y_k\}$ com limite y , $y_k \in B(x_k)$. Como $B(x)$ é fechado, então $y \in B(x)$. Além disso, já que $C(B(x))$ é fechado e $y \in B(x)$, então C é fechado em y , o que implica que $z \in C(y)$. Portanto, $z \in C(y)$ e $C(y) \in C(B(x))$, isto é, $z \in C(B(x)) = A(x)$.

□

Teorema 2.10. (*Convergência de Algoritmos Representados por Aplicações Compostas*) Considere $X \subset \mathbb{R}^n$ um conjunto não-vazio e fechado. Seja $\Omega \subseteq X$ um conjunto solução não-vazio. Ademais, defina $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função contínua e considere a aplicação ponto-conjunto $C : X \rightarrow X$. satisfazendo a seguinte propriedade: Dado $x \in X \implies \alpha(y) \leq \alpha(x)$ para $y \in C(x)$. Seja $B : X \rightarrow X$ uma aplicação ponto-conjunto fechada em $X \setminus \Omega$ que satisfaz $\alpha(y) < \alpha(x)$ para cada $y \in B(x)$ se $x \notin \Omega$. Agora, considere o algoritmo definido pela aplicação composta $A = CB$. Dado $x_1 \in X$, suponha que a sequência $\{x_k\}$ é gerada da seguinte forma: Se $x_k \in \Omega$, pare. Caso contrário, seja $x_{k+1} \in A(x_k)$, substitua k por $k + 1$ e repita o processo. Suponha que $\Lambda = \{x; \alpha(x) \leq \alpha(x_1)\}$ seja compacto. Então, o algoritmo para em um número finito de passos com Ω ou todo ponto de acumulação de $\{x_k\}$ pertence ao conjunto Ω .

Demonstração. Se em qualquer k -ésima iteração $x_k \in \Omega$, então pare. Suponha que a sequência x_k é gerada pelo algoritmo e considere $\{x_k\}_{\mathcal{K}}$ uma subsequência convergente com limite x . Logo, $\alpha(x_k) \rightarrow \alpha(x)$ para $k \in \mathcal{K}$. Pela continuidade de α , segue que:

$$\lim_{k \rightarrow \infty} \alpha(x_k) = \alpha(x)(*)$$

Afirma-se que $x \in \Omega$. Com efeito, suponha por absurdo que $x \notin \Omega$ e considere a sequência $\{x_{k+1}\}_{\mathcal{K}}$. Pela definição da aplicação composta $A(x)$, sabe-se que $x_{k+1} \in A(x_k)$, isto é, $x_{k+1} \in A(x_k) \implies x_{k+1} \in (CB)(x_k)$, onde $x_{k+1} \in C(y_k) \implies \alpha(x_{k+1}) \leq \alpha(y_k)$ e $y_k \in B(x_k) \implies \alpha(y_k) < \alpha(x_k)$ já que $x_k \notin \Omega$, pois caso contrário, o algoritmo iria parar em x_k . Note que $\alpha(x_{k+1}) \leq \alpha(x_1)$ e como $\alpha(y_k) < \alpha(x_k) \leq \alpha(x_1)$, segue que $x_{k+1}, y_k \in \Lambda$. Por hipótese, Λ é compacto. Assim, toda sequência de pontos Λ possui uma subsequência convergente com limite sendo algum ponto de Λ . Tome as sequências $\{x_{k+1}\}_{\mathcal{K}}$ e $\{y_k\}_{\mathcal{K}}$ em Λ . Consequentemente, $\exists \mathcal{K}' \subseteq \mathcal{K}$ tal que $y_k \rightarrow y$ e $x_{k+1} \rightarrow x'$ para $k \in \mathcal{K}'$. Do fato de que $y_k \rightarrow y$ e $y_k \in B(x_k) \subset B(X \setminus \Omega)$, então o valor $y \in B(X \setminus \Omega)$, pois tal conjunto é fechado por hipótese. Veja que $\alpha(y_k) < \alpha(x_k) \implies \alpha(y) < \alpha(x)$, pois $\alpha \in C^1$. Portanto, das condições expostas no enunciado para aplicação $B, y \in B(x)$. Além disso, lembre que $\alpha(x_{k+1}) \leq \alpha(y_k)$. Se $k \rightarrow \infty$, então $\alpha(x') \leq \alpha(y)$. Ora, decorre que $\alpha(y) < \alpha(x) < \alpha(x')$, o que contradiz (*). \square

2.4.5.2 Convergência do Método dos Gradientes Conjugados

No teorema 3.10 foi mostrado que a aplicação composta que representa o algoritmo $A = CB$ converge para um ponto no conjunto solução Ω se as seguintes propriedades são válidas:

1. $B(x)$ é fechado em $X \setminus \Omega$.
2. Se $y \in B(x) \implies f(y) < f(x)$ para $x \in X \setminus \Omega$.
3. Se $z \in C(y) \implies f(z) \leq f(y)$.
4. $\Lambda = \{x : f(x) \leq f(x_1)\}$ é compacto, onde x_1 é o ponto inicial do algoritmo.

Para uma direção conjugada (Método Quase-Newton ou Gradiente Conjugado), a função B é da seguinte forma: Dado x , então $y \in B(x)$ significa afirmar que y é obtido por meio de uma minimização de f iniciando de x e indo ao longo da direção $d = -D\nabla f(x)$, sendo D positiva definida. Se o método é o dos gradientes conjugados, então $D = \mathcal{I}$. Caso seja um método Quase-Newton, então D é definida positiva e arbitrária. Após a obtenção do ponto através da aplicação B , a aplicação C é caracterizada por minimizar a função f ao longo das direções particulares inerentes à cada método. Dessa maneira, a função C cumpre o item 3.

Defina $\Omega = \{x; \nabla f(x) = 0\}$. Prova-se que a função B satisfaz os requisitos 1 e 2. Com efeito, seja $x \in \Omega$, $x_k \rightarrow x$, $y_k \in B(x_k)$ e $y_k \rightarrow y$. É preciso mostrar que $y \in B(x)$. Ora, pela definição de y_k , tem-se que $y_k = x_k - \lambda_k D\nabla f(x_k)$ para $\lambda_k \geq 0$ tal que $f(y_k) \leq f(x_k - \lambda D\nabla f(x_k))$, $\forall \lambda \geq 0$. Devido ao fato de que $\nabla f(x) \neq 0 \implies y_k - x_k = -\lambda_k D\nabla f(x_k) \implies \|y_k - x_k\| = \|-\lambda_k D\nabla f(x_k)\| \implies \lambda_k = \frac{\|y_k - x_k\|}{\|D\nabla f(x_k)\|}$. Fazendo $k \rightarrow \infty$,

segue que $\bar{\lambda} = \frac{\|y - x\|}{\|D\nabla f(x)\|} \geq 0$. Por consequência, $y = x - \bar{\lambda}D\nabla f(x)$. Tomando a desigualdade (*) e realizando $k \rightarrow \infty$, decorre que $f(y) \leq f(x - \lambda D\nabla f(x))$, $\forall \lambda \geq 0$. Portanto, y é realmente obtido pela minimização da f começando por x e indo ao longo da direção $-D\nabla f(x)$. Ou seja, $y \in B(x)$ e B é fechado. Ademais, o requisito 2 é satisfeito, pois $-\nabla f(x)D\nabla f(x) < 0 (D > 0)$, o que implica que $-D\nabla f(x)$ é direção de descida. Por fim, assumindo que o conjunto na parte 4 é compacto, segue que os métodos de direções conjugadas dos algoritmos discutidos nesta seção convergem a um ponto estacionário.

2.4.5.3 Característica da Razão de Convergência do Método dos Gradientes Conjugados

Considere a função quadrática $f(x) = c^T x + \frac{1}{2}x^T Hx$, $H = H^T$, $H > 0$, logo seus autovalores são reais e estritamente positivos. Suponha que os autovalores de H sejam agrupados em dois conjuntos, um dos conjuntos por m grandes autovalores e talvez distantes entre si, e o outro formado pelos restantes $n-m$ autovalores relativamente pequenos. Assuma que $m + 1 < n$ e denote α como a razão entre o maior e o menor autovalor do último conjunto descrito. Sabemos que a convergência do Método dos Gradientes Conjugados ao ponto ótimo é estabelecida em no máximo n passos (Teorema 2.6). No entanto, suponha que o Método dos Gradientes Conjugados reinicie após $m + 1$ buscas lineares, isto é, ao invés de continuarmos a descida, tome o ponto inicial obtido após $m + 1$ iterações como o novo x_1 e inicie o método novamente. Esta metodologia é chamada de Método dos Gradientes Conjugados Parcial.

Em outras palavras, iniciando com x_1 , considere x_k a sequência gerada pela versão parcial do método, logo para cada $k \geq 1$, o ponto x_{k+1} é obtido aplicando $m + 1$ vezes o Método dos Gradientes Conjugados até que reinicie o método com x_k sendo o ponto inicial. Dessa forma, definindo uma função erro $E(x) = \frac{1}{2}(x - x^*)^T H(x - x^*) = f(x) - \frac{1}{2}x^{*T} Hx^*$, onde é possível mostrar, segundo (LUENBERGER; YE, 2010), que:

$$e(x_{k+1}) \leq \frac{(\alpha - 1)^2}{(\alpha + 1)^2} e(x_k)$$

Consequentemente, isto estabelece uma ordem de convergência linear para o processo acima como no caso já exposto anteriormente para $m = 0$ (Método do Gradiente). Porém, a razão α que controla a taxa de convergência é independente dos m maiores autovalores comentados. Assim, o efeito provocado pelos m maiores autovalores é eliminada. No entanto, há um acréscimo de $m + 1$ passos em comparação ao passo único da metodologia tradicional.

Em seguida, considere o caso geral não-quadrático no qual o usual Método dos Gradientes Conjugados é utilizado. Intuitivamente, já que o gradiente conjugado realiza em n passos o que o Método de Newton realiza em um único passo, pela sua convergência quadrática local, espera-se que semelhantemente o Método dos Gradientes Conjugados venha a convergir quadraticamente. Ou seja, $\|x_{k+1} - x^*\| \leq \beta \|x_k - x^*\|^2$ para algum $\beta > 0$. Com efeito, pode ser demonstrado que se a sequência $\{x_k\} \rightarrow x^*$, sendo f duas vezes diferenciável em alguma

vizinhança de x^* e $\nabla^2 f(x^*) > 0$, então o Método dos Gradientes Conjugados possui convergência superlinear. Além disso, se a Hessiana da função possui a condição de Lipschitz em uma vizinhança de x^* , então a razão de convergência é quadrática.

3. Avaliações Computacionais

O objetivo deste capítulo é avaliar a eficiência dos métodos de minimização irrestrita quanto ao tempo de processamento na resolução de problemas de minimização global. Para tal, selecionaram-se dez funções de classe C^1 com domínio em \mathbb{R}^n , cujo os mínimos globais existem e no qual possuímos prévio conhecimento. Da análise matemática realizada sobre os métodos acima, vemos que o ponto inicial $x_0 \in \mathbb{R}^n$ necessita ser dado para que o algoritmo comece seu processo de iteração.

Conseqüentemente, para cada um dos problemas, o desempenho de cada um dos cinco algoritmos (Gradiente, Newton, DFP, BFGS, Gradientes Conjugados) foram verificados nos dois cenários. O primeiro, para x_0 suficientemente próximo da solução x^* , e o segundo, para x_0 relativamente distante de x^* . Ademais, realizaram-se comparações entre os algoritmos em ambos os cenários por meio do uso de uma ferramenta conhecida como perfil de desempenho. Assim sendo, antes da exposição dos problemas e dos respectivos perfís, é necessária uma breve discussão sobre sua metodologia de criação, bem como sua correta interpretação.

3.1 Perfil de Desempenho

Os perfís de desempenho são uma técnica poderosa na comparação de softwares científicos. A comparação é apresentada de forma compacta, permitindo a rápida interpretação dos testes computacionais envolvendo um conjunto de softwares. Entretanto, a interpretação de um gráfico de perfil desempenho não é imediata. Por essa razão, certas observações iniciais serão feitas.

Seja \mathcal{P} um conjunto de problemas com n_p elementos e \mathcal{S} um conjunto de softwares de n_s elementos, onde o termo “software” refere-se à implementação computacional de determinado algoritmo. Para cada $p \in \mathcal{P}$ e $s \in \mathcal{S}$, define-se $t_{p,s}$ como o tempo computacional do problema p em relação ao software s . Embora o tempo seja aqui utilizado, é importante ressaltar que este critério é apenas um dos tipos de métrica de desempenho.

A razão de desempenho para um software $s_1 \in \mathcal{S}$ na resolução de um problema $p' \in \mathcal{R}$ serve para comparar o desempenho de s_1 em relação ao software de melhor desempenho na resolução de p' , isto é:

$$r_{p',s_1} = \frac{t_{p',s_1}}{\min\{t_{p',s} : \forall s \in \mathcal{S}\}}$$

Logo, note que $r_{p',s_1} \geq 1$. No entanto, há possibilidade de o software s falhar na resolução de um problema p (por exemplo, o tempo de processamento não é finito). Por essa razão, é definido r_M tal que $r_M \geq r_{\bar{p},\bar{s}}(*), \forall \bar{p} \in \mathcal{P}, \forall \bar{s} \in \mathcal{S}$, onde:

$$r_M = 1 + \max\{r_{\bar{p},\bar{s}} ; \bar{p} \in \mathcal{P}, \bar{s} \in \mathcal{S}\}$$

Conforme comentado em (MUNARI, 2009), r_M não influencia na razão de desempenho, contanto que $(*)$ seja satisfeita. Entretanto, o que se deseja é uma avaliação geral do desempenho do software. Para isso, é definido o perfil de desempenho do software, $\rho_s : \mathbb{R} \rightarrow [0, 1]$, dado por:

$$\rho_s(\tau) = \frac{1}{n_p} |\{p \in \mathcal{P} ; r_{p,s} \leq \tau\}|$$

Sendo a notação $|\cdot|$ a representação do número de elementos de um conjunto. Em outras palavras, veja que temos a associação para cada valor $\tau \in \mathbb{R}$ à fração de problemas resolvidos pelo software $s \in \mathcal{S}$ com um desempenho dentro de um fator τ do melhor desempenho obtido. Note que ρ_s é constante por partes, não decrescente e contínua à direita de cada ponto de descontinuidade.

Uma outra interpretação a ser dada ao perfil de desempenho, é que este pode ser encarado como a probabilidade de a razão de desempenho $r_{p,s}$ associada ao software s esteja dentro de um fator τ da melhor razão obtida. Ou seja, encaramos ρ_s como uma função de distribuição acumulada $\rho_s(\tau) = P(r_{p,s} \leq \tau : 1 \leq s \leq n_s)$. Assim, ρ_s é a probabilidade de o software s resolver um problema de \mathcal{P} em não mais do que τ vezes o tempo mínimo levado por qualquer software em \mathcal{S} . Se \mathcal{P} é suficientemente grande e capaz de representar os problemas encontrados na prática, então os softwares com grande probabilidade $\rho_s(\tau)$ são favoritos.

3.1.1 Análise do Perfil de Desempenho

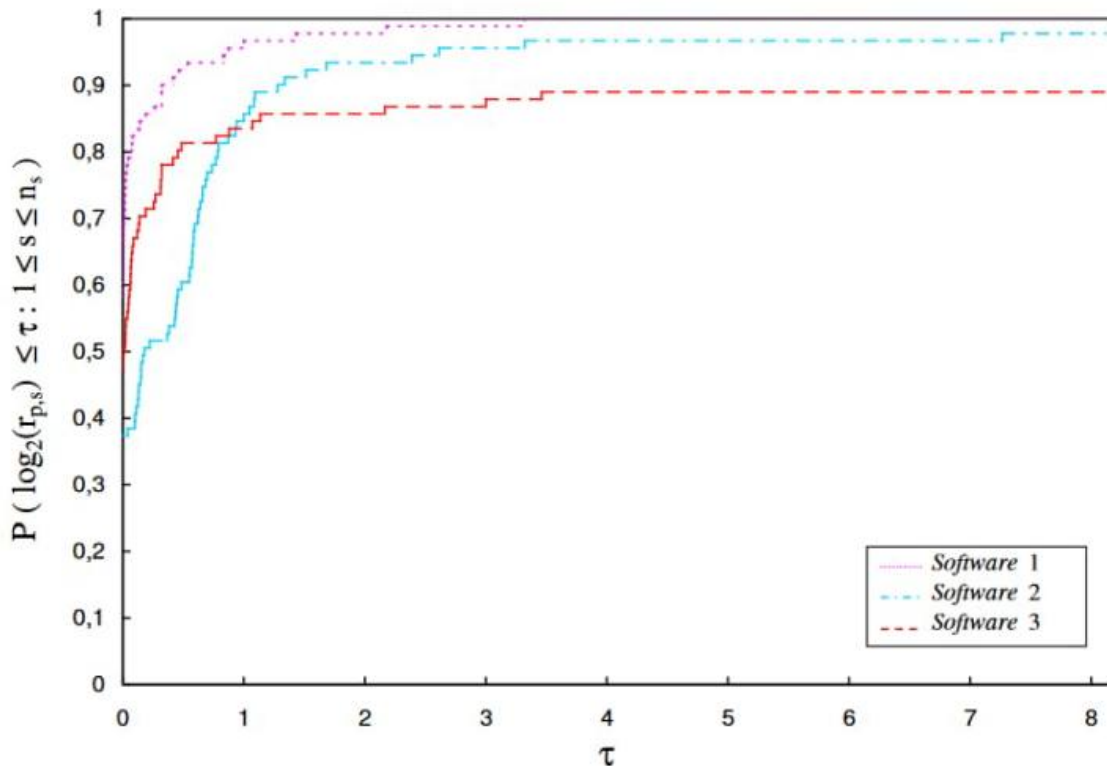
No perfil de desempenho, $\rho_s(1)$ é a fração de problemas para os quais s obteve melhor desempenho que outros softwares. Isto é, $\rho_s(1)$ é a probabilidade que o software ganhe dos demais. Logo, **se o interesse é saber em quantas vezes o software foi melhor que os restantes, analise $\rho_s(1)$.**

Analisando o outro extremo, temos que **se o interesse for em robustez**, ou seja, a capacidade de resolver o problema em si, não necessariamente contabilizando o tempo de processamento, **então é relevante analisar** $\rho_s(r_M) = 1$, já que $r_M \geq r_{\bar{p},\bar{s}}, \forall \bar{p} \in \mathcal{P} \text{ e } \bar{s} \in \mathcal{S}$. Recorde que $r_{p,s} = r_M$ quando o problema p não é resolvido pelo software s . Equivalentemente, a probabilidade que o software s resolva um problema p qualquer é dada por $\rho_s^* = \rho_s(r_M - \epsilon)$ quando $\epsilon \rightarrow 0, \epsilon > 0$. Logo, ρ_s^* é o maior valor assumido por $\rho_s(\tau)$ quando τ está próximo de r_M pela esquerda com $\tau \neq r_M$. Note que ρ_s^* não depende de τ , assim para escolher o software de maior probabilidade de resolver o conjunto de problemas \mathcal{P} , analise ρ_s^* para cada $s \in \mathcal{S}$.

3.1.2 Gráficos de Perfil de Desempenho

Considere $[1, \tau)$. Para cada software, tem-se uma respectiva curva com seu perfil de desempenho conforme a variação de τ . Já que $r_M \gg 1$, recomenda-se uso de escala logarítmica. Em geral, o uso de um logaritmo de base 2 é utilizado na literatura. Ou seja, $\rho_s : [1, r_M) \rightarrow \frac{1}{n_p} |\{p \in \mathcal{P}; \log_2(r_{p,s}) \leq \tau\}|$. Assim, $\tau \in [0, r_M)$ com $r_M = 1 + \max\{\log_2(r_{p,s}); \bar{p} \in \mathcal{P}, \bar{s} \in \mathcal{S}\}$. Ademais, perceba que agora haverá de ter cuidado na interpretação nos extremos dos perfis de desempenho, visto que houve mudança de escala. Segue abaixo um exemplo de gráfico de perfil de desempenho.

Figura 3.1 – Exemplo de Perfil de Desempenho (MUNARI, 2009)



Observe a curva de perfil de desempenho do Software 1. Quanto ao tempo de processamento, analise $\tau = 0$. Veja que o software 1 obteve melhor desempenho em 60 % dos casos.

Além disso, quando τ se aproxima de 3, o valor $\rho_{s_1}(\tau) = 1$, logo o Software 1 resolve todos os problemas.

Embora a curva do Software 1 domine as demais curvas, isto não indica que tal software obteve melhor desempenho que os demais, mas sim, que ele conseguiu resolver mais problemas dentro de um fator τ do desempenho de qualquer outro software.

O Software 2 e o Software 3 não resolvem todos os problemas, pois $\rho_{s_2}(\tau) \neq 1$ e $\rho_{s_3}(\tau) \neq 1, \forall \tau \in [0, r_M)$. Observe que o Software 3 resolveu acima de 95 % dos problemas e o Software 3 resolveu entre 85 e 90 % dos problemas. Ou seja, a probabilidade de sucesso utilizando o Software 3 é acima de 0,95 e a do Software 2 é entre 0,85 e 0,90.

3.1.3 Implementação dos Algoritmos e Especificações

A implementação dos algoritmos relacionados aos métodos de minimização irrestrita foram construídos conforme a teoria vista nos capítulos anteriores, onde a busca linear por Armijo foi utilizada em cada um dos métodos. Ademais, as ferramentas utilizadas foram Python (versão 3.9.12) para a construção dos métodos e obtenção dos respectivos tempos de processamento, além da linguagem Julia (versão 1.7.3) para a construção dos perfis de desempenho. Por fim, a compilação dos scripts foi feita através da ferramenta on-line Jupyter.

3.1.4 Banco de Problemas

Conforme já dito na introdução do capítulo, os dez problemas nada mais são que dez funções $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe C^1 , onde cada uma possui mínimo global. Portanto, descreveremos tais funções e seus respectivos mínimos globais.

Tabela 3.1 – Problemas Seleccionados

Problema	Lei da Função	Mínimo Global
(1)	$g(x, y) = x^2 + 4y^2 - 4x + 8y$	(2, 1)
(2)	$griewank(x, y) = 1 + \frac{1}{4000}(x^2 + y^2) - \cos(x)\cos\left(\frac{y}{\sqrt{2}}\right)$	(0, 0)
(3)	$rosenbrock(x, y) = 100(y - x^2)^2 + (1 - x)^2$	(1, 1)
(4)	$s(x, y) = x^2 + y^2$	(0, 0)
(5)	$tcamel(x, y) = 2x^2 - 1.05x^4 + \frac{x^6}{6} + xy + y^2$	(0, 0)
(6)	$h(x, y) = 20 + (x^2 - 10\cos(2\pi x)) + (y^2 - 10\cos(2\pi y))$	(0, 0)
(7)	$booth(x, y) = (x + 2y - 7)^2 + (2x + y - 5)^2$	(1, 3)
(8)	$matyas(x, y) = 0.26(x^2 + y^2) - 0,48xy$	(0, 0)
(9)	(*)	(0, -1)
(10)	$mccormick(x, y) = \text{sen}x + (x - y)^2 - 1,5x + 2,5y + 1$	(-0.54719, -1.54719)

(*) *goldstein – price*(x, y) =

$$[1 + (x + y + 1)^2(19 - 14x + 3x^2 - 14y + 6xy + 3y^2)][30 + (2x - 3y)^2(18 - 32x + 12x^2 + 48y - 36xy + 27y^2)]$$

3.1.5 Pontos Iniciais Arbitrados

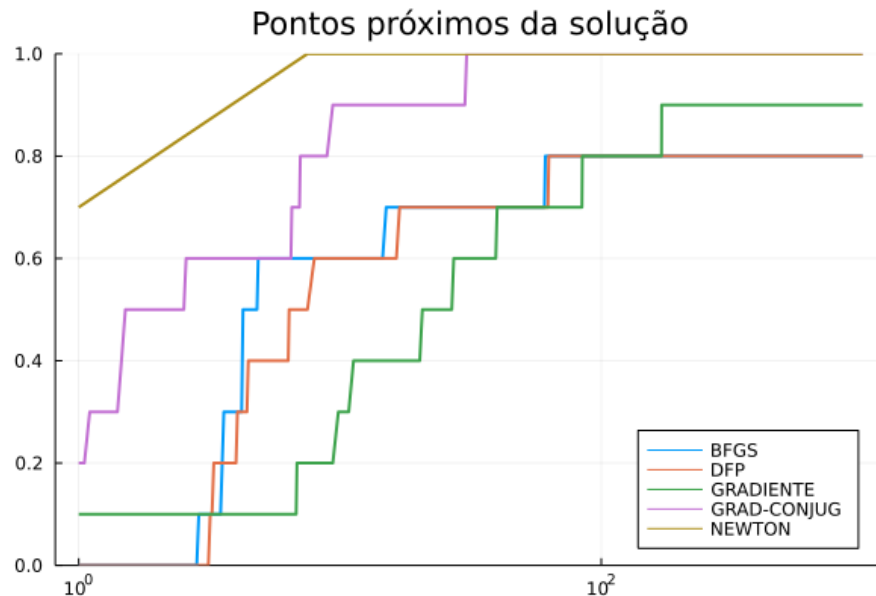
Para as análises feitas a seguir, os pontos próximos(x_0 próximo da solução x^*) e distantes(x_0 distantes da solução x^*) serão, respectivamente, denotados por x_{0_P} e x_{0_D} .

Tabela 3.2 – Pontos iniciais arbitrados

Problema	x_{0_P}	x_{0_D}
(1)	(1.98, 0.97)	(1000, 2000)
(2)	(0.01, 0.002)	(20, 50)
(3)	(0.99, 0.92)	(10, 30)
(4)	(1, 0.5)	(100, 200)
(5)	(0.5, 0.1)	(200, 300)
(6)	(0.01, 0.02)	(150, 200)
(7)	(0.95, 2.5)	(120, 100)
(8)	(1, 3)	(50, 100)
(9)	(0.0001, -0.9999)	(90, 60)
(10)	(-0.54715, -1.90)	(20, 30)

3.1.6 Perfil de Desempenho dos Métodos com Ponto Inicial Próximo da Solução

Figura 3.2 – Perfil de desempenho dos métodos tomando ponto inicial próximo da solução



Primeiramente, considere o conjunto \mathcal{S} de softwares e o conjunto \mathcal{P} de problemas

$\mathcal{S} = \{\text{Método do Gradiente, Método de Newton, DFP, BFGS, Método dos Gradientes Conjugados}\}$

$$\mathcal{P} = \{(i); i = 1, \dots, 10\}$$

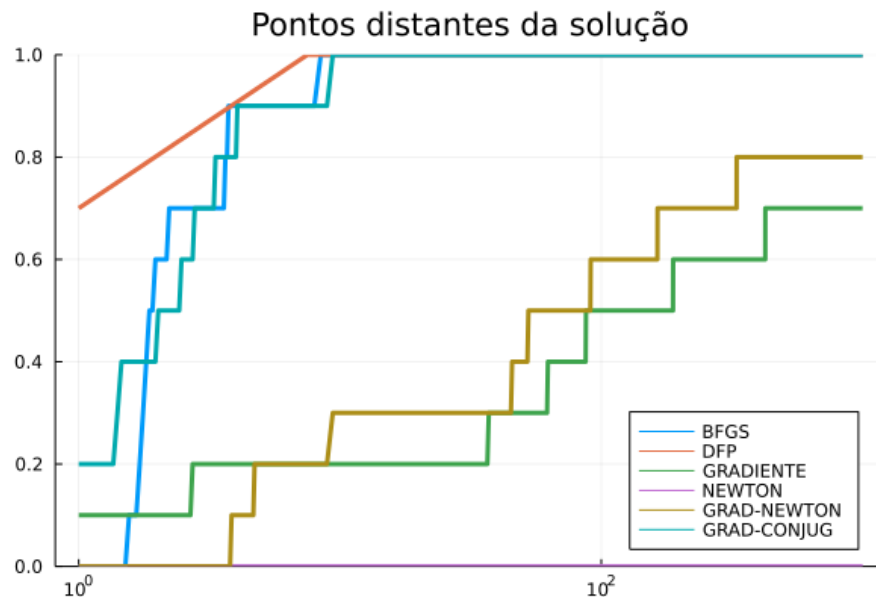
onde (i) refere-se aos problemas exibidos em 3.1.4.

A escala logarítmica decimal foi utilizada em cada curva do gráfico da Figura 2. Em outras palavras, a função relativa ao perfil de desempenho do contexto em questão é dada por $\rho_s : [1, r_M) \rightarrow \frac{1}{n_p} |\{p \in \mathcal{P}; \log_{10}^{(r_p, s)} \leq \tau\}|$. Assim, $\tau \in [0, r_M)$ com $r_M = 1 + \max\{\log_{10}^{(r_p, s)}\}$; $\bar{p} \in \mathcal{P}$, $\bar{s} \in \mathcal{S}$.

Ademais, para $\tau = 1$, verificamos que o Método de Newton foi o mais rápido em 70 % dos problemas, o Método dos Gradientes Conjugados foi o mais rápido em 20 % dos problemas e o Método do Gradiente foi o mais rápido em 10 % dos problemas. Quanto à robustez, analisando $\tau \rightarrow r_M$, o Método de Newton e o Método dos Gradientes Conjugados resolvem todos os problemas. O Método do Gradiente resolve 90 % dos problemas. O BFGS e o DFP resolvem em 80 % dos casos. Em suma, o Método de Newton foi o vencedor da competição, resultado esperado teoricamente, devido à sua convergência quadrática local.

3.1.7 Perfil de Desempenho dos Métodos com Ponto Inicial Distante da Solução

Figura 3.3 – Perfil de desempenho dos métodos tomando ponto inicial distante da solução



Agora, considere o conjunto de softwares:

$\mathcal{S} = \{\text{Método do Gradiente, Método de Newton, DFP, BFGS, M.G.N., M.G.C.}\}$

E o conjunto de problemas:

$$\mathcal{P} = \{(i); i = 1, \dots, 10\}$$

onde (i) refere-se aos problemas exibidos em 3.1.4., M.G.N é o Método Gradiente-Newton e M.G.C é o Método dos Gradientes Conjugados.

A escala logarítmica decimal foi utilizada em cada curva do gráfico da Figura 3. Novamente, a função relativa ao perfil de desempenho de cada método é dada por $\rho_s : [1, r_M) \rightarrow \frac{1}{n_p} |\{p \in \mathcal{P}; \log_{10}^{(r_p, s)} \leq \tau\}|$. Assim, $\tau \in [0, r_M)$ com $r_M = 1 + \max\{\log_{10}^{(r_p, s)}; \bar{p} \in \mathcal{P}, \bar{s} \in \mathcal{S}\}$.

Analisando a rapidez dos métodos, veja que para $\tau = 1$, verificamos que o DFP foi o mais rápido em 70 % dos problemas e o Método do Gradiente foi o mais rápido em 10% dos problemas. Embora a literatura (BAZARAA, 2006), (MARTÍNEZ; AUGUSTA, 1998) e (LUNENBERGER; YE, 2010) indique o BFGS como o de melhor desempenho computacional, isto não invalida a análise acima, visto que o conjunto de problemas analisados é pequeno em comparação ao feito tradicionalmente, isto é, as particularidades das funções do banco de problemas tendeciam ao resultado apresentado. Além disso, comparando ao cenário anterior, percebe-se que neste contexto, o Método de Newton é ineficiente, devido ao chute inicial distante da solução do problema. Ademais, o Método dos Gradientes Conjugados é o mais rápido em 20 % dos problemas, o que é plausível, visto que a dimensão do domínio é pequena, o que proporciona vantagens quanto à rapidez de processamento.

Quanto à robustez, analisando $\tau \rightarrow r_M$, o DFP, o BFGS e o Método dos Gradientes Conjugados resolvem todos os problemas. O Método do Gradiente resolve 70 % dos problemas. Já o Método de Newton sequer converge em tempo viável. Portanto, o DFP é o vencedor da competição, pois ganha por menor tempo de processamento e robustez. Em segundo lugar, teremos o Método dos Gradientes Conjugados. E, por sua vez, o BFGS viria em terceiro lugar, pois embora não haja apresentado rapidez em comparação aos outros, ele possui probabilidade de sucesso é 1.

Por fim, testou-se uma combinação entre os métodos do Gradiente e de Newton (Método Gradiente-Newton) no intuito de verificar se este seria mais eficiente quando comparado aos dois anteriores. A metodologia foi a seguinte: Inicialmente, o chute inicial é computado com o método do Gradiente, pois este apresentou maior eficiência ao de Newton com base no conjunto de problemas \mathcal{P} . Porém, em um dos problemas, ao se chegar em um ponto x_k próximo da solução, o decréscimo do valor da função era irrisório ao se utilizar apenas a descida na direção do gradiente. Neste momento, o algoritmo toma x_k como x_0 e inicia o método de Newton. Dessa maneira, aproveita-se a eficiência local de Newton, o que nos possibilita esperar (hipoteticamente), uma convergência ao ponto desejado. Com efeito, em um dos problemas houve sucesso na estratégia, confirmando sua superioridade do Método Gradiente-Newton \mathcal{P} relação do Método de Newton (nenhuma robustez) e Método do Gradiente (70 % de robustez), possuindo uma robustez de 80 %.

4. Conclusão

Neste trabalho, apresentaram-se resultados relacionados ao embasamento da Programação Não-Linear, bem como desenvolveu-se teoricamente alguns métodos de minimização irrestrita, havendo ao término suas decorrentes implementações computacionais. De fato, houve a exposição de cinco algoritmos distintos que possibilitam minimizar uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^1$, a partir de um ponto dado $x_0 \in \mathbb{R}^n$ com $\nabla f(x_0) \neq 0$. Obviamente, no Método de Newton necessitamos que a hessiana da função seja invertível em x_k a cada k -ésima iteração. Ademais, embora uma comparação tenha sido feita, percebeu-se claramente que não houve um destaque entre os métodos trabalhados. Com efeito, o Método do Gradiente é bastante eficiente no estudo de quadráticas com hessianas que possuem autovalores próximos, pois dependendo do ponto inicial tomado, há convergência em um único passo. Entretanto, é indubitável que para os pontos suficientemente próximos da conhecida solução, o Método de Newton mostrou-se superior. Porém, o desempenho dos Métodos Quase-Newton verificou-se de maneira notável e, em geral, eis a preferência de implementação destes métodos, em especial do BFGS, em solvers com fins comerciais, pois o não conhecimento da solução global é um fator presente, ou seja, não há conhecimento prévio do minimizador global da função como o aqui apresentado, o que traz aos Métodos Quase-Newton maior eficiência geral em robustez e tempo de processamento se comparados ao Método de Newton e ao Método do Gradiente. Por último, o Método dos Gradientes Conjugados utilizando a escolha de Hestenes e Stiefel para o parâmetro α_j demonstrou sua eficiência computacional(veja o capítulo de avaliações computacionais) quanto à robustez e ao tempo de processamento em comparação às escolhas α^{PR} (Polak e Ribiere) e α^{FR} (Fletcher e Reeves) comentados em (NOCEDAL; WRIGHT, 2006) e (BAZARAA, 2006) quando buscas lineares inexatas são feitas, fato este confirmado nos experimentos acima.

REFERÊNCIAS

- BAZARAA, M. S. **Nonlinear programming : theory and algorithms**. Hoboken, N.J: Wiley-Interscience, 2006.
- DENNIS, J. E.; SCHNABEL, R. B. **Numerical Methods for Unconstrained Optimization and Nonlinear Equations**. [S.l.]: Society for Industrial e Applied Mathematics, jan. 1996. P. 239–258.
- FRIEDLANDER, A. **Elementos de Programação Não-Linear**. Campinas: [s.n.].
- HOFFMAN, K.; KUNZE, R. **Linear Algebra**. 2. ed. Upper Saddle River, NJ: Pearson, abr. 1971.
- LIMA, E. **Curso de Análise Vol. 1**. Rio de Janeiro: IMPA, 2008.
- _____. **Curso de Análise Vol. 2**. Rio de Janeiro: IMPA, 2020.
- LUENBERGER, D. G.; YE, Y. **Linear and nonlinear programming**. New York, NY: Springer, nov. 2010. (International Series in Operations Research & Management Science).
- MARTÍNEZ, J. M.; AUGUSTA, S. **Métodos Computacionais de Otimização**. Campinas: [s.n.], 1998.
- MUNARI, P. A. **Comparação de softwares científicos utilizando perfis de desempenho: automatização dos cálculos pela planilha perfis.xls**. [S.l.: s.n.], 2009.
- NOCEDAL, J.; WRIGHT, S. **Numerical Optimization**. 2. ed. New York, NY: Springer, jul. 2006. (Springer Series in Operations Research and Financial Engineering).