

UNIVERSIDADE FEDERAL DO PARANÁ

FRANCINE LUISE MILIAN RIBAS

MÉTODOS DE REGIÃO DE CONFIANÇA PARA MINIMIZAÇÃO IRRESTRITA

CURITIBA

2017

FRANCINE LUISE MILIAN RIBAS

MÉTODOS DE REGIÃO DE CONFIANÇA PARA MINIMIZAÇÃO IRRESTRITA

Monografia apresentada à disciplina de Projeto de Matemática Industrial como requisito à conclusão do curso de Bacharelado em Matemática Industrial, Setor de Ciências Exatas, Universidade Federal do Paraná.

Orientador: Prof. Dr. Abel Soares Siqueira

CURITIBA

2017

Dedico este trabalho a minha avó.

Agradecimentos

Agradeço primeiramente a minha mãe, por toda a dedicação e carinho ao longo de todos estes anos de vida.

Agradeço ao meu orientador pela paciência e pelo acompanhamento durante o andamento deste trabalho, e por fim agradeço aos amigos que me acompanharam durante esta longa caminhada, especialmente a Janaina, Ana Carolina e Oksana, que estiveram sempre me dando o apoio necessário para superar os obstáculos que surgiram.

Resumo

Neste trabalho foi realizado um estudo a respeito de métodos de região de confiança, revendo antes alguns aspectos teóricos e um pouco sobre os métodos de Newton, Cauchy e Busca Linear. Também foram realizadas implementações de combinações de alguns destes métodos para resolver problemas de minimização irrestrita. Os algoritmos implementados, em linguagem Julia, foram testados e a partir dos resultados, foram comparados entre si.

Palavras-chave: Região de confiança, programação não linear, minimização irrestrita.

Abstract

In this work we performed a study about trust-region methods, before reviewing some theoretical aspects and a little about the methods of Newton, Cauchy and Linear Search. In this work we also performed implementations of combinations of some of these methods to solve problems of unconstrained minimization. The implemented algorithms, in Julia language, were tested and from the results, were compared among themselves.

Keywords: Trust-Region, nonlinear programming, unrestricted minimization.

Lista de ilustrações

Figura 1.1 – Exemplo mínimo local(ambos) e global(esquerda).	13
Figura 1.2 – Exemplo mínimo local estrito.	13
Figura 2.1 – Ilustração de 6 iterações do método de Newton	19
Figura 2.2 – Ilustração de 11 iterações do método de Cauchy	21
Figura 3.1 – Aproximação por Dogleg	36
Figura 4.1 – Perfil de desempenho - tempo para 28 problemas	43
Figura 4.2 – Perfil de desempenho - avaliações de função para 28 problemas	43
Figura 4.3 – Perfil de desempenho - tempo para 173 problemas	44
Figura 4.4 – Perfil de desempenho - avaliações de função para 173 problemas	45

Lista de tabelas

Tabela 4.1 – Saídas dos algoritmos.	40
Tabela 4.2 – Newton-Steihaug para 28 problemas.	41
Tabela 4.3 – LBFGS-Dogleg para 28 problemas.	42
Tabela A.1 – Newton-Steihaug para 173 problemas.	52
Tabela A.2 – LBFGS-Dogleg para 173 problemas.	57

Lista de algoritmos

1	Método de Newton	19
2	Método de Cauchy	21
3	Método de Região de Confiança	28
4	Método de Moré-Sorensen	35
5	Método Dogleg	37
6	Método de Steihaug	38

Sumário

	Introdução	11
1	OTIMIZAÇÃO IRRESTRITA	12
1.1	O problema de otimização não linear	12
1.2	Condições de otimalidade	14
2	MÉTODOS DE OTIMIZAÇÃO IRRESTRITA	18
2.1	Método de Newton	18
2.2	Método de Cauchy	20
2.3	Busca Linear	22
2.3.1	Tamanho do passo	23
2.4	Métodos quase-Newton	24
3	MÉTODOS DE REGIÃO DE CONFIANÇA	26
3.1	O Passo de Cauchy	27
3.2	Convergência	30
3.3	Resolução do subproblema	34
3.3.1	Método de Moré-Sorensen	34
3.3.2	Método Dogleg	35
3.3.3	Método de Steihaug	37
4	TESTES PRÁTICOS	40
	Conclusão	46
	Bibliografia	47
A	RESULTADOS	48

Introdução

Neste trabalho, realizamos um estudo sobre os métodos de Região de Confiança para resolução de problemas irrestritos. Foram implementados dois algoritmos em linguagem Julia. O primeiro algoritmo é o método LBFGS com região de confiança e a direção encontrada pelo método Dogleg. O segundo algoritmo é o método de Newton com região de confiança e a direção dada pelo método de Steihaug. Ambos os algoritmos foram testados com problemas irrestritos da biblioteca CUTEst e comparados entre si.

No primeiro capítulo apresentamos o problema de otimização não linear, algumas definições de minimizadores e algumas condições de otimalidade que nos permitem reconhecer e classificar estes minimizadores.

No segundo capítulo, mostraremos alguns métodos de otimização irrestrita, começando pelos métodos do Gradiente e de Newton, falaremos então um pouco sobre Busca Linear e os métodos Quase-Newton.

No terceiro capítulo faremos um estudo sobre métodos de região de confiança, começando pela teoria de região de confiança e depois veremos alguns métodos de região de confiança.

No quarto capítulo, veremos uma melhor descrição dos algoritmos implementados e os resultados dos testes realizados com os problemas irrestritos da biblioteca CUTEst. Vamos também comparar os algoritmos através de um perfil de desempenho, para avaliarmos a eficiência e robustez dos algoritmos.

1 Otimização irrestrita

A otimização é uma ferramenta importante para tomada de decisões. Para fazer uso desta ferramenta, precisamos definir um objetivo. Neste caso estamos interessados no problema de minimizar uma função $f(x)$, com $x \in \mathbb{R}^n$, sendo x pertencente a um conjunto fechado Ω de pontos viáveis. Se Ω for igual a todo o \mathbb{R}^n , teremos um problema de otimização irrestrita, já se Ω for igual a um subconjunto próprio do \mathbb{R}^n , teremos um problema de otimização restrita.

Neste capítulo, veremos algumas definições de minimizadores que nos permitem verificar se o minimizador é local ou global. Veremos também como podemos encontrar um minimizador e como podemos caracterizá-lo. As definições e teoremas aqui apresentados também podem ser vistos em [1], [2] e [3].

1.1 O problema de otimização não linear

Considere o problema

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & x \in \Omega, \end{aligned} \tag{1.1}$$

sendo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $\Omega \subset \mathbb{R}^n$. Se o conjunto Ω não for citado, então $\Omega = \mathbb{R}^n$ e estaremos lidando com um problema sem restrições.

DEFINIÇÃO 1.1.1. Um ponto x^* é um minimizador global de f em Ω , quando $f(x^*) \leq f(x)$, para todo $x \in \Omega$.

A solução ideal para o problema (1.1) seria encontrar um ponto $x^* \in \Omega$ que fosse um minimizador global da função f , onde a função assume seu menor valor em Ω .

DEFINIÇÃO 1.1.2. Um ponto x^* é um minimizador local de f em Ω , quando existir uma vizinhança \mathcal{V} de x^* tal que $f(x^*) \leq f(x)$, para todo $x \in \mathcal{V} \subset \Omega$.

Encontrar o mínimo global da função f pode não ser muito fácil, porém podemos acabar encontrando um mínimo local da função, onde f assume seu menor valor dentro de uma vizinhança do ponto.

Temos ainda que o ponto x^* pode ser um minimizador local estrito, definido a seguir.

DEFINIÇÃO 1.1.3. Um ponto x^* é um minimizador local estrito de f em Ω , quando existir uma vizinhança \mathcal{V} de x^* tal que $f(x^*) < f(x)$, para todo $x \in \mathcal{V} \cap \Omega$ com $x \neq x^*$.

O ponto em vermelho na Figura 1.1 representa um minimizador local e o ponto em verde representa o minimizador global.

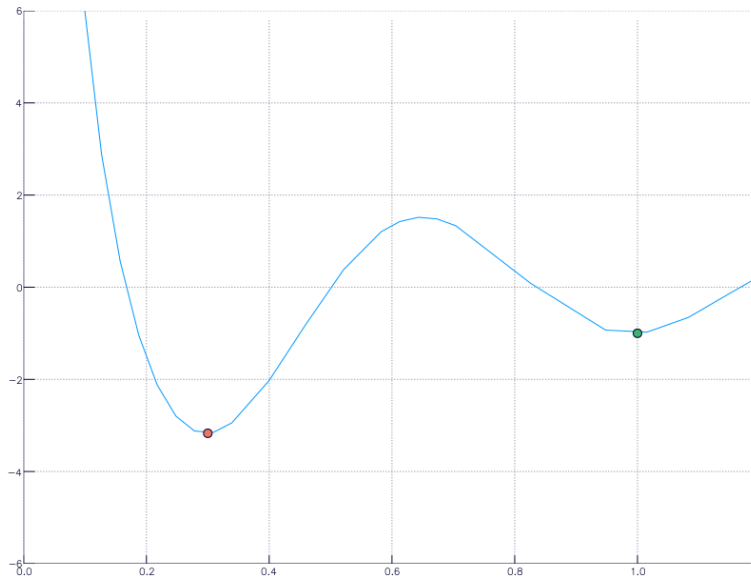


Figura 1.1: Exemplo mínimo local(ambos) e global(esquerda).

Podemos ver, na Figura 1.2, um exemplo de minimizador local estrito.

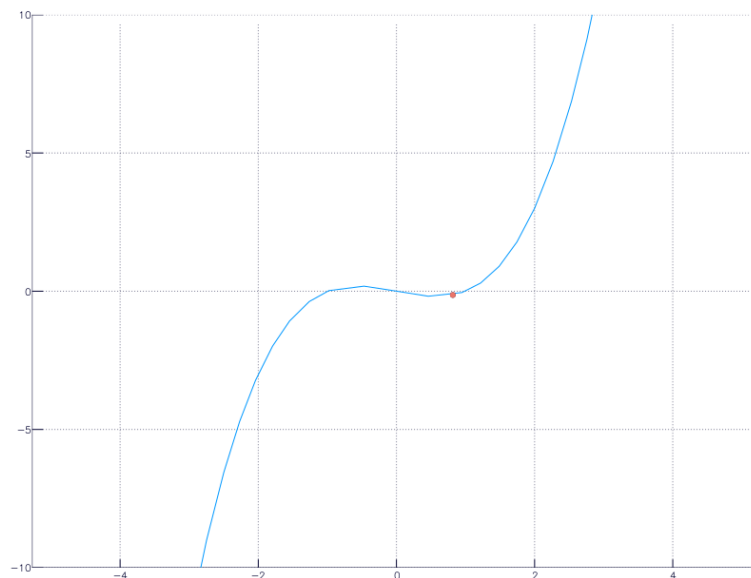


Figura 1.2: Exemplo mínimo local estrito.

1.2 Condições de otimalidade

As definições vistas na seção anterior não apresentam como encontrar ou como reconhecer um minimizador. Uma ferramenta importante para a caracterização dos minimizadores são os Teoremas de Taylor.

TEOREMA 1.2.1 (Taylor de 1ª ordem). *Considere $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável em $\bar{x} \in \mathbb{R}^n$, então*

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + r(x),$$

com

$$\lim_{x \rightarrow \bar{x}} \frac{r(x)}{\|x - \bar{x}\|} = 0.$$

TEOREMA 1.2.2 (Taylor de 2ª ordem). *Considere $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função duas vezes diferenciável em $\bar{x} \in \mathbb{R}^n$, então*

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + \frac{1}{2} (x - \bar{x})^T \nabla^2 f(\bar{x}) (x - \bar{x}) + r(x),$$

com

$$\lim_{x \rightarrow \bar{x}} \frac{r(x)}{\|x - \bar{x}\|^2} = 0.$$

TEOREMA 1.2.3 (Taylor com resto de Lagrange). *Considere $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe C^1 e $\bar{x}, d \in \mathbb{R}^n$. Se f é duas vezes diferenciável no segmento $(\bar{x}, \bar{x} + d)$, então existe uma constante $\alpha \in (0, 1)$ tal que*

$$f(\bar{x} + d) = f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + \frac{1}{2} d^T \nabla^2 f(\bar{x} + \alpha d) d.$$

Agora veremos algumas condições importantes para que um determinado ponto possa ser caracterizado como um minimizador local do Problema (1.1), considerando $\Omega = \mathbb{R}^n$.

TEOREMA 1.2.4 (Condição necessária de 1ª ordem). *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciável no ponto $x^* \in \mathbb{R}^n$. Se x^* é um minimizador local de f , então*

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

Demonstração. Seja $d \in \mathbb{R}^n$ com $d \neq 0$. Como x^* é minimizador local de f , então

$$f(x^*) \leq f(x^* + td), \tag{1.2}$$

para um t suficientemente pequeno. Pelo Teorema 1.2.1,

$$f(x^* + td) = f(x^*) + t \nabla f(x^*)^T d + r(t),$$

com $\lim_{t \rightarrow 0} \frac{r(t)}{t} = 0$. Dividindo por t , temos

$$\nabla f(x^*)^T d + \frac{r(t)}{t} = \frac{f(x^* + td) - f(x^*)}{t}.$$

Usando (1.2)

$$\nabla f(x^*)^T d + \frac{r(t)}{t} \geq 0.$$

Passando o limite, temos

$$\lim_{t \rightarrow 0} \nabla f(x^*)^T d + \lim_{t \rightarrow 0} \frac{r(t)}{t} \geq 0.$$

Como

$$\lim_{t \rightarrow 0} \frac{r(t)}{t} = 0,$$

temos

$$\nabla f(x^*)^T d \geq 0.$$

Suponha por contradição que $\nabla f(x^*) \neq 0$ e defina $d = -\nabla f(x^*)$. Observe que $\nabla f(x^*)^T d = -\|\nabla f(x^*)\|^2 < 0$. Portanto, $\nabla f(x^*) = 0$. \square

O ponto que satisfaz a condição $\nabla f(x^*) = 0$ é chamado de ponto crítico, ou ponto estacionário de f .

DEFINIÇÃO 1.2.5. Considere uma função diferenciável $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $\bar{x} \in \mathbb{R}^n$ um ponto crítico de f . Dizemos que \bar{x} é um ponto de sela de f quando para todo $\epsilon > 0$ existem $x, y \in \mathcal{B}(\bar{x}, \epsilon)$ tal que

$$f(x) < f(\bar{x}) < f(y).$$

TEOREMA 1.2.6 (Condição necessária de 2ª ordem). *Se x^* é um minimizador local e f é duas vezes diferenciável no ponto x^* , então a Hessiana $\nabla^2 f(x^*)$ é semidefinida positiva, ou seja,*

$$d^T \nabla^2 f(x^*) d \geq 0,$$

para todo $d \in \mathbb{R}^n$.

Demonstração. Seja $d \in \mathbb{R}^n$ com $d \neq 0$. Do Teorema 1.2.2, temos

$$f(x^* + td) = f(x^*) + t \nabla f(x^*)^T d + \frac{1}{2} t^2 d^T \nabla^2 f(x^*) d + r(t), \quad (1.3)$$

com

$$\lim_{t \rightarrow 0} \frac{r(t)}{t^2} = 0.$$

Para um t suficientemente pequeno, temos

$$f(x^* + td) - f(x^*) \geq 0,$$

de (1.3), e como x^* é um minimizador local de f e f é diferenciável em x^* temos, pelo Teorema 1.2.4, $\nabla f(x^*) = 0$, que

$$\frac{t^2}{2} d^T \nabla^2 f(x^*) d + r(t) \geq 0.$$

Dividindo por t^2 ,

$$\frac{d^T \nabla^2 f(x^*) d}{2} + \frac{r(t)}{t^2} \geq 0.$$

Passando o limite,

$$\lim_{t \rightarrow 0} \frac{d^T \nabla^2 f(x^*) d}{2} + \lim_{t \rightarrow 0} \frac{r(t)}{t^2} \geq 0$$

Disto, temos que $d^T \nabla^2 f(x^*) d \geq 0$, para todo $d \in \mathbb{R}^n$. □

TEOREMA 1.2.7 (Condição suficiente de 2ª ordem). *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ duas vezes continuamente diferenciável numa vizinhança de x^* . Se x^* é ponto crítico de f e $\nabla^2 f(x^*)$ é definida positiva, então x^* é um minimizador local estrito de f .*

Demonstração. Como $\nabla^2 f$ é contínua e definida positiva em x^* , conseguimos escolher um raio $r > 0$ de forma que $\nabla^2 f$ é contínua definida positiva para todo x em uma bola aberta $\mathcal{B} = \{z \mid \|z - x^*\| < r\}$. Tomando $d \in \mathbb{R}^n$ não nulo com $\|d\| < r$, temos que $x^* + d \in \mathcal{B}$ e então, pelo Teorema de Taylor com resto de Lagrange, temos

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

sendo $z = x^* + td$ e $t \in (0, 1)$. Do Teorema 1.2.4, $\nabla f(x^*) = 0$ e então

$$f(x^* + d) = f(x^*) + \frac{1}{2} d^T \nabla^2 f(z) d.$$

Como $z \in \mathcal{B}$, temos que $d^T \nabla^2 f(z) d > 0$. Portanto $f(x^* + d) > f(x^*)$. □

COROLÁRIO 1.2.8. *Se x^* é um ponto crítico de f e $\nabla^2 f(x^*)$ é definida negativa, então x^* é um maximizador de f . se x^* é um ponto crítico de f e a Hessiana $\nabla^2 f(x^*)$ é indefinida, então o ponto x^* é um ponto de sela.*

EXEMPLO 1.2.9. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por $f(x_1, x_2) = 3x_1^2 + 3x_1x_2 + 2x_2^2 + x_1 + x_2$, verifique que a função f tem um minimizador local.*

Começamos calculando o gradiente de f

$$\nabla f(x) = \begin{bmatrix} 6x_1 + 3x_2 + 1 \\ 3x_1 + 4x_2 + 1 \end{bmatrix}.$$

Fazendo $\nabla f(x) = 0$, temos o seguinte sistema

$$\begin{cases} 6x_1 + 3x_2 + 1 = 0 \\ 3x_1 + 4x_2 + 1 = 0. \end{cases}$$

Resolvendo este sistema, obtemos

$$x^* = \begin{bmatrix} -\frac{1}{15} \\ -\frac{1}{5} \end{bmatrix},$$

que é um ponto crítico de f . Calculando a Hessiana no ponto x^* temos,

$$\nabla^2 f(x^*) = \begin{bmatrix} 6 & 3 \\ 3 & 4 \end{bmatrix}.$$

Como o determinante e todos os subdeterminantes de $\nabla^2 f(x^*)$ são maiores que 0, $\nabla^2 f(x^*)$ é definida positiva temos pelo Teorema 1.2.7 que x^* é minimizador local estrito de f .

2 Métodos de Otimização Irrestrita

O modelo geral de um problema irrestrito é dado por

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & x \in \mathbb{R}^n. \end{aligned} \tag{2.1}$$

Neste capítulo veremos duas estratégias utilizadas para minimizar o Problema (2.1). Primeiramente veremos métodos de busca linear e depois veremos alguns métodos de região de confiança. Para a construção deste capítulo foram usadas como referências [1], [2] e [3].

2.1 Método de Newton

O Método de Newton é um dos mais conhecidos em otimização, e inspirou vários outros métodos. É um método iterativo para encontrar raízes do sistema de equações não lineares diferenciáveis

$$F(x) = 0, \tag{2.2}$$

sendo $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de classe C^1 .

Vamos aproximar F pelo polinômio de Taylor de primeira ordem, em uma vizinhança de um ponto $\bar{x} \in \mathbb{R}^n$,

$$F(x) \approx F(\bar{x}) + J(\bar{x})(x - \bar{x}), \tag{2.3}$$

sendo $J(\bar{x})$ a matriz Jacobiana de F . Como encontrar um ponto $x \in \mathbb{R}^n$ que satisfaça (2.2) é difícil, vamos encontrar um ponto x que anule a aproximação de Taylor de F dada em (2.3). Para tanto, vamos resolver o sistema

$$F(\bar{x}) + J(\bar{x})(x - \bar{x}) = 0. \tag{2.4}$$

Se $J(\bar{x})$ for inversível, o sistema (2.4) tem solução

$$\tilde{x} = \bar{x} - J(\bar{x})^{-1}F(\bar{x}).$$

O método de Newton se baseia nesta solução para encontrar raízes de um sistema de equações não lineares diferenciáveis.

O Problema (2.2) pode ser relacionado com o Problema (2.1), em que desejamos encontrar um ponto crítico, ou seja, fazendo $F(x) = \nabla f(x)$. Neste caso, vamos buscar a solução do sistema

$$\nabla f(\bar{x}) + \nabla^2 f(\bar{x})(x - \bar{x}) = 0.$$

Se $\nabla^2 f(\bar{x})$ for inversível, obtemos a seguinte solução do sistema

$$\tilde{x} = \bar{x} - \nabla^2 f(\bar{x})^{-1} \nabla f(\bar{x}).$$

O método de Newton é apresentado no Algoritmo 1.

Algoritmo 1 Método de Newton

- 1: Dado $x^0 \in \mathbb{R}^n$
 - 2: $k = 0$
 - 3: **enquanto** $\nabla f(x^k) \neq 0$ **faça**
 - 4: $d^k = -[\nabla^2 f(x^k)]^{-1} \nabla f(x^k)$
 - 5: $x^{k+1} = x^k + d^k$
 - 6: $k = k + 1$
 - 7: **fim enquanto**
-

Na Figura 2.1 podemos ver algumas iterações do método de Newton, para encontrar o mínimo da função $f(x) = \frac{1}{2}(x_2 - x_1^2)^2 + \frac{1}{2}(1 - x_1)^2$ com $x^0 = (-1.2, 1.0)$ e $x^* = (1, 0, 1.0)$.

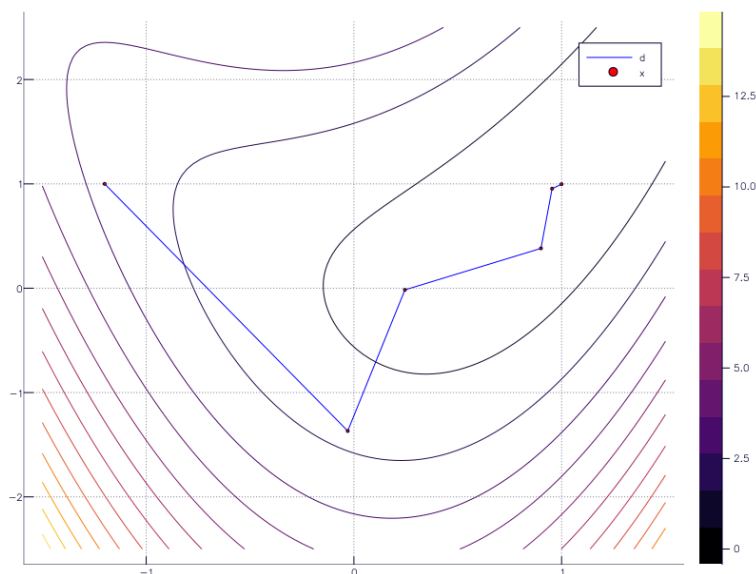


Figura 2.1: Ilustração de 6 iterações do método de Newton

O método de Newton encontra um pontos crítico da função f , porém o método não nos dá nenhuma garantia de que esse ponto será o minimizador de f . Além disso, caso a Hessiana não seja inversível, a direção pode não estar bem definida.

O método de Newton possui variações, em uma destas variações o método utiliza a condição de Armijo, que será vista mais adiante.

2.2 Método de Cauchy

O método de Cauchy, também conhecido como método do Gradiente, é um método de minimização de uma função, que, a partir de x^k , caminha ao longo da direção de busca d^k dada por

$$d^k = -\nabla f(x^k).$$

Uma justificativa para esta tanto é que a direção do gradiente aponta para o maior crescimento da função, neste caso a direção de busca é a direção oposta ao gradiente, que aponta para o maior decréscimo da função. De fato, se $d = -\nabla f(x)$ e sendo $v \in \mathbb{R}^n$ tal que $\|v\| = \|d\|$, então

$$\begin{aligned} \frac{\partial f}{\partial d}(x) &= \nabla f(x)^T d \\ &= -\|\nabla f(x)\|^2 \\ &= -\|\nabla f(x)\| \|\nabla f(x)\|, \end{aligned}$$

como $\|d\| = \|-\nabla f(x)\| = \|\nabla f(x)\|$ e $\|v\| = \|d\|$

$$\frac{\partial f}{\partial d}(x) = -\|\nabla f(x)\| \|v\|,$$

usando a desigualdade de Cauchy-Schwarz, temos

$$-\|\nabla f(x)\| \|v\| \leq \nabla f(x)^T v,$$

onde

$$\nabla f(x)^T v = \frac{\partial f}{\partial v}(x).$$

Logo,

$$\frac{\partial f}{\partial d}(x) \leq \frac{\partial f}{\partial v}(x),$$

o que comprova a afirmação anterior de que a direção $d = -\nabla f(x)$ é a direção de maior decréscimo da função.

Observando o Algoritmo 2 podemos ver que cada iterando k do método de Cauchy é dado por $x^{k+1} = x^k + t_k d^k$. A necessidade da escolha para t_k está relacionada com o fato de que se utilizarmos cada iterando $x^{k+1} = x^k + d^k$, pode resultar em perder a oportunidade de encontrar o minimizador de f , pois a direção utilizada integralmente pode ser muito grande. Para evitar que isto aconteça, definimos um tamanho do passo t_k , de modo a controlar o quanto caminharemos na direção d^k . Uma forma para obter um valor para t_k é minimizando $f(x^k + t d^k)$.

Algoritmo 2 Método de Cauchy

- 1: Dado $x^0 \in \mathbb{R}^n$
 - 2: $k = 0$
 - 3: **enquanto** $\nabla f(x^k) \neq 0$ **faça**
 - 4: $d^k = -\nabla f(x^k)$
 - 5: Obter $t_k = \operatorname{argmin} f(x^k + t d^k)$
 - 6: $x^{k+1} = x^k + t_k d^k$
 - 7: $k = k + 1$
 - 8: **fim enquanto**
-

Na Figura 2.2 podemos ver algumas iterações do método de Cauchy, sendo $f(x) = 3x_1^2 + 3x_1x_2 + 2x_2^2 + x_1 + x_2$ com $x^0 = (0, 0)$ e $x^* = (-\frac{1}{15}, -\frac{1}{5})$.

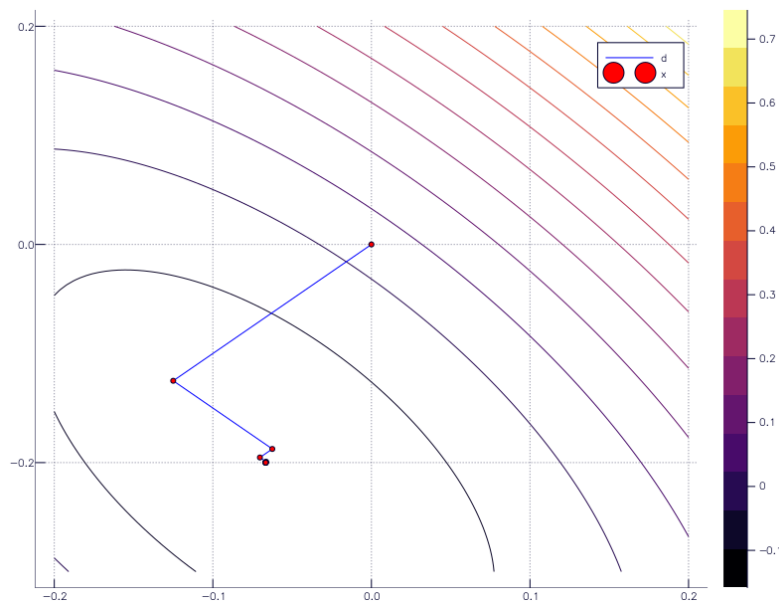


Figura 2.2: Ilustração de 11 iterações do método de Cauchy

Veremos agora o caso particular em que a função objetivo é quadrática. Considerando

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

com $A \in \mathbb{R}^{n \times n}$ definida positiva e $b \in \mathbb{R}^n$. Como f é quadrática, o minimizador x^* satisfaz

$$\nabla f(x^*) = Ax^* - b = 0,$$

sendo x^* a solução única do sistema linear $Ax = b$. Para calcular o tamanho do passo t_k , temos que resolver

$$\begin{aligned} \min \quad & f(x^k + td^k) \\ \text{s.a} \quad & t > 0 \end{aligned}$$

Suponha que x^k não é um ponto crítico. Para encontrar o ponto crítico t_k fazemos $\nabla f(x^k + t_k d^k) = 0$, temos

$$\begin{aligned} (d^k)^T [A(x^k + t_k d^k) - b] &= 0 \\ (d^k)^T [Ax^k - b + t_k d^k A] &= 0 \\ (d^k)^T [\nabla f(x^k) + t_k d^k A] &= 0. \end{aligned}$$

Utilizando o fato de que $d^k = -\nabla f(x^k)$, temos

$$t_k (d^k)^T A d^k - (d^k)^T d^k = 0,$$

e então obtemos

$$t_k = \frac{(d^k)^T d^k}{(d^k)^T A d^k}. \quad (2.5)$$

Como A é definida positiva

$$\nabla^2 f(x^k + t_k d^k) = (d^k)^T A d^k > 0,$$

portanto t_k é minimizador de $f(x^k + t_k d^k)$.

2.3 Busca Linear

Na seção anterior, vimos que para melhorar a capacidade de encontrar uma solução para problemas da forma do problema (2.1), precisamos de uma forma para controlar o quanto caminharemos ao longo de uma direção d^k . Para isso, apresentamos a Busca Linear que faz esse controle a cada iteração. A iteração da busca linear é dada por

$$x^{k+1} = x^k + t_k d^k,$$

onde $t_k > 0$ é denominado tamanho do passo. O sucesso da busca linear depende da escolha adequada de uma direção d^k e um tamanho de passo t_k . Vimos anteriormente

que umas das possibilidades é fazer a minimização de $f(x^k + td^k)$, porém dependendo do problema, isto pode ser muito custoso.

A maioria dos algoritmos que utilizam a busca linear requerem que a direção d^k seja uma direção de descida como definida a seguir.

DEFINIÇÃO 2.3.1. Considere uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$, um ponto $\bar{x} \in \mathbb{R}^n$ e uma direção $d \in \mathbb{R}^n$ com $d \neq 0$. Dizemos que d é uma direção de descida para f , a partir de \bar{x} , quando existe um $\delta > 0$ tal que $f(\bar{x} + td) < f(\bar{x})$, para tod $t \in (0, \delta)$.

TEOREMA 2.3.2. Se $\nabla f(\bar{x})^T d < 0$, então d é uma direção de descida para f , a partir de \bar{x} .

Demonstração. A. A. Ribeiro e E. W. Karas, página 47. □

Isto garante que o valor da função f é reduzido caminhando ao longo desta direção. Além de que, a direção de busca muitas vezes é da forma

$$d^k = -B_k^{-1} \nabla f(x^k),$$

onde B_k é uma matriz simétrica não singular que aproxima $\nabla^2 f(x^k)$. A direção definida desta maneira, sendo B_k definida positiva, será uma direção de descida, pois

$$\nabla f(x^k)^T d^k = -\nabla f(x^k)^T B_k^{-1} \nabla f(x^k) < 0.$$

Note que para o método de Cauchy, a matriz B_k é igual à matriz identidade, e para o método de Newton a matriz B_k é a própria Hessiana.

2.3.1 Tamanho do passo

Para termos um bom passo, temos que escolher um t_k de modo a obter uma redução de f , sem gastar muito tempo computacional. Um tamanho de passo adequado seria o minimizador de $f(x^k + td^k)$. Mas para encontrar um minimizador, são necessárias muitas avaliações de f e talvez de ∇f .

Algumas estratégias utilizam da busca linear inexata para escolher um tamanho para o passo t_k , que proporcione um decréscimo suficiente de f com baixo custo computacional. Para isso são testados vários valores para t_k até que algumas condições sejam satisfeitas.

Uma das condições é que t_k proporcione um decréscimo simples no valor da função f comparado ao valor de f em t_k , isto é, que satisfaça

$$f(x^k + td^k) < f(x^k).$$

Somente isto não é suficiente, pois não garante a convergência. A busca linear inexata possui uma condição que determina que t_k deve proporcionar um decréscimo suficiente de f , conforme a seguinte desigualdade

$$f(x^k + td^k) \leq f(x^k) + c_1 t \nabla f(x^k)^T d^k, \quad (2.6)$$

para $c_1 \in (0, 1)$, esta condição é conhecida como condição de Armijo. Outra condição que pode ser imposta a t_k é a chamada condição de curvatura

$$\nabla f(x^k + td^k)^T d^k \geq c_2 \nabla f(x^k)^T d^k, \quad (2.7)$$

para $c_2 \in (c_1, 1)$. Em conjunto, as condições (2.6) e (2.7) juntas, são chamadas de condições de Wolfe.

2.4 Métodos quase-Newton

Os métodos quase-Newton são uma família de métodos que, como o método de Cauchy, requerem por iteração apenas a avaliação do gradiente da função objetivo e de uma aproximação da matriz Hessiana. Estes métodos apresentam uma melhora de desempenho com relação ao método de Cauchy e possuem uma estrutura muito parecida com a do método de Newton. A direção de busca é dada por

$$d^k = -B_k^{-1} \nabla f(x^k),$$

e o novo iterando é

$$x^{k+1} = x^k + t_k d^k,$$

sendo que o passo t_k deve satisfazer a condição (2.6) mais opcionalmente (2.7).

A diferença para o método de Newton está na utilização de uma aproximação da matriz Hessiana B_k , no lugar da verdadeira Hessiana, o que dispensa a utilização de segundas derivadas, tornando os métodos quase-Newton mais baratos em alguns casos.

Um dos métodos quase-Newton mais conhecido é o método BFGS (Broyden-Fletcher-Goldfarb-Shanno), que utiliza no lugar de B_k , uma aproximação da inversa da matriz Hessiana denotada por

$$H_k = B_k^{-1}.$$

A atualização de H_k é feita da seguinte forma

$$H_{k+1} = (I - \rho_k s_k y_k^T) H_k (I - \rho_k s_k y_k^T) + \rho_k s_k s_k^T,$$

sendo

$$\rho_k = \frac{1}{y_k^T s_k}, \quad y_k = x^{k+1} - x^k, \quad s_k = \nabla f(x^{k+1}) - \nabla f(x^k). \quad (2.8)$$

Disto temos que a aproximação da matriz Hessiana é dada por

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}.$$

Para problemas muito grandes, as matrizes podem ser muito densas, tornando muito custoso computacionalmente calculá-las e armazená-las iteração após iteração. Para contornar este problema, surgiu um método variante do método BFGS, o método LBGFS (Limited-memory BFGS).

O método LBGFS utiliza uma versão modificada da matriz Hessiana, que armazena somente um determinado número de pares de vetores $\{s_k, y_k\}$ dados em (2.8). O produto $B_k \nabla f(x^k)$, utilizado no cálculo da direção, pode ser obtido a partir de uma sequência de produtos internos e somas de vetores utilizando $\nabla f(x^k)$ e os pares de vetores $\{s_k, y_k\}$. Assim que o novo iterando é armazenado, o último par de vetores $\{s_k, y_k\}$ é substituído por um novo, e assim são mantidas apenas informações das iterações mais recentes.

3 Métodos de Região de Confiança

Para a construção deste capítulo foram usadas como referências [1], [2] e [3].

O Método de Região de Confiança também é utilizado para resolver (2.1). Este define um modelo a partir da função objetivo e uma região ao redor do ponto corrente, na qual esperamos que o modelo seja uma representação adequada da função nesta região. Definidos o modelo e a região de confiança, vamos minimizar o modelo e usar a solução do mesmo como próximo iterado.

Considerando uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^2 , dado um ponto $x^k \in \mathbb{R}^n$, o modelo quadrático de f é dado por

$$q_k(x) = f(x^k) + \nabla f(x^k)^T (x - x^k) + \frac{1}{2}(x - x^k)^T B_k (x - x^k),$$

onde $B_k \in \mathbb{R}^{n \times n}$ pode ser a própria Hessiana $\nabla^2 f(x^k)$ como uma matriz simétrica que a aproxime. Vamos considerar $\Delta_k > 0$ e a região

$$\{x \in \mathbb{R}^n \mid \|x - x^k\| \leq \Delta_k\}.$$

Chamamos Δ_k de raio da região de confiança e esta região de região de confiança, no sentido de que esperamos que o modelo represente f adequadamente nessa região. Para facilitar, usaremos

$$d = x - x^k$$

e também,

$$m_k(d) = q_k(x^k + d).$$

Para obter um passo d^k , precisamos resolver o seguinte subproblema,

$$\begin{aligned} \min \quad & m_k(d) = f(x^k) + g_k^T d + \frac{1}{2} d^T B_k d \\ \text{s. a} \quad & \|d\| \leq \Delta_k, \end{aligned} \tag{3.1}$$

sendo $g_k = \nabla f(x^k)$.

No entanto, resolver exatamente (3.1) pode ser muito custoso, então usaremos uma aproximação do minimizador. Comparamos o decréscimo da função objetivo com o decréscimo obtido pelo modelo a partir da razão

$$\rho_k = \frac{f(x^k) - f(x^k + d^k)}{m_k(0) - m_k(d^k)}. \quad (3.2)$$

Definimos

$$\text{ared} = f(x^k) - f(x^k + d^k),$$

que é a redução real, e

$$\text{pred} = m_k(0) - m_k(d^k),$$

que é a redução predita. Note que $m_k(0) = f(x^k)$.

O controle do tamanho da região de confiança deve ser adequado, uma vez que, se o raio da região for muito pequeno, o algoritmo pode perder a chance de encontrar um passo que pode chegar a um ponto mais próximo do minimizador da função objetivo. Se o raio da região de confiança for grande, o modelo pode não representar a função objetivo adequadamente. A medida (3.2) auxilia no controle da região de confiança.

Se ρ_k for muito pequeno, a redução de $f(x^k + d^k)$ não será suficiente, nesse caso o passo é rejeitado e então reduzimos Δ_k . Porém, se ρ_k está próximo de 1, o modelo m_k é uma boa aproximação de f , então aumentamos Δ_k . Para valores intermediários de ρ_k a redução é boa, mas não é excepcional, então aceitamos o passo, mas mantemos Δ_k .

O método de região de confiança é apresentado no Algoritmo 3.

3.1 O Passo de Cauchy

O passo de Cauchy proporciona uma redução suficiente do modelo m_k e que já garante a convergência do algoritmo. Vamos definir o passo de Cauchy como minimizador do modelo m_k na direção oposta ao gradiente e sujeito à região de confiança. Temos,

$$d_C^k = -t_k g_k, \quad (3.3)$$

onde $t_k > 0$ é solução do problema

$$\begin{aligned} \min \quad & m_k(-t g_k) = f(x^k) - t \|g_k\|^2 + \frac{1}{2} t^2 g_k^T B_k g_k \\ \text{s. a} \quad & \|t g_k\| \leq \Delta_k, \quad t \geq 0. \end{aligned} \quad (3.4)$$

LEMA 3.1.1. *O passo de Cauchy, definido por (3.3), satisfaz*

$$m_k(0) - m_k(d_C^k) \geq \frac{1}{2} \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}.$$

Algoritmo 3 Método de Região de Confiança

-
- 1: Dados $x^0 \in \mathbb{R}^n$, $\Delta_0 > 0$, $\eta_1 \in \left[0, \frac{1}{4}\right)$, $\eta_2 \geq \eta_1$ e $0 < \sigma_1 < 1 \leq \sigma_2$.
 - 2: $k = 0$
 - 3: **enquanto** $\nabla f(x^k) \neq 0$ **faça**
 - 4: Obter d^k resolvendo aproximadamente (3.1)
 - 5: $\rho_k = \frac{f(x^k) - f(x^k + d^k)}{m_k(0) - m_k(d^k)}$
 - 6: **se** $\rho_k > \eta_1$ **então**
 - 7: $x^{k+1} = x^k + d^k$
 - 8: **senão**
 - 9: $x^{k+1} = x^k$
 - 10: **fim se**
 - 11: **se** $\rho_k < \eta_1$ **então**
 - 12: $\Delta_{k+1} = \sigma_1 \Delta_k$
 - 13: **senão se** $\rho > \eta_2$ e $\|d^k\| = \Delta_k$ **então**
 - 14: $\Delta_{k+1} = \sigma_2 \Delta_k$
 - 15: **senão**
 - 16: $\Delta_{k+1} = \Delta_k$
 - 17: **fim se**
 - 18: $k = k + 1$
 - 19: **fim enquanto**
-

Demonstração. A primeira coisa a se fazer é encontrar a solução t_k do problema (3.4), ou seja, encontrar o minimizador da função quadrática

$$\psi(t) = f(x^k) - t \|g_k\|^2 + \frac{1}{2} t^2 g_k^T B_k g_k,$$

com $t \geq 0$. Da restrição do problema (3.4), temos

$$\begin{aligned} \|tg_k\| &\leq \Delta_k, \\ t \|g_k\| &\leq \Delta_k, \\ t &\leq \frac{\Delta_k}{\|g_k\|}, \end{aligned}$$

o intervalo onde devemos encontrar o minimizador.

Primeiramente vamos considerar o caso em que $g_k^T B_k g_k > 0$. Neste caso sabemos que a função ψ é convexa em t , pois $\psi''(t) = g_k^T B_k g_k > 0$, e portanto tem um minimizador t^* irrestrito. Para encontrar este minimizador fazemos $\psi'(t) = 0$,

$$\begin{aligned} -\|g_k\|^2 + t g_k^T B_k g_k &= 0, \\ t g_k^T B_k g_k &= \|g_k\|^2 \end{aligned}$$

obtendo

$$t^* = \frac{\|g_k\|^2}{g_k^T B_k g_k}. \quad (3.5)$$

Se $t^* \leq \frac{\Delta_k}{\|g_k\|}$, podemos fazer $t_k = t^*$ pois t^* está dentro do intervalo que restringe o problema (3.4). Então,

$$m_k(0) - m_k(d_C^k) = \frac{1}{2} \frac{\|g_k\|^4}{g_k^T B_k g_k}.$$

Pela desigualdade de Cauchy-Schwarz e pela consistência da norma 2, temos

$$\begin{aligned} g_k^T B_k g_k &\leq \|g_k\| \|B_k g_k\| \\ &\leq \|g_k\| \|B_k\| \|g_k\|. \end{aligned}$$

Logo,

$$\begin{aligned} m_k(0) - m_k(d_C^k) &\geq \frac{1}{2} \frac{\|g_k\|^4}{\|g_k\|^2 \|B_k\|} \\ &= \frac{1}{2} \frac{\|g_k\|^2}{\|B_k\|}. \end{aligned} \quad (3.6)$$

Se $t^* > \frac{\Delta_k}{\|g_k\|}$, quer dizer que o minimizador da função $\psi(t)$ está na fronteira. Então,

$$t_k = \frac{\Delta_k}{\|g_k\|} < \frac{\|g_k\|^2}{g_k^T B_k g_k}. \quad (3.7)$$

De (3.7), $t_k g_k^T B_k g_k < \|g_k\|^2$, então $t_k^2 g_k^T B_k g_k < t_k \|g_k\|^2 = \Delta_k \|g_k\|$. Daí,

$$\begin{aligned} m_k(d_C^k) &= f(x^k) - t_k \|g_k\|^2 + \frac{1}{2} t_k^2 g_k^T B_k g_k \\ &< f(x^k) - \Delta_k \|g_k\| + \frac{1}{2} \Delta_k \|g_k\|, \end{aligned}$$

e então

$$m_k(0) - m_k(d_C^k) > \frac{1}{2} \Delta_k \|g_k\|. \quad (3.8)$$

No caso em que $g_k^T B_k g_k \leq 0$, temos que $\psi(t)$ é decrescente para $t \geq 0$. Neste caso não existem pontos críticos no intervalo $0 \leq t \leq \frac{\Delta_k}{\|g_k\|}$, então $t_k = \frac{\Delta_k}{\|g_k\|}$. Disto, temos

$$\begin{aligned} m_k(0) - m_k(d_C^k) &= t_k \|g_k\|^2 - \frac{1}{2} t_k^2 g_k^T B_k g_k \\ &\geq t_k \|g_k\|^2, \end{aligned}$$

e então

$$m_k(0) - m_k(d_C^k) \geq \frac{1}{2} \Delta_k \|g_k\|. \quad (3.9)$$

De (3.6), (3.8) e (3.9), temos

$$m_k(0) - m_k(d_C^k) \geq \frac{1}{2} \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}.$$

□

3.2 Convergência

Nesta subseção serão mostrados alguns resultados de convergência, para isso assumiremos que as seguintes hipóteses são satisfeitas.

H1 A função f é de classe C^1 , com ∇f Lipschitz, com constante $L > 0$.

H2 A solução aproximada d^k de (3.1) satisfaz a inequação

$$m_k(0) - m_k(d^k) \geq c_1 \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}, \quad \text{para } c_1 \in (0, 1), \quad (3.10)$$

H3 As aproximações das Hessianas B_k são uniformemente limitadas, de forma que existe uma constante $\beta > 0$ tal que $\|B_k\| \leq \beta$ para todo $k \in \mathbb{N}$.

H4 A função f é limitada inferiormente no conjunto de nível

$$S = \{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\}. \quad (3.11)$$

LEMA 3.2.1. *Suponha que as Hipóteses H1-H3 são satisfeitas. Então existe uma constante $c > 0$ tal que*

$$|\rho_k - 1| \leq \frac{c\Delta_k^2}{\|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\beta} \right\}}.$$

Demonstração. Pelo Teorema de Taylor com resto de Lagrange, sabemos que existe um $\alpha_k \in (0, 1)$ tal que

$$f(x^k + d^k) = f(x^k) + \nabla f(x^k + \alpha_k d^k)^T d^k.$$

Então,

$$\begin{aligned} \text{ared} - \text{pred} &= f(x^k) - f(x^k + d^k) - m_k(0) + m_k(d^k) \\ &= f(x^k) - [f(x^k) + \nabla f(x^k + \alpha_k d^k)^T d^k] - m_k(0) + m_k(d^k) \\ &= -\nabla f(x^k + \alpha_k d^k)^T d^k + g_k^T d^k + \frac{1}{2}(d^k)^T B_k d^k \\ &= \frac{1}{2}(d^k)^T B_k d^k - [\nabla f(x^k + \alpha_k d^k) - g_k]^T d^k. \end{aligned}$$

Da desigualdade de Cauchy-Schwarz, temos

$$|\text{ared} - \text{pred}| \leq \frac{1}{2} \|d^k\|^2 \|B_k\| + \|d^k\| \|\nabla f(x^k + \alpha_k d^k) - g_k\|.$$

Por H1,

$$\|\nabla f(x^k + \alpha_k d^k) - g_k\| \leq L \|\alpha_k d^k\|,$$

então,

$$|\text{ared} - \text{pred}| \leq \frac{1}{2} \|d^k\|^2 \|B_k\| + L\alpha_k \|d^k\|^2.$$

De H3,

$$\begin{aligned} |\text{ared} - \text{pred}| &\leq \frac{1}{2} \|d^k\|^2 \beta + L\alpha_k \|d^k\|^2, \\ &\leq \|d^k\|^2 \left(L\alpha_k + \frac{\beta}{2} \right). \end{aligned}$$

Sabendo que $\|d^k\| \leq \Delta_k$, então existe um c_0 tal que

$$|\text{ared} - \text{pred}| \leq c_0 \Delta_k^2.$$

Por (3.10),

$$|\rho_k - 1| = \left| \frac{\text{ared} - \text{pred}}{\text{pred}} \right| \leq \frac{c_0 \Delta_k^2}{c_1 \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\beta} \right\}}.$$

Fazendo $\frac{c_0}{c_1} = c$, temos

$$|\rho_k - 1| = \left| \frac{\text{ared} - \text{pred}}{\text{pred}} \right| \leq \frac{c \Delta_k^2}{\|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\beta} \right\}}.$$

□

Mostraremos agora o primeiro resultado de convergência global, que diz que a sequência de gradientes g_k tem pelo menos um ponto limite em zero.

TEOREMA 3.2.2. *Suponha que as Hipóteses H1-H4 são satisfeitas, então*

$$\liminf_{k \rightarrow \infty} \|\nabla f(x^k)\| = 0.$$

Demonstração. Suponha que essa afirmação é falsa. Então, existe um $\epsilon > 0$ e um $K \in \mathbb{N}$ tal que $\|\nabla f(x^k)\| \geq \epsilon$ para todo $k \geq K$. Considerando $\tilde{\Delta} = \min \left\{ \frac{\epsilon}{\beta}, \frac{\epsilon}{2c} \right\}$, onde c é definida pelo Lema 3.2.1. Se $\Delta_k \leq \tilde{\Delta}$, então

$$\Delta_k \leq \frac{\epsilon}{\beta} \leq \frac{\|\nabla f(x^k)\|}{\beta} \tag{3.12}$$

e

$$\Delta_k \leq \frac{\epsilon}{2c}.$$

Assim,

$$\frac{c\Delta_k}{\epsilon} \leq \frac{1}{2}.$$

Pelo Lema 3.2.1, do fato que $\|\nabla f(x^k)\| \geq \epsilon$ e por (3.12), temos

$$|\rho_k - 1| \leq \frac{c\Delta_k}{\epsilon} \leq \frac{1}{2}.$$

Disso temos que $\rho_k \geq \frac{1}{2} > \eta_1$, e do Algoritmo 3, temos $\Delta_{k+1} \geq \Delta_k$, o que significa que o raio somente será reduzido se $\Delta_k > \tilde{\Delta}$, de modo que $\Delta_{k+1} = \sigma_1 \Delta_k > \sigma_1 \tilde{\Delta}$. Então

$$\Delta_k \geq \min \{ \Delta_0, \sigma_1 \tilde{\Delta} \}, \quad (3.13)$$

para todo $k \geq K$. Considerando o conjunto $\mathcal{K} = \{k \geq K \mid \rho_k \geq \eta_1\}$. Dado $k \in \mathcal{K}$, pelo Algoritmo 3 e por (3.10) temos

$$\begin{aligned} f(x^k) - f(x^{k+1}) &= f(x^k) - f(x^k + d^k) \\ &\geq \eta_1(m_k(0) - m_k(d^k)) \\ &\geq \eta_1 \|\nabla f(x^k)\| \min \left\{ \Delta_k, \frac{\|\nabla f(x^k)\|}{\|B_k\|} \right\} \\ &\geq \eta_1 c_1 \epsilon \min \left\{ \Delta_k, \frac{\epsilon}{\beta} \right\}. \end{aligned} \quad (3.14)$$

De (3.13), temos que existe $\delta > 0$ tal que

$$f(x^k) - f(x^{k+1}) \geq \eta_1 c_1 \epsilon \min \left\{ \Delta_k, \frac{\epsilon}{\beta} \right\} \geq \delta > 0, \quad (3.15)$$

para todo $k \in \mathcal{K}$. Como a sequência $f(x^k)$ não é crescente, e devido à Hipótese H4, temos que $f(x^k) - f(x^{k+1}) \rightarrow 0$. De (3.15), podemos ver que o conjunto \mathcal{K} é finito. Daí, $\rho_k < \eta_1$ para todo k suficientemente grande, de modo que $\Delta_{k+1} = \sigma_1 \Delta_k$ em cada iteração do Algoritmo 3. O que significa que $\Delta_k \rightarrow 0$, que contradiz (3.13). Logo a afirmação do teorema é verdadeira. \square

A seguir veremos um resultado de convergência global mais forte que o anterior, que exige que a redução real de f seja no mínimo um fração da redução predita.

TEOREMA 3.2.3. *Supondo que as Hipótese H1-H4 são satisfeitas e que $\eta_1 > 0$ no Algoritmo 3, então*

$$\nabla f(x^k) \rightarrow 0.$$

Demonstração. Supondo por contradição que para algum $\epsilon > 0$ o conjunto

$$\mathcal{K} = \{k \in \mathbb{N} \mid \|\nabla f(x^k)\| \geq \epsilon\}$$

seja infinito. Dado $k \in \mathcal{K}$, considerando o primeiro índice $\ell_k > k$ tal que $\|\nabla f(x^{\ell_k})\| \leq \frac{\epsilon}{2}$. O Teorema 3.2.2 garante a existência de ℓ_k . Como ∇f é Lipschitz e

$$\begin{aligned} \|\nabla f(x^k) - \nabla f(x^{\ell_k})\| &\geq \|\nabla f(x^k)\| - \|\nabla f(x^{\ell_k})\| \\ &\geq \epsilon - \frac{\epsilon}{2} \\ &= \frac{\epsilon}{2} \end{aligned}$$

temos,

$$\frac{\epsilon}{2} \leq \|\nabla f(x^k) - \nabla f(x^{\ell_k})\| \leq L \|x^k - x^{\ell_k}\|, \quad (3.16)$$

para algum $L > 0$. Agora, definindo

$$\mathcal{S}_k = \{j \in \mathbb{N} \mid k \leq j < \ell_k, x^{j+1} \neq x^j\},$$

de (3.16) temos que

$$L \|x^k - x^{\ell_k}\| \geq \frac{\epsilon}{2},$$

logo

$$\|x^k - x^{\ell_k}\| \geq \frac{\epsilon}{2L}.$$

Também temos que

$$\|x^k - x^{\ell_k}\| \leq \sum_{j \in \mathcal{S}_k} \|x^j - x^{j+1}\|,$$

e então

$$\frac{\epsilon}{2L} \leq \|x^k - x^{\ell_k}\| \leq \sum_{j \in \mathcal{S}_k} \|x^j - x^{j+1}\| \leq \sum_{j \in \mathcal{S}_k} \Delta_j. \quad (3.17)$$

Do Algoritmo 3, de (3.14), da Hipótese H3, do fato que $k \in \mathcal{K}$ e da definição de ℓ_k , temos

$$\begin{aligned} f(x^k) - f(x^{\ell_k}) &= \sum_{j \in \mathcal{S}_k} f(x^j) - f(x^{j+1}) \\ &> \sum_{j \in \mathcal{S}_k} \eta_1 (m_j(0) - m_j(d^j)) \\ &\geq \sum_{j \in \mathcal{S}_k} \eta_1 c_1 \frac{\epsilon}{2} \min \left\{ \Delta_j, \frac{\epsilon}{2\beta} \right\} \\ &\geq \eta_1 c_1 \frac{\epsilon}{2} \min \left\{ \sum_{j \in \mathcal{S}_k} \Delta_j, \frac{\epsilon}{2\beta} \right\}. \end{aligned}$$

Por fim, usando (3.17), temos

$$f(x^k) - f(x^{\ell_k}) \geq \eta_1 c_1 \frac{\epsilon}{2} \min \left\{ \frac{\epsilon}{2L}, \frac{\epsilon}{2\beta} \right\} > 0. \quad (3.18)$$

Como a sequência $f(x^k)$ não é crescente, e pela Hipótese H4, temos que $f(x^k) - f(x^{\ell_k}) \rightarrow 0$, que contradiz (3.18). Logo a afirmação do Teorema é verdadeira. \square

O Teorema 3.2.3 nos diz que todo ponto limite de uma sequência produzida pelo Algoritmo 3 é um ponto crítico.

3.3 Resolução do subproblema

Vimos anteriormente que o Algoritmo 3 requer a solução aproximada do subproblema (3.1). Nesta seção serão apresentados alguns métodos que podem ser utilizados para encontrar tal solução.

3.3.1 Método de Moré-Sorensen

Moré e Sorensen propõem em [4], um método para encontrar uma solução aproximada para (3.1) em um número finito de passos. Podemos provar que a solução exata de (3.1), para algum $\lambda \geq 0$, deve satisfazer as seguintes condições

$$(B + \lambda I)d^* = -g, \quad (3.19a)$$

$$\lambda(\Delta - \|d^*\|) = 0, \quad (3.19b)$$

$$(B + \lambda I) \text{ semidefinida positiva.} \quad (3.19c)$$

J. Nocedal and S. J. Wright, página 70.

Para não sobrecarregar a notação, dado $k \in \mathcal{N}$ temos $g = g_k$ e $B = B_k$. A condição (3.19b), nos diz que ou $\lambda = 0$, ou $(\Delta - \|d^*\|) = 0$, isto é, $\|d^*\| = \Delta$. Quando a solução estiver estritamente dentro da região de confiança, temos $\lambda = 0$, caso contrário precisamos encontrar um λ de forma que a solução esteja na fronteira da região de confiança. O Método de Moré-Sorensen consiste então em, dado $\lambda \geq 0$, encontrar uma direção d^* que seja solução de (3.19a), e a partir da diferença $\Delta - \|d^*\|$, escolher um novo valor para λ .

O método de Moré e Sorensen é apresentado no Algoritmo 4.

Algoritmo 4 Método de Moré-Sorensen

-
- 1: Dados $\lambda \geq 0$, $\Delta > 0$, $\epsilon > 0$, B e g .
 - 2: $d(\lambda) = -(B + \lambda I)^{-1}g$
 - 3: **se** $\|d(0)\| \leq \Delta$ **então**
 - 4: $d^* = d(0)$
 - 5: **fim se**
 - 6: Encontre λ_l^0 e λ_h^0 tais que $\|\lambda_l\| > \Delta$ e $\|\lambda_h\| < \Delta$
 - 7: Calcule $\lambda^k = \frac{\lambda_l^k + \lambda_h^k}{2}$
 - 8: **enquanto** $\left| \|d(\lambda^k)\| - \Delta \right| > \epsilon$ **faça**
 - 9: **se** $\|d(\lambda^k)\| > \Delta$ **então**
 - 10: $\lambda_l^{k+1} = \lambda^k$
 - 11: $\lambda_h^{k+1} = \lambda_h^k$
 - 12: **senão**
 - 13: $\lambda_l^{k+1} = \lambda_l^k$
 - 14: $\lambda_h^{k+1} = \lambda^k$
 - 15: **fim se**
 - 16: Calcule $\lambda^k = \frac{\lambda_l^k + \lambda_h^k}{2}$
 - 17: **fim enquanto**
-

3.3.2 Método Dogleg

O método Dogleg nos dá uma solução aproximada para o Subproblema (3.1). A ideia do método consiste em encontrar uma trajetória que sai da origem até o passo de Cauchy, e então segue para o passo de Newton, ambos sem a região de confiança. Começamos calculando o passo de Cauchy da seguinte forma

$$d_u^k = -\frac{g_k^T g_k}{g_k^T B_k g_k} g_k. \quad (3.20)$$

Se $\|d_u^k\| > \Delta$, adotamos para d^k a solução aproximada para (3.1) pelo passo de Cauchy, a qual é dada por

$$d^k = -\frac{\Delta_k}{\|g_k\|} g_k, \quad (3.21)$$

Quando $\|d_u^k\| < \Delta$, podemos tentar um passo melhor que o de Cauchy. Se B_k é definida positiva, o minimizador irrestrito do modelo m_k é dado por $d^N = -B_k^{-1}g_k$, que é a direção de Newton. Se d^N é viável, isto é, $\|d^N\| \leq \Delta$, temos a solução de (3.1), então consideramos

$$d^k = d^N.$$

A trajetória em vermelho na Figura 3.1 representa a solução ótima de (3.1) variando Δ . O método Dogleg aproxima essa solução a partir da substituição da trajetória

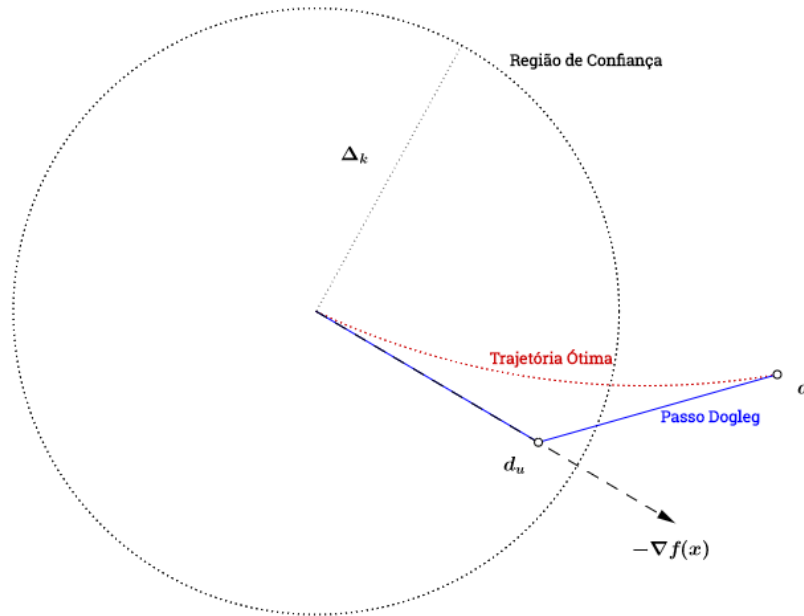


Figura 3.1: Aproximação por Dogleg

ótima pela trajetória aproximada encontrada pelo método, que é formada por dois segmentos de reta. O primeiro segmento vai da origem até o passo de Cauchy d_u^k . O segundo segmento vai de d_u^k até a direção de Newton $d_N^k = -B_k^{-1}g_k$. Esta trajetória é dada por

$$\tilde{d}(\tau) = \begin{cases} \tau d_u^k, & 0 \leq \tau \leq 1, \\ d_u^k + (\tau - 1)(d_N^k - d_u^k), & 1 \leq \tau \leq 2, \end{cases}$$

sendo $\tau \in [0, 2]$.

O método Dogleg encontra um d^k que minimiza o modelo m_k ao longo desta trajetória, sujeito a uma região de confiança. O Lema 3.3.1 mostra que o minimizador ao longo desta trajetória pode ser facilmente encontrado.

LEMA 3.3.1. *Seja B definida positiva, então*

- $\|\tilde{d}(\tau)\|$ é uma função crescente de τ , e
- $m(\tilde{d}(\tau))$ é uma função decrescente de τ .

Demonstração. J. Nocedal and S. J. Wright, página 75. \square

A convergência do método é garantida pelo seguinte teorema.

TEOREMA 3.3.2. *Supondo que d^k satisfaz (3.10), então*

$$m_k(0) - m_k(d^k) \geq m_k(0) - m_k(d_C^k) \geq c_1 \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}. \quad (3.22)$$

Demonstração. J. Nocedal and S. J. Wright, página 79. \square

O método Dogleg é apresentado no Algoritmo 5.

Algoritmo 5 Método Dogleg

- 1: Dados $x^0 \in \mathbb{R}^n$ e $\Delta_k > 0$
 - 2: $d_u^k = -\frac{g_k^T g_k}{g_k^T B_k g_k} g_k$
 - 3: **se** $\|d_u^k\| > \Delta_k$ **então**
 - 4: $d^k = -\frac{\Delta_k}{\|g_k\|} g_k$
 - 5: **senão**
 - 6: Calcular d_N^k tal que $B_k d_N^k = -g_k$
 - 7: **se** $\|d_N^k\| \leq \Delta_k$ **então**
 - 8: $d^k = d_N^k$
 - 9: **senão**
 - 10: Encontrar $\alpha_k \in [0, 1]$ tal que $\|d_u^k + \alpha_k(d_N^k - d_u^k)\| = \Delta_k$
 - 11: $d^k = d_u^k + \alpha_k(d_N^k - d_u^k)$
 - 12: **fim se**
 - 13: **fim se**
-

Note que α_k pode ser encontrado resolvendo

$$\|d_u^k + \alpha_k(d_N^k - d_u^k)\|^2 = \Delta_k^2.$$

3.3.3 Método de Steihaug

O método proposto por Steihaug em [5] é baseado no método de Gradientes Conjugados e busca uma solução aproximada para o Subproblema (3.1).

O método de Steihaug difere do método de Gradientes Conjugados pelo fato de não requerer uma matriz B_k definida positiva e pelo acréscimo de duas condições que encerram o método quando é encontrada uma direção de curvatura negativa ao longo de B_k , pois B_k pode assumir valor negativos, ou quando a norma da aproximação for muito

grande e violar a fronteira da região de confiança. Nestes casos é retornado a interseção entre direção atual e fronteira da região de confiança. O método de Steihaug é apresentado no Algoritmo 6.

Algoritmo 6 Método de Steihaug

```

1: Dados  $z_0 = 0$ ,  $r_0 = g_k$ ,  $d_0 = -r_0$  e uma tolerância  $\epsilon > 0$ 
2: enquanto  $p_k$  não for calculado faça
3:   se  $d_j^T B_k d_j \leq 0$  então
4:     Encontrar  $\tau > 0$  tal que  $p^k = z_j + \tau d_j$  minimiza  $m_k$  em (??) e satisfaz  $\|p^k\| = \Delta$ 
5:     Retorna  $p^k$ 
6:   fim se
7:    $\alpha_j = \frac{r_j^T r_j}{d_j^T B_k d_j}$ 
8:    $z_{j+1} = z_j + \alpha_j d_j$ 
9:   se  $\|z_{j+1}\| \geq \Delta$  então
10:    Encontrar  $\tau \geq 0$  tal que  $p^k = z_j + \tau d_j$  satisfaz  $\|p^k\| = \Delta$ 
11:    Retorna  $p^k$ 
12:   fim se
13:    $r_{j+1} = r_j + \alpha_j B_k d_j$ 
14:   se  $\|r_{j+1}\| < \epsilon \|r_0\|$  então
15:     Retorna  $p^k = z_{j+1}$ 
16:   fim se
17:    $\beta_{j+1} = \frac{r_{j+1}^T r_{j+1}}{r_j^T r_j}$ 
18:    $d_{j+1} = -r_{j+1} + \beta_{j+1} d_j$ 
19: fim enquanto

```

O método Dogleg, apresentado na seção anterior, começa com o passo de Cauchy, envolvendo apenas um produto de matriz vetor, em seu pior caso requer que seja resolvido um sistema linear. O método de Steihaug, também começa com o passo de Cauchy, então para casos em que Δ_k seja suficientemente pequeno, os dois métodos se equiparam. No entanto, o método de Steihaug melhora passo a passo, porém no pior caso requer a solução do mesmo sistema que o método Dogleg.

A solução aproximada p^k encontrada pelo método, satisfaz a condição (3.22) do Teorema 3.3.2, logo o método de Região de Confiança com o passo calculado pelo método de Steihaug converge. O método possui ainda uma propriedade interessante, oriunda da inicialização $z_0 = 0$, onde é garantido que cada iterando z_j possui menor norma que seu antecessor de modo que nenhum iterando posterior dará uma redução maior do modelo m_k dentro da região de confiança. Essa propriedade é assegurada pelo teorema seguinte.

TEOREMA 3.3.3. *A sequência de vetores z_j gerada pelo Algoritmo 6 satisfaz*

$$0 = \|z_0\| < \dots < \|z_j\| < \|z_{j+1}\| < \dots < \|p^k\| \leq \Delta_k.$$

Demonstração. J. Nocedal and S. J. Wright, página 172.

□

4 Testes práticos

Neste trabalho foram implementadas dois algoritmos, em linguagem Julia, sendo mesclados alguns métodos apresentados anteriormente, e para realizar os testes utilizamos os problemas da biblioteca CUTEst. O computador utilizado para realização dos testes foi um Lenovo G400s, com processador Intel[®] Core[™] i7-3612QM.

O primeiro algoritmo, chamado de `newton_stg`, é a implementação do método de Newton com região de confiança, e com o subproblema resolvido pelo método de Steihaug. Para o método de Steihaug foi definido para o número máximo de iterações como sendo 10 vezes o tamanho da norma do gradiente.

O segundo algoritmo, chamado de `lbfgs_dog`, é a implementação do método LBFGS com região de confiança e o subproblema resolvido pelo método Dogleg. O método LBFGS é utilizado por meio do pacote `LinearOperators`. Para o método Dogleg, foi utilizada a função `cg` do pacote `Krylov`, que devolve a solução do sistema linear $Bd = -g$.

Em ambos os algoritmos foram utilizados $\eta_1 = 0.25$, $\eta_2 = 0.75$, $\sigma_1 = 0.5$, $\sigma_2 = 2.0$ e $\Delta_0 = \min(\max(0.1 \|g\|, 1), 100)$. O máximo de iterações foi definido como 1000 iterações, e o tempo máximo para resolução dos problemas selecionados do CUTEst foi definido como 30 segundos.

Para os testes realizados temos uma quantidade saída, que refere-se à forma que o algoritmo encerrou as iterações. Sendo estas de acordo com o fato de atingir o número máximo de iterações estipuladas ou atingir o tempo limite para a resolução de cada problema selecionado, ou encontrou uma solução aproximada para os mesmos.

Saída	
0	uma aproximação da solução foi encontrada
1	atingiu o número máximo de iterações
2	atingiu o tempo máximo
3	o algoritmo falhou

Tabela 4.1: Saídas dos algoritmos.

A biblioteca CUTEst possui 173 problemas irrestritos que foram utilizados para testar os algoritmos implementados. Para dar uma ideia de como os testes foram realizados, vemos nas tabelas (4.2) e (4.3) os resultados dos testes com problemas irrestritos de até duas variáveis, totalizando 28 problemas. As tabelas com os resultados dos testes

realizados podem ser encontradas no Apêndice A deste trabalho.

Problema	$f(x)$	$\ g(x)\ $	k	Saída	Avaliações	t
AKIVA	6.2e+00	3.7e-04	5	0	17	0.23680377
BEALE	6.5e-09	3.9e-04	7	0	23	0.00342989
BRKMCC	1.7e-01	6.1e-06	2	0	8	0.00005507
BROWNBS	1.1e-09	5.6e+01	32	0	98	0.00039005
CLIFF	2.2e+04	6.2e+05	10	0	32	0.00021982
CUBE	1.1e-02	1.1e-01	41	0	125	0.00056982
DENSCHNA	2.2e-12	3.0e-06	5	0	17	0.00018406
DENSCHNB	1.3e-15	1.0e-07	5	0	17	0.00011396
DENSCHNC	2.7e-06	8.4e-03	8	0	26	0.00016785
DENSCHNF	2.3e-10	3.7e-04	5	0	17	0.00010204
DJTL	-7.6e+03	4.8e+06	1000	1	3005	0.04808998
EXPFIT	2.4e-01	4.9e-04	9	0	29	0.00033498
HAIRY	2.0e+01	2.3e-07	58	0	176	0.00109005
HILBERTA	1.1e-30	1.7e-15	3	0	11	0.00005794
HIMMELBB	6.3e-05	1.9e+00	19	0	59	0.00054598
HIMMELBG	3.5e-15	1.8e-07	8	0	26	0.00033808
HIMMELBH	-1.0e+00	6.0e-08	4	0	14	0.00009894
HUMPS	4.5e+03	3.2e+01	1000	1	3005	0.03094888
JENSMP	1.2e+02	3.4e-02	8	0	26	0.00079393
LOGHAIRY	6.2e+00	4.8e-03	1000	1	3005	0.02440095
MARATOSB	1.0e+00	7.8e-01	3	0	11	0.00005984
MEXHAT	2.5e+00	3.3e+04	8	0	26	0.00012898
ROSENBR	1.3e-08	7.9e-04	38	0	116	0.00088501
S308	7.7e-01	9.4e-03	13	0	41	0.00022292
SINEVAL	4.4e-17	3.5e-07	68	0	206	0.00102997
SISSER	7.0e-06	7.2e-04	8	0	26	0.00011897
SNAIL	2.9e-13	1.1e-06	108	0	326	0.00173998
ZANGWIL2	-1.8e+01	0.0e+00	2	0	8	0.00004578

Tabela 4.2: Newton-Steihaug para 28 problemas.

Problema	$f(x)$	$\ g(x)\ $	k	Saída	Avaliações	t
AKIVA	-Inf	NaN	1	3	5	0.15582800
BEALE	5.8e-08	8.8e-04	31	0	95	0.00181890
BRKMCC	1.7e-01	1.4e-05	8	0	26	0.00030804
BROWNB	5.8e-03	1.5e-01	152	0	458	0.00666308
CLIFF	4.2e+04	1.2e+06	21	0	65	0.00651598
CUBE	1.1e-03	1.4e-01	62	0	188	0.00374293
DENSCHNA	1.0e-07	1.0e-03	6	0	20	0.00032401
DENSCHNB	7.6e-10	5.6e-05	8	0	26	0.00038815
DENSCHNC	3.9e-04	1.0e-01	26	0	80	0.00087905
DENSCHNF	6.4e-09	1.8e-03	21	0	65	0.00040698
DJTL	-9.0e+03	1.9e-02	382	0	1148	0.03065801
EXPFIT	2.4e-01	1.9e-03	32	0	98	0.00276399
HAIRY	2.0e+01	5.5e-03	48	0	146	0.00194812
HILBERTA	1.3e-08	1.8e-04	6	0	20	0.00049901
HIMMELBB	1.7e-01	1.8e+01	13	0	41	0.00045705
HIMMELBG	1.3e-12	3.7e-06	20	0	62	0.00064206
HIMMELBH	-1.0e+00	2.1e-05	12	0	38	0.00063896
HUMPS	1.2e-03	5.6e-03	373	0	1121	0.01783919
JENSMP	1.2e+02	4.1e+00	34	0	104	0.00129700
LOGHAIRY	5.8e+00	1.9e-03	1000	1	3003	0.03988600
MARATOSB	9.8e-01	7.8e+01	18	0	56	0.00025105
MEXHAT	1.4e-01	4.6e+03	20	0	62	0.00031281
ROSENBR	5.6e-08	2.2e-04	52	0	158	0.00179887
S308	7.7e-01	3.6e-03	11	0	35	0.00049305
SINEVAL	9.3e-12	2.2e-06	131	0	395	0.01165390
SISSER	5.5e-06	5.9e-04	11	0	35	0.00089288
SNAIL	8.7e-10	5.9e-05	166	0	500	0.01450586
ZANGWIL2	-1.8e+01	0.0e+00	2	0	8	0.00022101

Tabela 4.3: LBFGS-Dogleg para 28 problemas.

Para estes 28 problemas, podemos ver que o algoritmo `lbfgs_dog` encontrou a solução aproximada de 26 dos 28 problemas, sem atingir o máximo de iterações ou o tempo limite, enquanto o algoritmo `newton_steihaug` encontrou a solução aproximada de 25 dos 28 problemas. Porém se avaliarmos com relação ao tempo que cada algoritmo

levou para resolver cada um dos 28 problemas de até duas variáveis, podemos observar no perfil de desempenho, visto na Figura 4.1, que o algoritmo `newton_steihaug` se mostrou mais rápido que o algoritmo `lbfgs_dog`.

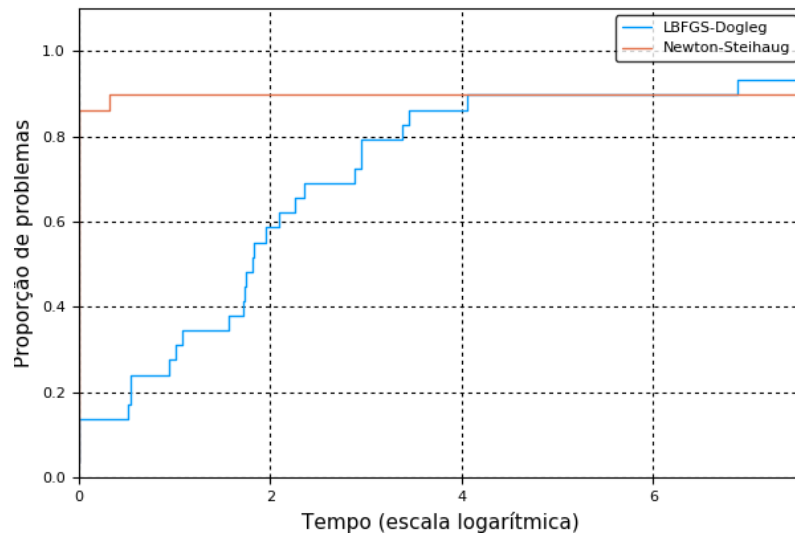


Figura 4.1: Perfil de desempenho - tempo para 28 problemas

Outro fator a ser considerado é número de avaliações de funções que cada algoritmo realiza. Para isto vamos observar a Figura 4.2. Podemos ver que ambos os algoritmos ficam próximos com relação ao número de avaliações de funções realizadas, mas o algoritmo `newton_steihaug` mostra ser capaz de realizar, na maioria dos problemas, um menor número avaliações que o algoritmo `lbfgs_dog`.

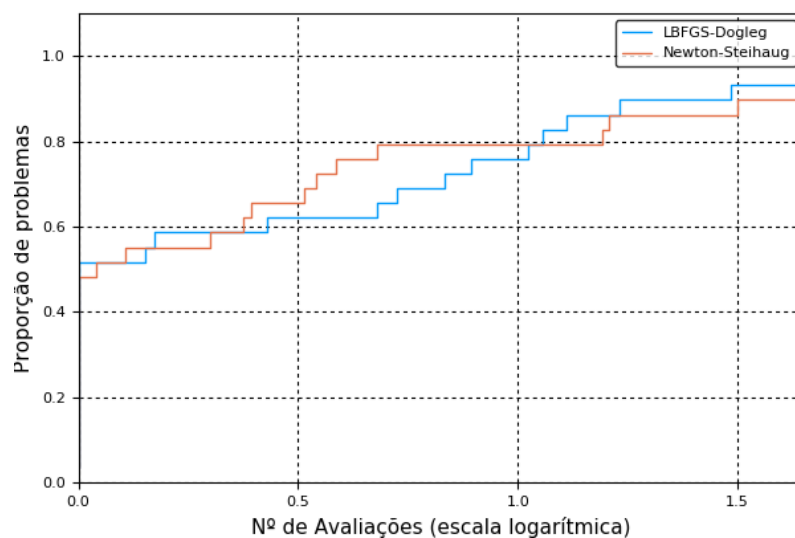


Figura 4.2: Perfil de desempenho - avaliações de função para 28 problemas

Vimos que para os 28 problemas de até duas variáveis o algoritmo `newton_stg`, mostrou um desempenho melhor que o do algoritmo `lbfgs_dog`. Agora vamos avaliar o desempenho dos algoritmos para os 173 problemas irrestritos da biblioteca CUTEst, sendo que o algoritmo `lbfgs_dog` encontrou a solução aproximada de 141 dos 173 problemas, enquanto o algoritmo `newton_steihaug` encontrou a solução aproximada de 143 dos 173 problemas.

Avaliando com relação ao tempo, podemos ver na Figura 4.3, que para cada um dos 173 problemas irrestritos `newton_steihaug` se mostrou mais rápido que o algoritmo `lbfgs_dog`.

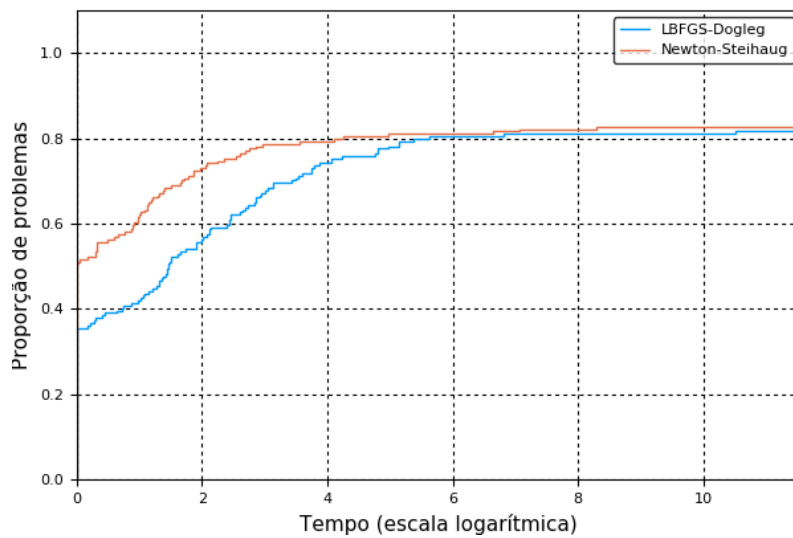


Figura 4.3: Perfil de desempenho - tempo para 173 problemas

Com relação ao número de avaliações de função realizadas por cada um dos algoritmos, podemos observar na Figura 4.4, que o algoritmo `lbfgs_dog` resolveu os problemas realizando um menor número de avaliações de função.

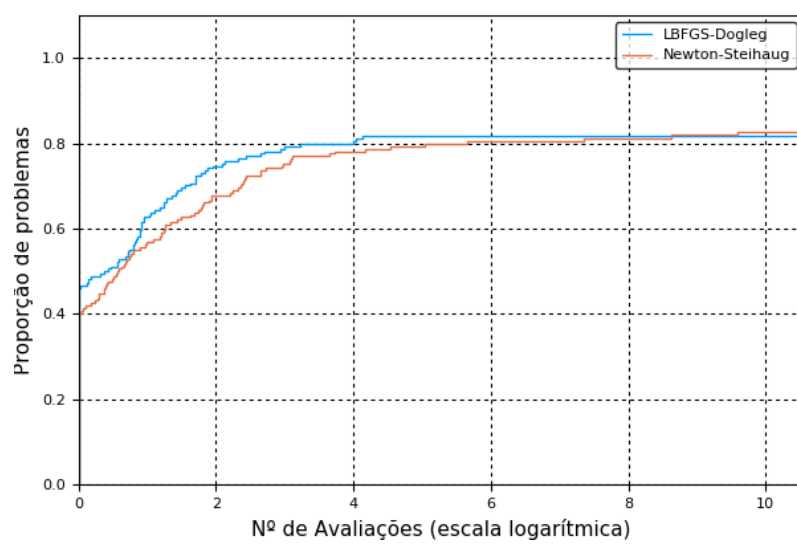


Figura 4.4: Perfil de desempenho - avaliações de função para 173 problemas

Conclusão

Nosso interesse neste trabalho foi estudar métodos de região de confiança para minimização de problemas irrestritos. Primeiramente revisamos conceitos básicos sobre minimização, passando então, aos métodos de otimização que podem utilizar a Busca Linear, uma estratégia de baixo custo computacional. E por fim estudamos os métodos de região de confiança. Por fim mesclamos alguns métodos vistos neste trabalho a fim de implementar dois algoritmos testá-los e compará-los, sendo estes algoritmos chamados de `newton_stg`: implementação do método de Newton com região de confiança, e com o subproblema resolvido pelo método de Steihaug; e `lbfgs_dog`, é a implementação do método LBFGS com região de confiança e o subproblema resolvido pelo método Dogleg. Ambos os algoritmos implementados foram testados com 173 problemas irrestritos da biblioteca CUTEst.

Para os 28 problemas de até duas variáveis da coleção CUTEst, o algoritmo `newton_stg` sem mostrou mais eficiente que o algoritmo `lbfgs_dog`, tanto no tempo que levou para resolver aproximadamente cada problema, quanto no número de avaliações de função que realiza. Ao expandir os testes para os 173 problemas irrestritos, com maiores dimensões, o algoritmo `lbfgs_dog`, ainda perde em relação ao tempo para encontrar uma solução aproximada de cada problema, porém mostra que realiza um menor número de avaliações de função do que o algoritmo `newton_stg`.

A realização da implementação dos algoritmos `newton_stg` e `lbfgs_dog` possibilitaram um aprofundamento de conhecimentos com relação a programação em linguagem Julia e utilização de seus pacotes e uma melhor visualização do funcionamento dos métodos combinados na implementação dos mesmos.

Bibliografia

- [1] J. Nocedal e S. J. Wright, *Numerical Optimization*, 2^a ed. New York: Springer, 2006.
- [2] A. A. Ribeiro e E. W. Karas, *Otimização Contínua - Aspectos Teóricos e Computacionais*, 1^a ed. São Paulo: Cengage Learning, 2013.
- [3] A. R. Conn, N. I. M. Gould e P. L. Toint, *Trust-Region Methods*. MPS-SIAM Series on Optimization, 2000.
- [4] J. J. Moré e D. C. Sorensen, *Computing a Trust Region Step*. Cern Libraries, Geneva, 1981.
- [5] T. Steihaug, *The Gradient Method and Trust Regions in Large Scale Optimization*, No.3. Vol. 20: SIAM J. NUMER. ANAL., 1983.

A Resultados

Problema	$f(x)$	$\ g(x)\ $	k	Saída	Avaliações	t
AKIVA	6.2e+00	3.7e-04	5	0	32	0.16748691
ALLINITU	5.7e+00	9.4e-07	5	0	34	0.00038815
ARGLINA	2.0e+02	1.1e-13	3	0	17	0.00613499
ARGLINB	1.0e+02	3.0e+00	1	0	9	0.00390291
ARGLINC	1.0e+02	1.9e+00	1	0	9	0.00415397
ARWHEAD	2.7e-04	2.4e+00	4	0	26	0.01531196
BARD	8.2e-03	7.8e-03	8	0	58	0.00056505
BDQRTIC	2.0e+04	1.7e+01	6	0	167	0.19648409
BEALE	6.5e-09	3.9e-04	7	0	44	0.00015211
BIGGS6	2.4e-01	9.9e-07	250	0	2060	0.01490307
BOX	-1.9e+03	1.3e-05	7	0	45	0.07917404
BOX3	2.7e-07	3.1e-04	6	0	44	0.00028419
BOXPOWER	7.8e-02	1.8e+00	2	0	18	0.02118397
BRKMCC	1.7e-01	6.1e-06	2	0	14	0.00011992
BROWNAL	2.7e+00	2.9e+01	19	0	141	0.03500295
BROWNBS	1.1e-09	5.6e+01	32	0	195	0.00097084
BROWNDEN	8.6e+04	2.0e+01	6	0	52	0.00041699
BROYDN7D	1.9e+03	4.0e-02	652	0	5956	9.51422191
BRYBND	1.6e-03	6.0e-01	38	0	411	0.52873397
CHAINWOO	1.6e+04	4.2e+01	23	0	1800	1.32968187
CHNROSNB	8.4e-04	2.1e-01	74	0	2047	0.04241014
CHNRSNBM	7.2e-03	3.9e-01	97	0	2332	0.04975605
CLIFF	2.2e+04	6.2e+05	10	0	70	0.00027299
COSINE	-1.0e+04	1.4e-05	17	0	172	0.24498010
CRAAGLVY	1.7e+03	2.3e+01	9	0	417	0.63709307
CUBE	1.1e-02	1.1e-01	41	0	247	0.00120687
CURLY10	-1.0e+06	7.4e-03	15	2	49824	69.07564282
CURLY20	-1.0e+06	3.5e+01	9	2	75907	160.09857416
CURLY30	-1.0e+06	2.5e-01	21	2	81814	231.04145408
DENSCHNA	2.2e-12	3.0e-06	5	0	32	0.00021911
DENSCHNB	1.3e-15	1.0e-07	5	0	32	0.00010800

DENSCHNC	2.7e-06	8.4e-03	8	0	50	0.00014210
DENSCHND	9.1e+02	2.3e+03	25	0	162	0.00068712
DENSCHNE	1.8e-06	2.7e-03	8	0	49	0.00013900
DENSCHNF	2.3e-10	3.7e-04	5	0	32	0.00014210
DIXMAANA	1.0e+00	3.0e-05	5	0	35	0.02230501
DIXMAANB	1.0e+00	1.2e-03	18	0	149	0.07127190
DIXMAANC	1.0e+00	2.3e-06	14	0	144	0.07439113
DIXMAAND	1.0e+00	6.6e-03	25	0	219	0.12481999
DIXMAANE	1.0e+00	7.3e-04	14	0	596	0.36262894
DIXMAANF	1.0e+00	4.5e-02	31	0	259	0.14975190
DIXMAANG	1.1e+00	2.9e-01	32	0	247	0.13033819
DIXMAANH	1.0e+00	1.3e-01	33	0	292	0.19249892
DIXMAANI	1.0e+00	4.3e-03	18	0	4910	3.11861992
DIXMAANJ	1.0e+00	1.8e-01	36	0	623	0.43221402
DIXMAANK	1.1e+00	3.1e-01	34	0	313	0.18042207
DIXMAANL	1.7e+00	6.6e-01	28	0	220	0.12744904
DIXMAANM	1.0e+00	8.9e-07	6	0	8093	5.18682003
DIXMAANN	1.0e+00	5.8e-02	27	0	393	0.26294994
DIXMAANO	1.0e+00	1.3e-01	30	0	412	0.24761915
DIXMAANP	1.1e+00	1.5e-01	28	0	371	0.20134807
DIXON3DQ	1.1e-04	3.3e-04	8	0	13086	8.77768898
DJTL	-7.6e+03	4.8e+06	1000	1	6932	0.03369093
DQDRTIC	3.3e-23	1.6e-10	2	0	16	0.00939417
DQRTIC	2.2e+12	1.0e+09	16	0	21041	6.22180200
EDENSCH	1.2e+04	1.8e+00	9	0	204	0.07554507
EG2	-1.0e+03	6.0e-09	3	0	17	0.00403690
EIGENALS	6.3e-01	7.5e+00	73	2	2846	31.52150702
EIGENBLS	7.6e-02	9.3e-01	85	2	2389	30.26519394
EIGENCLS	6.6e+01	4.0e+01	167	2	2427	30.22245789
ENGVAL1	5.5e+03	3.2e-01	5	0	91	0.07707906
ENGVAL2	7.3e-10	1.8e-03	19	0	127	0.00046897
ERRINROS	4.1e+01	7.4e+00	28	0	1444	0.01447701
ERRINRSM	3.9e+01	4.5e+00	28	0	1368	0.01611686
EXPFIT	2.4e-01	4.9e-04	9	0	55	0.00045800
EXTROSNB	7.6e-01	2.1e+00	6	0	103	0.01058888
FLETBV3M	-2.5e+05	6.5e-04	45	0	258	0.36495781

FLETGBV2	-5.0e-01	1.0e-06	4	0	6306	6.91715598
FLETGBV3	-4.0e+08	3.8e+01	1000	1	5447	5.68987083
FLETCHBV	-4.1e+16	3.8e+09	1000	1	5459	5.66398811
FLETCHCR	5.5e+02	5.6e+00	1000	1	28827	3.98923492
FMINSRF2	1.0e+00	9.6e-07	865	0	7875	14.38127708
FMINSURF	1.0e+00	1.1e-05	885	0	5988	9.91517401
FREUROTH	6.1e+05	1.1e-01	14	0	115	0.10208607
GENHUMPS	2.5e+07	3.9e+03	1000	1	5527	8.25042820
GENROSE	1.0e+00	4.1e-05	558	0	6092	0.46617699
GROWTHLS	2.2e+00	8.3e+01	20	0	124	0.00172591
GULF	1.2e-04	2.6e-03	29	0	190	0.01019311
HAIRY	2.0e+01	2.3e-07	58	0	337	0.00085306
HATFLDD	4.5e-07	1.3e-03	19	0	131	0.00142097
HATFLDE	1.0e-06	3.9e-03	20	0	136	0.00120807
HATFLDFL	6.3e-05	9.6e-07	27	0	190	0.00055122
HEART6LS	8.7e-10	8.7e-07	281	0	2617	0.02184200
HEART8LS	3.6e-10	4.8e-04	141	0	1268	0.01385307
HELIX	1.1e-05	2.8e-02	23	0	156	0.00059986
HIELOW	-Inf	NaN	2	3	14	0.08309484
HILBERTA	1.1e-30	1.7e-15	3	0	18	0.00006509
HILBERTB	9.9e-19	4.5e-09	1	0	10	0.00008512
HIMMELBB	6.3e-05	1.9e+00	19	0	116	0.00027800
HIMMELBF	3.2e+02	1.4e-01	14	0	107	0.00038815
HIMMELBG	3.5e-15	1.8e-07	8	0	46	0.00013304
HIMMELBH	-1.0e+00	6.0e-08	4	0	23	0.00016999
HUMPS	4.5e+03	3.2e+01	1000	1	5520	0.02313709
HYDC20LS	1.0e+02	6.3e+04	23	0	803	0.06446290
INDEF	-3.8e+06	1.7e+02	1000	1	5669	4.10475302
INDEFM	-9.7e+06	2.5e+03	273	2	1560	30.07751417
JENSMP	1.2e+02	3.4e-02	8	0	50	0.00028801
JIMACK	8.7e-01	1.2e-01	2	2	211	32.15866899
KOWOSB	3.1e-04	7.2e-07	9	0	69	0.00059414
LIARWHD	8.5e-04	4.7e+00	10	0	63	0.04810309
LOGHAIRY	6.2e+00	4.8e-03	1000	1	5494	0.02064085
MANCINO	2.7e+02	4.6e+04	5	0	52	0.20018888
MARATOSB	1.0e+00	7.8e-01	3	0	20	0.00007987

MEXHAT	2.5e+00	3.3e+04	8	0	55	0.00024414
MEYER3	3.9e+05	7.5e+06	13	0	86	0.00063992
MODBEALE	7.8e+00	1.4e+01	4	0	112	0.68990493
MOREBV	2.9e-11	9.9e-07	1	0	483	0.32089186
MSQRTALS	5.5e-05	2.5e-02	34	0	4284	11.13491893
MSQRTBLS	1.6e-05	2.4e-02	30	0	3210	8.38169098
NCB20	-1.5e+03	8.2e-03	82	0	871	8.92369390
NCB20B	7.4e+03	2.8e-01	19	2	2634	32.58180189
NONCVXU2	8.8e+05	3.3e+02	194	0	1135	1.91370296
NONCVXUN	8.7e+05	3.5e+02	167	0	981	1.64910817
NONDIA	1.3e-03	3.0e-02	2	0	15	0.00782704
NONDQUAR	1.2e-02	1.2e+00	8	0	43248	16.28148985
NONMSQRT	9.7e+02	1.6e+02	27	2	1678	38.66590309
OSBORNEA	8.8e-01	4.2e+02	1000	1	6008	0.05913091
OSBORNEB	4.0e-02	2.0e-05	22	0	286	0.01745391
OSCIGRAD	6.4e+00	2.0e+04	17	0	474	8.80709100
OSCIPATH	1.0e+00	1.8e-05	7	0	99	0.00039196
PALMER1C	1.1e+03	9.1e+03	1	0	14	0.00014806
PALMER1D	9.9e+02	2.2e+01	1	0	12	0.00008512
PALMER2C	2.2e-02	9.3e-01	1	0	21	0.00021195
PALMER3C	3.5e-02	6.5e-01	1	0	20	0.00011015
PALMER4C	1.4e-01	8.1e+00	1	0	20	0.00012803
PALMER5C	2.1e+00	4.6e-13	1	0	12	0.00008416
PALMER5D	8.7e+01	3.7e-12	2	0	18	0.00010991
PALMER6C	7.2e-02	1.6e-01	1	0	19	0.00011611
PALMER7C	3.9e+00	6.9e-01	1	0	20	0.00009990
PALMER8C	5.3e-01	2.5e-01	1	0	19	0.00009513
PARKCH	2.2e+03	2.5e+04	1	2	26	42.98738599
PENALTY1	2.4e+11	1.4e+09	13	0	79	0.00895977
PENALTY2	4.7e+13	6.1e+01	8	0	792	0.07391405
PENALTY3	7.7e-02	7.8e+01	17	0	466	5.98979998
POWELLSG	4.7e-01	9.5e-01	8	0	66	0.01761484
POWER	5.8e+09	6.9e+09	8	0	6721	4.75023890
QUARTC	2.2e+12	1.0e+09	16	0	21041	6.26550412
ROSENBR	1.3e-08	7.9e-04	38	0	230	0.00065112
S308	7.7e-01	9.4e-03	13	0	80	0.00026202

SBRYBND	6.8e+04	1.2e+07	1000	1	18060	26.09726501
SCHMVETT	-1.5e+04	3.0e-05	4	0	123	0.46189499
SCOSINE	2.1e+03	3.9e+06	1000	1	6422	4.46919990
SCURLY10	-6.0e+05	7.4e+07	176	2	24398	31.23577809
SCURLY20	-5.7e+05	1.5e+08	154	2	15505	30.21536517
SCURLY30	-5.4e+05	4.9e+08	169	2	11435	30.11310911
SENSORS	-2.0e+03	2.0e-03	21	0	147	0.42989206
SINEVAL	4.4e-17	3.5e-07	68	0	406	0.00147104
SINQUAD	-6.8e+06	8.2e-03	23	0	133	0.14702010
SISSER	7.0e-06	7.2e-04	8	0	50	0.00013399
SNAIL	2.9e-13	1.1e-06	108	0	648	0.00215316
SPARSINE	4.6e+00	1.8e+02	23	0	483	0.72463417
SPARSQUR	3.3e+01	7.4e+01	8	0	858	2.41805720
SPMSRTLS	1.6e-07	6.6e-04	19	0	589	0.79056883
SROSENBR	1.3e-04	4.8e-01	10	0	62	0.02031994
SSBRYBND	2.2e+04	3.2e+04	1000	1	15045	22.02918601
SSCOSINE	6.8e+02	8.3e+03	1000	1	5439	3.82822704
STRATEC	-1.3e+06	3.0e+06	12	2	70	30.60889792
TESTQUAD	9.6e-15	6.9e-07	1	0	1678	0.59036493
TOINTGOR	1.4e+03	3.5e-03	5	0	283	0.00337291
TOINTGSS	1.0e+01	6.1e-03	16	0	110	0.13765216
TOINTPSP	2.3e+02	1.9e-05	50	0	367	0.00614810
TOINTQOR	1.2e+03	7.0e-07	1	0	36	0.00041604
TQUARTIC	2.3e-13	5.2e-05	9	0	57	0.02610803
TRIDIA	1.9e-16	9.6e-07	1	0	816	0.28820801
VARDIM	7.5e+10	9.4e+11	8	0	66	0.00242209
VAREIGVL	3.0e-07	3.8e-03	18	0	351	0.01261806
VIBRBEAM	8.2e+00	3.1e+03	65	0	407	0.02502394
WATSON	2.4e-05	1.4e-02	20	0	208	0.00773406
WOODS	7.9e+03	4.9e+01	7	0	57	0.03319311
YATP1LS	1.4e+01	3.0e+01	34	0	206	0.41527104
YATP2LS	6.8e+02	2.0e+01	114	2	19015	32.72676802
YFITU	1.9e-07	1.6e-01	70	0	489	0.00315189
ZANGWIL2	-1.8e+01	0.0e+00	2	0	12	0.00005603

Tabela A.1: Newton-Steihaug para 173 problemas.

Problema	$f(x)$	$\ g(x)\ $	k	Saída	Avaliações	t
AKIVA	-Inf	NaN	1	3	5	0.15582800
ALLINITU	5.7e+00	2.4e-04	27	0	83	0.45193982
ARGLINA	2.0e+02	1.6e-13	3	0	11	0.00250006
ARGLINB	1.0e+02	1.4e-02	5	0	17	0.00341392
ARGLINC	1.0e+02	1.5e-02	5	0	17	0.00362015
ARWHEAD	5.0e-06	1.8e-02	15	0	47	0.04427385
BARD	9.9e-03	3.3e-03	9	0	29	0.00067091
BDQRTIC	2.0e+04	1.5e+02	92	0	278	0.53943419
BEALE	5.8e-08	8.8e-04	31	0	95	0.00119209
BIGGS6	5.7e-03	1.2e-05	80	0	242	0.00707197
BOX	-1.9e+03	3.3e-03	264	0	794	3.53407884
BOX3	2.5e-06	1.9e-04	8	0	26	0.00044918
BOXPOWER	6.8e-01	9.6e+00	37	0	113	0.61140418
BRKMCC	1.7e-01	1.4e-05	8	0	26	0.00035191
BROWNAL	1.9e-03	1.1e+01	31	0	95	0.01810002
BROWNBS	5.8e-03	1.5e-01	152	0	458	0.01416993
BROWNDEN	8.6e+04	1.4e+02	50	0	152	0.00537992
BROYDN7D	2.7e+03	8.5e+00	1000	1	3003	11.15593600
BRYBND	7.0e-03	7.3e-01	35	0	107	0.15179801
CHAINWOO	2.4e+01	3.9e+01	47	0	143	0.21894503
CHNROSNB	6.4e-03	3.4e-01	408	0	1226	0.07181716
CHNRSNBM	4.5e-02	3.2e-01	787	0	2363	0.12206078
CLIFF	4.2e+04	1.2e+06	21	0	65	0.00134397
COSINE	-1.0e+04	5.3e-03	30	0	92	0.20998812
CRAGGLVY	1.7e+03	2.1e+01	60	0	182	0.39692307
CUBE	1.1e-03	1.4e-01	62	0	188	0.00231695
CURLY10	-1.0e+06	1.8e+00	1000	1	3003	19.37670302
CURLY20	-1.0e+06	6.1e+00	1000	1	3003	20.47419906
CURLY30	-1.0e+06	3.3e+01	1000	1	3003	21.27369714
DENSCHNA	1.0e-07	1.0e-03	6	0	20	0.00064898
DENSCHNB	7.6e-10	5.6e-05	8	0	26	0.00093484
DENSCHNC	3.9e-04	1.0e-01	26	0	80	0.00167704
DENSCHND	1.0e+04	5.7e+03	45	0	137	0.00463605
DENSCHNE	1.0e+00	1.5e-02	11	0	35	0.00067902

DENSCHNF	6.4e-09	1.8e-03	21	0	65	0.00069594
DIXMAANA	1.0e+00	2.5e-02	8	0	26	0.01925111
DIXMAANB	1.0e+00	1.8e-01	19	0	59	0.06245995
DIXMAANC	1.0e+00	1.3e-01	17	0	53	0.04871392
DIXMAAND	1.1e+00	4.8e-01	22	0	68	0.06853819
DIXMAANE	1.1e+00	9.2e-02	37	0	113	0.18612599
DIXMAANF	1.1e+00	1.5e-01	16	0	50	0.05729103
DIXMAANG	1.1e+00	2.4e-01	23	0	71	0.07775092
DIXMAANH	1.2e+00	3.9e-01	14	0	44	0.03918505
DIXMAANI	1.3e+00	9.6e-02	49	0	149	0.25266385
DIXMAANJ	1.1e+00	1.5e-01	24	0	74	0.09793615
DIXMAANK	1.2e+00	2.7e-01	16	0	50	0.06233501
DIXMAANL	1.9e+00	6.4e-01	13	0	41	0.03361607
DIXMAANM	1.1e+00	4.2e-02	53	0	161	0.30639911
DIXMAANN	1.0e+00	8.8e-02	25	0	77	0.12552905
DIXMAANO	1.2e+00	1.7e-01	17	0	53	0.07750797
DIXMAANP	1.4e+00	2.9e-01	15	0	47	0.05523396
DIXON3DQ	6.8e-03	9.3e-03	1000	1	3003	17.70584202
DJTL	-9.0e+03	1.9e-02	382	0	1148	0.02277303
DQDRTIC	5.7e+00	8.2e+00	26	0	80	0.10240197
DQRTIC	3.0e+11	2.9e+08	17	0	53	0.06741714
EDENSCH	1.2e+04	9.3e+00	28	0	86	0.05694699
EG2	-1.0e+03	1.3e-06	10	0	32	0.00734591
EIGENALS	4.1e+01	3.4e+00	1000	1	3003	15.81127214
EIGENBLS	4.3e-02	1.2e-01	1000	1	3003	15.80159903
EIGENCLS	2.3e+02	5.7e+01	1000	1	3003	17.04024100
ENGVAL1	5.5e+03	6.3e-01	22	0	68	0.09759092
ENGVAL2	6.1e-05	3.6e-02	69	0	209	0.00415397
ERRINROS	4.3e+01	7.5e+00	26	0	80	0.00308895
ERRINRSM	5.1e+01	1.2e+01	19	0	59	0.00168610
EXPFIT	2.4e-01	1.9e-03	32	0	98	0.00144601
EXTROSNB	3.8e-01	2.9e+00	32	0	98	0.03894305
FLETBV3M	-2.5e+05	3.6e-03	49	0	149	0.39137602
FLETBVB2	-5.0e-01	1.2e-04	1000	1	3003	9.77825904
FLETBVB3	-2.9e+14	1.6e+02	1000	1	3003	8.17477107
FLETCHBV	-9.9e+24	1.1e+11	1000	1	3003	10.54185891

FLETCHCR	9.4e+02	5.6e+00	1000	1	3003	1.78366208
FMINSRF2	1.0e+00	3.3e-05	305	0	917	3.15173697
FMINSURF	1.0e+00	3.2e-05	434	0	1304	4.60164189
FREUROTH	6.1e+05	4.2e+00	162	0	488	1.12905312
GENHUMPS	1.5e+04	3.4e+01	1000	1	3003	9.66962409
GENROSE	2.5e+02	8.1e+00	1000	1	3003	1.06330609
GROWTHLS	1.3e+01	9.5e+01	36	0	110	0.00204921
GULF	3.1e-04	1.5e-03	29	0	89	0.00549102
HAIRY	2.0e+01	5.5e-03	48	0	146	0.00200081
HATFLDD	5.1e-06	6.0e-04	11	0	35	0.00107908
HATFLDE	1.1e-04	1.4e-03	14	0	44	0.00153613
HATFLDFL	6.7e-05	9.9e-06	11	0	35	0.00024605
HEART6LS	8.0e+00	4.3e+00	1000	1	3003	0.09311795
HEART8LS	2.2e-01	4.6e-01	1000	1	3003	0.11658812
HELIX	5.9e-03	1.6e-01	96	0	290	0.01275516
HIELOW	1.0e+03	8.8e+03	1000	1	3003	2.18689609
HILBERTA	1.3e-08	1.8e-04	6	0	20	0.00068116
HILBERTB	4.1e-07	3.1e-03	14	0	44	0.00084019
HIMMELBB	1.7e-01	1.8e+01	13	0	41	0.00030518
HIMMELBF	3.2e+02	6.1e-01	285	0	857	0.03467202
HIMMELBG	1.3e-12	3.7e-06	20	0	62	0.00135899
HIMMELBH	-1.0e+00	2.1e-05	12	0	38	0.00061893
HUMPS	1.2e-03	5.6e-03	373	0	1121	0.01879406
HYDC20LS	1.4e+03	5.0e+04	344	0	1034	0.19048715
INDEF	-	-	-	3	-	-
INDEFM	-9.8e+06	4.8e+03	106	2	321	30.13111115
JENSMP	1.2e+02	4.1e+00	34	0	104	0.00274777
JIMACK	1.0e+00	4.5e-01	103	2	312	30.23456001
KOWOSB	3.1e-04	8.9e-06	74	0	224	0.00874805
LIARWHD	6.7e-03	1.6e-01	47	0	143	0.17531896
LOGHAIRY	5.8e+00	1.9e-03	1000	1	3003	0.06272292
MANCINO	2.4e+03	1.3e+05	21	0	65	0.20024109
MARATOSB	9.8e-01	7.8e+01	18	0	56	0.00027299
MEXHAT	1.4e-01	4.6e+03	20	0	62	0.00036502
MEYER3	3.0e+06	6.5e+05	42	0	128	0.00220585
MODBEALE	3.8e+02	3.1e+01	598	0	1796	28.38390112

MOREBV	1.1e-10	1.1e-06	390	0	1172	3.03699088
MSQRTALS	1.3e-02	2.5e-02	402	0	1208	1.60342503
MSQRTBLS	1.5e-02	3.1e-02	444	0	1334	1.76089811
NCB20	-1.4e+03	2.3e-02	402	0	1208	7.15749192
NCB20B	7.4e+03	2.1e-02	72	0	218	1.24724317
NONCVXU2	3.4e+05	3.2e+02	109	0	329	0.96970987
NONCVXUN	4.2e+05	3.1e+02	97	0	293	0.85306311
NONDIA	1.2e-02	3.7e+01	9	0	29	0.01661491
NONDQUAR	1.3e-01	7.9e-01	18	0	56	0.05660796
NONMSQRT	8.9e+02	2.5e+01	993	2	2982	30.02874303
OSBORNEA	1.2e-04	2.8e-02	268	0	806	0.02709794
OSBORNEB	4.0e-02	5.8e-04	349	0	1049	0.05846310
OSCIGRAD	2.2e+02	1.8e+05	31	0	95	4.59730911
OSCIPTH	1.0e+00	7.0e-05	36	0	110	0.00131392
PALMER1C	4.7e+04	4.9e+04	8	0	26	0.00020289
PALMER1D	3.9e+04	2.1e+03	21	0	65	0.00060797
PALMER2C	1.4e+03	5.8e+02	21	0	65	0.00061584
PALMER3C	2.7e+02	1.5e+02	21	0	65	0.00055909
PALMER4C	2.6e+02	2.0e+02	21	0	65	0.00054312
PALMER5C	2.1e+00	5.2e-02	30	0	92	0.00181389
PALMER5D	8.7e+01	1.3e-01	98	0	296	0.00711703
PALMER6C	1.8e+02	2.0e+01	59	0	179	0.00727201
PALMER7C	2.6e+02	2.8e+02	43	0	131	0.00431514
PALMER8C	1.7e+02	8.6e+01	21	0	65	0.00088000
PARKCH	2.2e+03	2.5e+04	809	2	2430	30.00289893
PENALTY1	2.5e+11	1.4e+09	16	0	50	0.01621199
PENALTY2	4.7e+13	1.4e+03	55	0	167	0.02606797
PENALTY3	4.1e-01	1.6e+02	95	0	287	0.83422494
POWELLSG	1.5e-01	1.3e+00	42	0	128	0.17867112
POWER	8.9e+09	3.9e+09	13	0	41	0.09426904
QUARTC	3.0e+11	2.9e+08	17	0	53	0.06350207
ROSENBR	5.6e-08	2.2e-04	52	0	158	0.00459385
S308	7.7e-01	3.6e-03	11	0	35	0.00042605
SBRYBND	6.9e+04	1.8e+06	1000	1	3003	14.26638103
SCHMVETT	-1.5e+04	6.3e-03	96	0	290	0.86957717
SCOSINE	2.4e+03	1.9e+06	1000	1	3003	29.90210509

SCURLY10	-2.6e+05	1.8e+08	476	2	1431	30.07315302
SCURLY20	-2.4e+05	3.6e+08	327	2	984	30.09923100
SCURLY30	-2.4e+05	1.3e+09	246	2	741	30.08814192
SENSORS	-2.1e+03	4.2e-03	39	0	119	0.38642907
SINEVAL	9.3e-12	2.2e-06	131	0	395	0.00830913
SINQUAD	-6.8e+06	2.5e-01	158	0	476	0.89640403
SISSER	5.5e-06	5.9e-04	11	0	35	0.00076389
SNAIL	8.7e-10	5.9e-05	166	0	500	0.01597404
SPARSINE	4.1e+01	2.7e+02	123	0	371	1.25001502
SPARSQUR	4.2e+01	8.8e+01	21	0	65	0.15207601
SPMSRTLS	1.9e-04	5.5e-03	161	0	485	1.43692207
SROSENBR	5.8e-04	5.6e-01	30	0	92	0.08632016
SSBRYBND	1.7e+04	7.1e+03	1000	1	3003	8.31454897
SSCOSINE	3.7e+02	3.1e+03	1000	1	3003	11.98205996
STRATEC	2.3e+03	2.6e+02	869	2	2610	30.02858996
TESTQUAD	4.2e+03	3.3e+03	88	0	266	0.75098205
TOINTGOR	1.4e+03	5.4e-02	141	0	425	0.05467796
TOINTGSS	1.0e+01	1.2e-02	18	0	56	0.08220696
TOINTPSP	2.3e+02	6.4e-03	212	0	638	0.03363395
TOINTQOR	1.2e+03	1.6e-02	46	0	140	0.00990415
TQUARTIC	1.9e-05	1.2e-04	50	0	152	0.18014193
TRIDIA	4.7e+00	3.0e+01	698	0	2096	6.66867995
VARDIM	7.5e+10	9.4e+11	12	0	38	0.00355816
VAREIGVL	3.5e-07	2.0e-03	30	0	92	0.00888491
VIBRBEAM	8.9e+02	5.9e+04	239	0	719	0.04310083
WATSON	2.0e-04	2.0e-02	186	0	560	0.03134608
WOODS	7.9e+03	4.8e+01	20	0	62	0.06787896
YATP1LS	4.0e+02	8.5e+01	28	0	86	0.15345907
YATP2LS	2.4e+06	1.2e+00	401	0	1205	2.47324204
YFITU	3.4e+00	3.2e-01	281	0	845	0.02203393
ZANGWIL2	-1.8e+01	0.0e+00	2	0	8	0.00014901

Tabela A.2: LBFGS-Dogleg para 173 problemas.