UNIVERSIDADE FEDERAL DO PARANÁ
EGMARA ANTUNES DOS SANTOS

ASPECTOS TEÓRICOS E COMPUTACIONAIS DE UM MÉTODO DE LAGRANGIANO AUMENTADO COM SUBPROBLEMAS DE CAIXA RESOLVIDOS POR REGIÃO DE CONFIANÇA

Curitiba
2020

UNIVERSIDADE FEDERAL DO PARANÁ
EGMARA ANTUNES DOS SANTOS

ASPECTOS TEÓRICOS E COMPUTACIONAIS DE UM MÉTODO DE LAGRANGIANO AUMENTADO COM SUBPROBLEMAS DE CAIXA RESOLVIDOS POR REGIÃO DE CONFIANÇA

Dissertação apresentada ao Programa de Pós-Graduação em Matemática do Setor de Ciências Exatas da Universidade Federal do Paraná, como requisito parcial à obtenção do grau de Mestre em Matemática.
Orientadora: Drª. Elizabeth Wegner Karas.
Coorientador: Dr. Abel Soares Siqueira.

Curitiba
2020

Catalogação na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

S237a Santos, Egmara Antunes dos
Aspectos teóricos e computacionais de um método de lagrangiano aumentado com subproblemas de caixa resolvidos por região de confiança [recurso eletrônico] / Egmara Antunes dos Santos. – Curitiba, 2020.

Dissertação - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Matemática, 2020.

Orientadora: Elizabeth Wegner Karas.
Coorientador: Abel Soares Siqueira.

1. Algoritmos computacionais. 2. Algoritmos. I. Universidade Federal do Paraná. II. Karas, Elizabeth Wegner. III. Siqueira, Abel Soares. IV. Título.

CDD: 006

Bibliotecária: Vanusa Maciel CRB- 9/1928

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **EGMARA ANTUNES DOS SANTOS** intitulada: **ASPECTOS TEÓRICOS E COMPUTACIONAIS DE UM MÉTODO DE LAGRANGIANO AUMENTADO COM SUBPROBLEMAS DE CAIXA RESOLVIDOS POR REGIÃO DE CONFIANÇA**, sob orientação da Profa. Dra. ELIZABETH WEGNER KARAS, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 26 de Agosto de 2020.

Assinatura Eletrônica
26/08/2020 16:42:56.0
ELIZABETH WEGNER KARAS
Presidente da Banca Examinadora (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica
26/08/2020 17:18:37.0
LUIZ CARLOS MATIOLI
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica
26/08/2020 17:27:29.0
LEANDRO DA FONSECA PRUDENTE
Avaliador Externo (UNIVERSIDADE FEDERAL DE GOIÁS)

Documento assinado eletronicamente de acordo com o disposto na legislação federal Decreto 8539 de 08 de outubro de 2015.
Gerado e autenticado pelo SIGA-UFPR, com a seguinte identificação única: 50822
Para autenticar este documento/assinatura, acesse https://www.prppg.ufpr.br/siga/visitante/autenticacaoassinaturas.jsp e insira o codigo 50822

"Nossas dúvidas são traidoras e quase sempre fazem que percamos o bem que poderíamos ganhar pelo temor de experimentá-lo."

William Shakespeare

"A vida é demasiado curta para nos permitir interessar-nos por todas as coisas, mas é bom que nos interessemos por tantas quantas forem necessárias para preencher os nossos dias."

Bertrand Russell

Agradecimentos

Agradeço, em especial, André, pelo amor e amizade em todos os momentos.

Aos meus familiares, pelo apoio.

Aos meus orientadores Elizabeth e Abel, pelo apoio, dedicação, ensinamentos e paciência em me orientar neste trabalho.

Aos meus amigos para toda a vida, Luciano, Bruno e Camila, pela amizade verdadeira.

Ao professor Roberto, pelos conselhos durante o curso de mestrado.

À Dominique Orban, pelas referências, pela inspiração e por disponibilizar seus projetos computacionais em parceria com meu coorientador, os quais foram essenciais para o desenvolvimento prático do meu trabalho.

Aos membros da banca e ao Programa de Pós-Graduação em Matemática da Universidade Federal do Paraná.

E finalmente, à CAPES pelo financiamento deste projeto.

RESUMO

Estudamos um método de Lagrangiano aumentado para resolver problemas com restrições de igualdade e caixa. Neste método os subproblemas são resolvidos por meio de uma abordagem de região de confiança para problemas com somente restrições de caixa, proposta por Lin e Moré, 1999. O desenvolvimento teórico desta proposta é feito supondo problemas de minimização de funções diferenciáveis em conjuntos convexos. Analisamos como o resultado de convergência global é estabelecido para os subproblemas e estudamos algumas propriedades geométricas do algoritmo proposto por Lin e Moré. Por fim, implementamos o método de Lagrangiano aumentado e realizamos testes numéricos para problemas da biblioteca CUTEst, comparando o desempenho do método com o algoritmo para otimização não linear, IPOPT, desenvolvido Wachter e Biegle, 2006.

Palavras-chaves: restrições de caixa, Lagrangiano aumentado, região de confiança, métodos computacionais.

ABSTRACT

We study an augmented Lagrangian method for solving equality-and-box constrained problems. In this method, the subproblems are solved by using a trust region approach for box constrained problems, proposed by Lin and Moré, 1999, which consists of minimizing differentiable functions onto convex sets. We prove global convergence results for the subproblems and study some geometric properties of the algorithm proposed by Lin and Moré. Finally, we implement the augmented Lagrangian method and perform numerical tests using the problems of the CUTEst library, comparing the performance of the algorithm with the nonlinear optimization solver, IPOPT, developed by Wachter e Biegle, 2006.

Keywords: bounded subproblems, augmented Lagrangian, trust region, computational methods.

Sumário

Introdução **6**

1 Método de Região de Confiança para problemas de minimização com restrições de caixa **8**
1.1 Definições e resultados preliminares 8
1.2 Método do gradiente projetado 14
1.2.1 Algoritmo básico do método do gradiente projetado 18
1.3 Método de região de confiança para problemas irrestritos 19
1.3.1 Estratégia de atualização no método de região de confiança 20
1.3.2 Algoritmo básico do método de região de confiança irrestrito 21
1.4 Método de região de confiança para problemas restritos 21
1.4.1 Passo de Cauchy 22
1.4.2 Algoritmo básico do método de região de confiança restrito 27
1.4.3 Convergência global 27
1.4.4 Propriedades geométricas do algoritmo de região de confiança restrito 39

2 Método de Lagrangiano aumentado para problemas com restrições de igualdade e de caixa **44**
2.1 Condições de otimalidade 45
2.1.1 Condições necessárias de primeira ordem 46
2.1.2 Condições necessárias e suficientes de segunda ordem 48
2.2 Método de Lagrangiano aumentado 49
2.2.1 Penalização quadrática para problemas com restrições de igualdade 49
2.2.2 Método clássico de Lagrangiano aumentado 50
2.2.3 Tratamento das restrições de caixa 53

3 Experimentos numéricos **55**
3.1 Detalhes de implementação 55
3.2 Resolução do subproblema do método de Lagrangiano aumentado 56
3.2.1 Cálculo do passo de Cauchy 56
3.2.2 Cálculo da solução aproximada para o subproblema 57
3.3 Resultados numéricos 58

Conclusão **65**

A Anexos **67**

Introdução

Neste trabalho estudamos problemas de otimização não linear, diferenciáveis e restritos. Mais precisamente, concentramos nossa atenção em uma técnica de resolução para os problemas com restrições de igualdade e de caixa, isto é, problemas da forma

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && h(x) = 0, \\ & && x \in \Omega, \end{aligned} \tag{1}$$

em que $f : \mathbb{R}^n \to \mathbb{R}$ e $h : \mathbb{R}^n \to \mathbb{R}^p$ são funções diferenciáveis e Ω é a caixa

$$\Omega = \{x \in \mathbb{R}^n | \ \ell \leq x \leq u\},$$

supondo os vetores $\ell \in (\mathbb{R} \cup \{-\infty\})^n$ e $u \in (\mathbb{R} \cup \{+\infty\})^n$.

Problemas de otimização da forma (1) aparecem em muitas aplicações, tais como de otimização de desenho de estruturas (veja [1, Seção 3]). Estes problemas são de grande escala, ou seja, envolvem um grande número de variáveis e restrições. O avanço da tecnologia contribuiu para o constante estudo de algoritmos para otimização de grande escala. Em [2], é feita uma revisão completa destes algoritmos e envolve problemas irrestritos e restritos. Além disso, questões práticas a respeito dos principais softwares implementados são apresentadas nesta referência.

Para esta dissertação, nos concentramos em estudar e implementar o método de Lagrangiano aumentado, que é um dos principais métodos estudados atualmente. Outras abordagens, tais como, programação quadrática sequencial, métodos de pontos interiores e métodos de filtro, são bastante estudados. Recomendamos [3]–[8] para estudos avançados destas abordagens clássicas em otimização não linear. Citamos [1] e [9] para exemplos de aplicações do método de Lagrangiano aumentado.

O método de Lagrangiano aumentado busca por pontos estacionários do problema (1), isto é, pontos que satisfazem as condições de Karush-Kuhn-Tucker. Mais precisamente, o problema (1) é resolvido de forma iterativa em que, a cada iteração, um subproblema é resolvido somente com as restrições de caixa. O tratamento das restrições de igualdade é feito através da penalização da inviabilidade nessas restrições. De forma genérica, podemos expressar os subproblemas no método de Lagrangiano aumentado na forma

$$\begin{aligned} &\text{minimizar} && \mathcal{L}(x, \lambda, c) \\ &\text{sujeito a} && \ell \leq x \leq u, \end{aligned} \tag{2}$$

em que $\mathcal{L}$ é a uma função real definida em termos de $x \in \mathbb{R}^n$ e dos parâmetros fixados $\lambda \in \mathbb{R}^p$, que é uma aproximação para o multiplicador de Lagrange associado às restrições

de igualdade, e $c > 0$, que é o parâmetro de penalidade para essas restrições. Esta função é chamada de função Lagrangiana aumentada e é dada pela expressão

$$\mathcal{L}(x, \lambda, c) = f(x) + \lambda^T h(x) + \frac{c}{2}\|h(x)\|^2.$$

Neste trabalho, veremos suas principais propriedades. Além disso, veremos como a estratégia de penalização de restrições é utilizada no método e como a abordagem clássica de região de confiança pode ser utilizada na resolução destes subproblemas. Uma boa introdução sobre métodos de região de confiança e análise de convergência é feita em [10] e estudos detalhados são feitos em [11].

Os subproblemas do método de Lagrangiano aumentado possuem um formato bastante particular. Inúmeros algoritmos de otimização não linear podem ser utilizados para a sua resolução, pois a estrutura das restrições é uma vantagem para o uso de abordagens geométricas, tais como o uso de projeções e região de confiança. Dentre os principais métodos para resolução de problemas com restrições de caixa destacam-se os métodos de conjunto ativo, gradiente projetado e pontos interiores. Neste trabalho apresentamos um estudo completo a respeito da abordagem proposta por Lin e Moré em [12] para a resolução de problemas com restrições de caixa. Esta abordagem utiliza um método de região de confiança. As principais referências que utilizamos foram [12]–[16]. Portanto, destacamos como um dos objetivos deste trabalho o estudo teórico do algoritmo de Lin e Moré. Apresentamos a teoria de convergência global e as principais propriedades geométricas do método.

Uma vantagem do uso do método de Lagrangiano aumentado, é que ele pode ser aplicado a problemas em um formato bastante geral, incluindo restrições de desigualdade (veja [8]) e problemas de grande escala. Daí, a importância do estudo deste método, tanto do ponto de vista teórico, quanto do ponto de vista computacional (veja [1]). Além disso, a utilização do algoritmo de Lin e Moré para a resolução dos subproblemas do método de Lagrangiano aumentado, torna a implementação deste livre de fatoração matricial. Sendo assim, destacamos também como um dos principais objetivos desta dissertação, a implementação do algoritmo de Lagrangiano aumentado.

Como referência para o desenvolvimento computacional utilizamos o trabalho [1], que apresenta um algoritmo em linguagem de programação Python. Utilizamos para a implementação dos algoritmos a linguagem de programação Julia. Toda esta implementação utilizou como base a infraestrutura JuliaSmoothOptimizers [17] para otimização não linear, e está disponível em [18]. Os experimentos numéricos para validação desta implementação foram realizados na resolução de todos os 384 problemas restritos da coleção CUTEst [19], com até 1000 variáveis e 1000 restrições de igualdade e/ou desigualdade, além das restrições de caixa. Comparamos os resultados obtidos nestes problemas com o algoritmo proposto em [20].

Organizamos este trabalho da seguinte forma: no Capítulo 1, revisamos alguns conceitos e teoremas que utilizamos ao longo de toda esta dissertação e, em seguida, apresentamos o estudo teórico do algoritmo de Lin e Moré. No Capítulo 2, estudamos o método de Lagrangiano aumentado. Por fim, no Capítulo 3, destacamos a implementação dos principais algoritmos e os resultados numéricos obtidos.

Capítulo 1

Método de Região de Confiança para problemas de minimização com restrições de caixa

Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável e $\Omega \subset \mathbb{R}^n$ um conjunto convexo e fechado. Considere o seguinte problema de minimização restrita

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

Neste capítulo, estudamos o método de região de confiança descrito em [12], [13] e [15] para resolver o problema acima, os principais resultados da teoria de convergência global do método e propriedades geométricas. O método estudado tem como característica a adaptação de algoritmos clássicos de região de confiança de modo que, a cada iteração, o iterando permaneça no conjunto Ω. Isto é feito por meio do cálculo de projeções no conjunto viável. Quando o conjunto Ω é da forma

$$\Omega = \{x \in \mathbb{R}^n \mid \ell \leq x \leq u\}, \tag{1.2}$$

para quaisquer vetores $\ell \in \mathbb{R}^n$ e $u \in \mathbb{R}^n$, veremos que o algoritmo torna-se mais simples. Os detalhes de implementação do algoritmo para o caso em que Ω é da forma (1.2), serão descritos no Capítulo 3. Nesse caso, o conjunto Ω é chamado de *caixa*.

Primeiramente, na Seção 1.1, revisamos algumas definições e conceitos preliminares para melhor entendimento do método. Em seguida, na Seção 1.2, descrevemos o método de gradiente projetado para resolver o problema (1.1). Na Seção 1.3, estudamos o método de região de confiança para problemas irrestritos. Por fim, na Seção 1.4, apresentamos o método de região de confiança para o problema (1.1), estabelecemos a convergência global do algoritmo e algumas características geométricas do método que motivam as implementações computacionais. As principais referências para este capítulo são [6], [7], [12]–[16].

1.1 Definições e resultados preliminares

Nesta seção, descrevemos algumas definições e resultados com o objetivo de facilitar o entendimento dos teoremas e demonstrações deste capítulo. A primeira definição que

apresentamos é a de *direção viável.*

Definição 1.1. *Sejam $\Omega \subset \mathbb{R}^n$ um conjunto qualquer, $d \in \mathbb{R}^n$ e $\bar{x} \in \Omega$. Dizemos que d é uma direção viável em relação a Ω a partir de $\bar{x}$ se existir $\varepsilon > 0$ tal que*

$$\bar{x} + td \in \Omega, \ \forall \ t \in [0, \varepsilon].$$

Denotamos por $V_\Omega(\bar{x})$ o conjunto de todas as direções viáveis em relação a Ω a partir de $\bar{x}$. É fácil ver que, se d é uma direção viável, então αd também é uma direção viável, para qualquer $\alpha \geq 0$. Esta característica nos permite afirmar que o conjunto $V_\Omega(\bar{x})$ é um *cone*, de acordo com a definição a seguir.

Definição 1.2. *Seja $K \subset \mathbb{R}^n$ um conjunto qualquer não vazio. Dizemos que K é um cone se, para qualquer $t \geq 0$,*

$$d \in K \Rightarrow td \in K.$$

Na Figura 1.1 representamos uma direção viável d a partir de $\bar{x}$, para um conjunto convexo Ω.

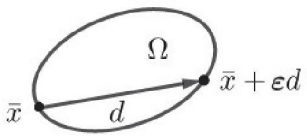

Figura 1.1: Direção viável. Fonte: a autora.

Além do conceito de direção viável, a definição de *direção tangente* é importante quando estudamos métodos para otimização com restrições. Neste trabalho, vamos utilizar a seguinte definição para uma direção tangente, como em [5].

Definição 1.3. *Sejam $\Omega \subset \mathbb{R}^n$ um conjunto qualquer, $d \in \mathbb{R}^n$ e $\bar{x} \in \Omega$. O vetor d é uma direção tangente ao conjunto Ω a partir de $\bar{x}$ quando é nula ou se existir uma sequência $(x^k) \subset \Omega$, com $x^k \to \bar{x}$, tal que*

$$\frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \frac{d}{\|d\|}.$$

A definição acima é equivalente a dizer que existem sequências $(t_k) \subset \mathbb{R}_{++}$ com $t_k \to 0_+$ e $(d^k) \subset \mathbb{R}^n$, com $d^k \to d$, tais que

$$\bar{x} + t_k d^k \in \Omega, \ \forall \ k \in \mathbb{N}. \tag{1.3}$$

De fato, caso $d = 0$, a expressão acima é válida de forma trivial. Se d é tangente e não nula, basta tomarmos

$$d^k = \frac{x^k - \bar{x}}{t_k} \text{ com } t_k = \frac{\|x^k - \bar{x}\|}{\|d\|},$$

para obtermos a relação (1.3) satisfeita. Reciprocamente, se existirem sequências (d^k) e (t_k) tais que vale (1.3) então, definimos

$$x^k = \bar{x} + t_k d^k, \ \forall \ k \in \mathbb{N}.$$

Denotamos por $T_\Omega(\bar{x})$ o conjunto de todas as direções tangentes a Ω a partir de $\bar{x}$. Este conjunto é chamado de *cone tangente*, uma vez que é um cone. Podemos dizer que as direções tangentes tangenciam ou penetram o conjunto viável. Por exemplo, para um conjunto convexo como o da Figura 1.1, ilustramos, na Figura 1.2, uma direção tangente d, a partir de $\bar{x}$.

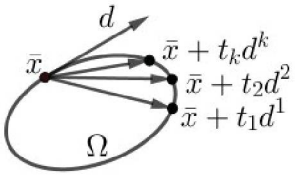

Figura 1.2: Direção tangente. Fonte: a autora.

Note que, o cone das direções viáveis está contido no cone das direções tangentes, isto é,

$$V_\Omega(\bar{x}) \subset T_\Omega(\bar{x}). \tag{1.4}$$

Além disso, quando o conjunto Ω é convexo, em [6, Teorema 3.1.8] mostra-se que o cone tangente é o fecho do cone das direções viáveis, isto é,

$$T_\Omega(\bar{x}) = \overline{V_\Omega(\bar{x})}. \tag{1.5}$$

Em [16], o cone tangente para conjuntos convexos é definido pela igualdade acima. Definimos também o *cone normal* para conjuntos convexos. Este conjunto será utilizado nas próximas seções para justificar algumas propriedades geométricas do algoritmo de região de confiança para problemas restritos.

Definição 1.4. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto convexo e $\bar{x} \in \Omega$. O cone normal a partir de $\bar{x}$ em relação a Ω é o conjunto de direções*

$$N_\Omega(\bar{x}) = \{d \in \mathbb{R}^n \mid d^T(x - \bar{x}) \leq 0, \ \forall\ x \in \Omega\}.$$

Também definimos o conceito de *projeção*, bastante importante para o desenvolvimento dos algoritmos.

Definição 1.5. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto qualquer. A projeção de $x \in \mathbb{R}^n$ sobre Ω é uma solução, quando existe, do problema*

$$\begin{aligned} &\textit{minimizar} \quad \|y - x\| \\ &\textit{sujeito a} \quad y \in \Omega. \end{aligned}$$

Da definição anterior, interpretamos a projeção de $x \in \mathbb{R}^n$ sobre o conjunto Ω, quando existe, como sendo um ponto em Ω cuja distância a x é mínima. Quando o conjunto Ω é convexo e fechado, o próximo teorema estabelece algumas propriedades da projeção.

Teorema 1.6. [7, Teorema 1.3.1] *Seja $\Omega \subset \mathbb{R}^n$ convexo e fechado. Então*

(a) Para todo $x \in \mathbb{R}^n$, a projeção de x sobre Ω existe e é única, denotada por $P_\Omega(x)$.

(b) Para todo $x \in \mathbb{R}^n$, temos que $\bar{x} = P_\Omega(x)$ se, e somente se,

$$(x - \bar{x})^T (y - \bar{x}) \leq 0, \forall\ y \in \Omega. \tag{1.6}$$

(c) Para todo $x, y \in \mathbb{R}^n$, vale

$$\|P_\Omega(x) - P_\Omega(y)\| \leq \|x - y\|.$$

(d) Para $x \in \mathbb{R}^n$ e $d \in \mathbb{R}^n$ vale que

$$0 < \alpha_1 \leq \alpha_2 \ \Rightarrow\ \|P_\Omega(x + \alpha_1 d) - x\| \leq \|P_\Omega(x + \alpha_2 d) - x\|.$$

(e) Para $x \in \mathbb{R}^n$ e $d \in \mathbb{R}^n$ vale que

$$0 < \alpha_1 \leq \alpha_2 \ \Rightarrow\ \frac{\|P_\Omega(x + \alpha_1 d) - x\|}{\alpha_1} \geq \frac{\|P_\Omega(x + \alpha_2 d) - x\|}{\alpha_2}.$$

Quando o conjunto Ω é da forma (1.2), o cálculo da projeção é feito de forma simples, isto é, dado $x \in \mathbb{R}^n$, a i-ésima coordenada da projeção de x em Ω é

$$\big(P_\Omega(x)\big)_i = \begin{cases} \ell_i, & \text{se} \quad x_i \leq \ell_i, \\ x_i, & \text{se} \quad x_i \in (\ell_i, u_i), \\ u_i, & \text{se} \quad x_i \geq u_i, \end{cases} \tag{1.7}$$

para cada $i = 1, \ldots, n$. A Figura 1.3 ilustra o cálculo da projeção para o caso $n = 2$.

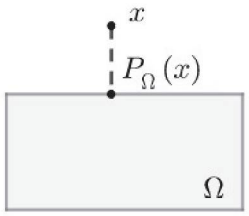

Figura 1.3: Projeção na caixa. Fonte: a autora.

Antes de estabelecermos condições de otimalidade de primeira ordem para o problema (1.1), recorde as seguintes definições.

Definição 1.7. *Um ponto $\bar{x} \in \Omega$ é um minimizador global para o problema (1.1) se*

$$f(\bar{x}) \leq f(x),$$

para todo $x \in \Omega$.

Definição 1.8. *Um ponto $\bar{x} \in \Omega$ é um minimizador local para o problema (1.1) se existir $\varepsilon > 0$ tal que*

$$f(\bar{x}) \leq f(x),$$

para todo $x \in \Omega$ com $\|x - \bar{x}\| \leq \varepsilon$.

As fórmulas de Taylor serão bastante úteis neste trabalho. Enunciamos as fórmulas de primeira e segunda ordem, demonstradas em [21, Capítulo 3, Teorema 5].

Teorema 1.9. *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função continuamente diferenciável em um aberto $U \subset \mathbb{R}^n$ e $\bar{x} \in U$. Então, para todo $d \in \mathbb{R}^n$ tal que $\bar{x} + d \in U$, vale que*

$$f(\bar{x} + d) = f(\bar{x}) + \nabla f(\bar{x})^T d + r(d), \tag{1.8}$$

em que

$$\lim_{d \to 0} \frac{r(d)}{\|d\|} = 0. \tag{1.9}$$

Teorema 1.10. *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função duas vezes continuamente diferenciável em um aberto $U \subset \mathbb{R}^n$ e $\bar{x} \in U$. Então, para todo $d \in \mathbb{R}^n$ tal que $\bar{x} + d \in U$, vale que*

$$f(\bar{x} + d) = f(\bar{x}) + \nabla f(\bar{x})^T d + \frac{1}{2} d^T \nabla^2 f(\bar{x}) d + s(d), \tag{1.10}$$

em que

$$\lim_{d \to 0} \frac{s(d)}{\|d\|^2} = 0. \tag{1.11}$$

Os limites (1.9) e (1.11) implicam que os erros $r(d)$ e $s(d)$ tendem a zero mais rápido do que $\|d\|$ e $\|d\|^2$ tendem a zero, respectivamente, ou seja, quando estamos nos aproximando de $\bar{x}$. Então, podemos reescrever as igualdades (1.8) e (1.10) como

$$f(\bar{x} + d) = f(\bar{x}) + \nabla f(\bar{x})^T d + o(\|d\|) \tag{1.12}$$

e

$$f(\bar{x} + d) = f(\bar{x}) + \nabla f(\bar{x})^T d + \frac{1}{2} d^T \nabla^2 f(\bar{x}) d + o(\|d\|^2), \tag{1.13}$$

respectivamente. Além disso, em alguns casos, é útil reescrever as fórmulas de Taylor (1.8) e (1.10) como

$$f(\bar{x} + d) = f(\bar{x}) + \nabla f(\bar{x})^T d + R(d)\|d\| \tag{1.14}$$

e

$$f(\bar{x} + d) = f(\bar{x}) + \nabla f(\bar{x})^T d + \frac{1}{2} d^T \nabla^2 f(\bar{x}) d + S(d)\|d\|^2,$$

respectivamente, sendo as funções $R(d) = r(d)/\|d\|$ e $S(d) = s(d)/\|d\|^2$ tais que

$$\lim_{\|d\| \to 0} R(d) = 0 \quad \text{e} \quad \lim_{\|d\| \to 0} S(d) = 0. \tag{1.15}$$

O lema a seguir facilita as próximas demonstrações.

Lema 1.11. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto convexo e $\bar{x} \in \Omega$. Então*

$$\{d \in \mathbb{R}^n \mid d = x - \bar{x}, \ \forall\, x \in \Omega\} \subset V_\Omega(\bar{x}).$$

Demonstração. Dados $x \in \Omega$ e $d = x - \bar{x}$ temos $x = \bar{x} + d \in \Omega$. Então, da convexidade de Ω, temos

$$\bar{x} + td \in \Omega,$$

para qualquer $t \in [0, 1]$. Logo, da Definição 1.1, segue que $d \in V_\Omega(\bar{x})$. □

Podemos demonstrar agora uma condição necessária de otimalidade para o problema (1.1).

Teorema 1.12. [6, Teorema 3.1.12] *Considere o problema (1.1). Se $\bar{x} \in \Omega$ é um minimizador local para este problema então*

$$\nabla f(\bar{x})^T(x - \bar{x}) \geq 0, \forall\ x \in \Omega. \tag{1.16}$$

Demonstração. Primeiramente, note que

$$\nabla f(\bar{x})^T d \geq 0, \forall\ d \in T_\Omega(\bar{x}).$$

De fato, para $d = 0 \in T_\Omega(\bar{x})$, a desigualdade acima é imediata. Considerando $d \in T_\Omega(\bar{x})$ não nulo, sabemos que existe uma sequência $(t_k) \subset \mathbb{R}_+$ tal que $t_k \to 0+$, e uma sequência $(d^k) \subset \mathbb{R}^n$ com $d^k \to d$, tais que vale (1.3). Como $\|d^k\| \to \|d\| > 0$, $\bar{x} + t_k d^k \to \bar{x}$ e $\bar{x}$ é minimizador local, vale que, para todo k suficientemente grande,

$$f(\bar{x}) \leq f(\bar{x} + t_k d^k) \quad \text{e} \quad \|d^k\| > 0.$$

Então, da fórmula de Taylor de primeira ordem (1.12), temos

$$0 \leq t_k \nabla f(\bar{x})^T d^k + o(t_k \|d^k\|).$$

Dividindo esta expressão por $t_k\|d^k\| > 0$ e passando o limite em k, obtemos

$$\nabla f(\bar{x})^T d \geq 0.$$

Além disso, o Lema 1.11 e a inclusão (1.4) implicam que

$$\{d \in \mathbb{R}^n \ \mid \ d = x - \bar{x},\ \forall\ x \in \Omega\} \subset T_\Omega(\bar{x}).$$

Portanto, vale a condição (1.16). □

Observe que a condição (1.16) é suficiente para $\bar{x}$ ser um minimizador global, quando f é convexa, pois, nesse caso,

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

para qualquer $x \in \Omega$.

O próximo teorema estabelece a condição necessária de otimalidade para $\bar{x}$, em termos da projeção.

Teorema 1.13. [6, Teorema 3.2.34] *Se $\bar{x} \in \Omega$ é um minimizador local do problema (1.1) então para qualquer $\alpha > 0$, temos*

$$P_\Omega\left(\bar{x} - \alpha \nabla f(\bar{x})\right) = \bar{x}. \tag{1.17}$$

Demonstração. Seja $\alpha > 0$. Então, fazendo $y = \bar{x}$ em (1.6), temos

$$\left(\bar{x} - \alpha\nabla f(\bar{x}) - P_\Omega\left(\bar{x} - \alpha\nabla f(\bar{x})\right)\right)^T \left(\bar{x} - P_\Omega\left(\bar{x} - \alpha\nabla f(\bar{x})\right)\right) \leq 0.$$

Isso implica que

$$\|\bar{x} - P_\Omega\left(\bar{x} - \alpha\nabla f(\bar{x})\right)\|^2 \leq \alpha\nabla f(\bar{x})^T \left(\bar{x} - P_\Omega\left(\bar{x} - \alpha\nabla f(\bar{x})\right)\right).$$

Usando o Teorema 1.12 para o lado direito da desigualdade acima, temos

$$\|\bar{x} - P_\Omega\left(\bar{x} - \alpha\nabla f(\bar{x})\right)\|^2 \leq 0.$$

Portanto, vale (1.17). □

A Figura 1.4 exemplifica a condição de otimalidade (1.17) em um ponto $\bar{x}$ para uma função quadrática f com matriz Hessiana definida positiva.

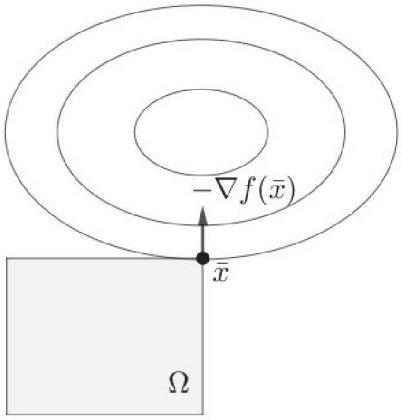

Figura 1.4: Condição necessária (1.17). Fonte: a autora.

Assim como a condição (1.16), para funções convexas, a condição (1.17) torna-se suficiente para $\bar{x}$ ser um minimizador (veja [6, Teorema 3.4.36]).

Neste capítulo, vamos considerar que um ponto $\bar{x}$ é *estacionário* para o problema (1.1) se satisfazer a condição necessária dada pelo Teorema 1.13 ou a condição (1.16), uma vez que estas condições são equivalentes quando Ω é convexo e fechado, como mostra [6, Teorema 3.4.36].

Para alguns resultados, vamos supor condições sobre a derivada da função objetivo. Uma condição bastante comum é pedir que a derivada seja *Lipschitz-contínua*.

Definição 1.14. *Dado um conjunto $\Omega \subset \mathbb{R}^n$, dizemos que uma aplicação $h : \Omega \to \mathbb{R}^p$ é Lipschitz-contínua com constante $L > 0$ se*

$$\|h(x) - h(y)\| \leq L\|x - y\|,$$

para quaisquer $x, y \in \Omega$.

Para finalizar esta seção, enunciamos um resultado que será utilizado mais adiante.

Lema 1.15. [7, Lema 1.5.4] *Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável. Suponha que f tem derivada Lipschitz-contínua com constante $L > 0$. Então, dado qualquer $x \in \mathbb{R}^n$ e qualquer direção $d \in \mathbb{R}^n$, vale que*

$$|f(x + d) - f(x) - \nabla f(x)^T d| \leq \frac{L\|d\|^2}{2}.$$

1.2 Método do gradiente projetado

O método do gradiente projetado busca de forma iterativa uma solução para o problema (1.1). A proposta deste método é garantir que as iterações se mantenham no conjunto viável. Supondo $x^0 \in \Omega$ ponto inicial, definimos, para cada $k \in \mathbb{N}$, o iterando

$$x^{k+1} = P_\Omega \left(x^k - \alpha_k \nabla f(x^k) \right), \tag{1.18}$$

em que $\alpha_k > 0$ satisfaz uma condição de decréscimo. Note que $x^k \in \Omega$ para todo $k \geq 0$. Vamos entender as condições requeridas para o passo α_k no método de gradiente projetado olhando, a cada iteração, para o caminho $d^k : \mathbb{R} \to \mathbb{R}^n$ dado por

$$d^k(\alpha) = P_\Omega \left(x^k - \alpha \nabla f(x^k)\right) - x^k. \tag{1.19}$$

Note que, quando o conjunto Ω é da forma (1.2), o cálculo da projeção é feito por meio da expressão (1.7). Neste caso, na Figura 1.5, para $n = 2$, exemplificamos uma direção da forma (1.19), considerando uma função f cuja matriz Hessiana é definida positiva.

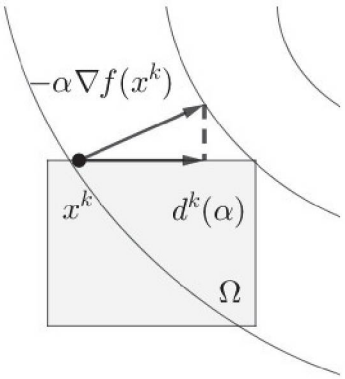

Figura 1.5: Direção (1.19). Fonte: a autora.

Queremos minimizar a função objetivo na direção (1.19). O tamanho do passo $\alpha_k > 0$ é atualizado de modo que um decréscimo suficiente para a função objetivo na direção $d^k(\alpha)$ seja satisfeito, isto é, dada uma constante $\mu \in (0, 1)$, pedimos que o passo $\alpha_k > 0$ satisfaça a condição

$$f(x^k + d^k(\alpha_k)) \leq f(x^k) + \mu \nabla f(x^k)^T d^k(\alpha_k). \tag{1.20}$$

O parâmetro μ está fixado e, como sugerido em [7], a obtenção de um valor de α_k positivo que satisfaça a condição acima pode ser feita utilizando um procedimento conhecido como *backtracking*. Nesse processo, são dados um parâmetro de redução $\theta \in (0, 1)$ e um valor $\hat{\alpha} > 0$ para o comprimento do passo. Caso este valor satisfaça a condição de decréscimo, isto é,

$$f(x^k + d^k(\hat{\alpha})) \leq f(x^k) + \mu \nabla f(x^k)^T d^k(\hat{\alpha}), \tag{1.21}$$

então escolhemos $\alpha_k = \hat{\alpha}$. Caso contrário o valor de α_k é obtido como

$$\alpha_k = \hat{\alpha}\theta^s, \tag{1.22}$$

em que s é o menor inteiro positivo tal que a condição (1.20) seja satisfeita.

Note que, quando $\Omega = \mathbb{R}^n$, a condição (1.20) se reduz à chamada *condição de Armijo* descrita em [5] pois, nesse caso, teríamos

$$f(x^k + \alpha_k d) \leq f(x^k) + \mu \alpha_k \nabla f(x^k)^T d, \tag{1.23}$$

sendo

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

Na Figura 1.6, podemos interpretar a desigualdade (1.23) para uma função diferenciável f. Destacamos no eixo horizontal os possíveis valores admissíveis para o passo α_k requerido, isto é, os valores que satisfazem a condição de Armijo (1.23). Note que, para estes valores o valor de $f(x^k + \alpha_k d)$ está abaixo ou é igual aos valores da reta tangente cujo coeficiente angular é $\alpha\mu$.

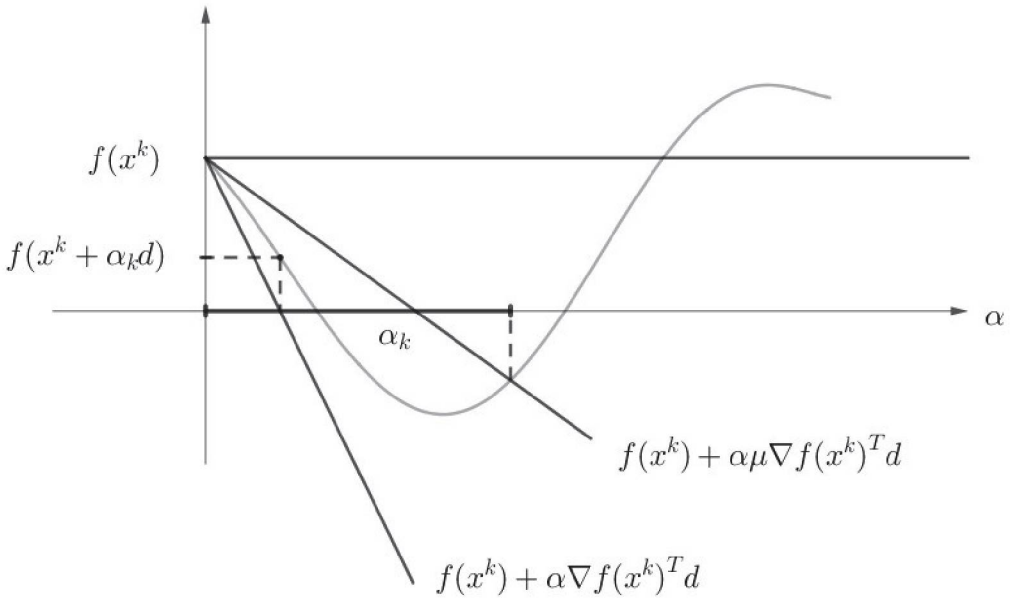

Figura 1.6: Condição de decréscimo (1.23). Fonte: a autora.

De forma semelhante à Figura 1.5, após obtermos um valor de α_k, ilustramos o iterando (1.18) do método do gradiente projetado na Figura 1.7.

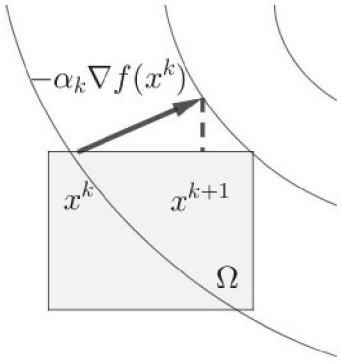

Figura 1.7: Iterando (1.18). Fonte: a autora.

Podemos garantir, sobre determinadas hipóteses, que existem valores positivos para α_k satisfazendo a condição (1.20).

Teorema 1.16. [7, Lema 4.1.1] *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável, $\Omega \subset \mathbb{R}^n$ convexo e fechado e $\mu \in (0, 1)$. Suponha que f tem derivada Lipschitz-contínua com constante $L > 0$. Então, para qualquer $x^k \in \Omega$, temos*

$$f(x^k + d^k(\alpha)) \leq f(x^k) + \mu\nabla f(x^k)^T d^k(\alpha)$$

para todo $\alpha \in (0, \bar{\alpha}]$, *sendo*

$$\bar{\alpha} = \frac{2(1-\mu)}{L} \tag{1.24}$$

e a direção $d^k(\alpha)$ *dada pela expressão (1.19).*

Demonstração. Dado $\alpha > 0$, escrevemos

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

Distribuindo os termos desta igualdade, temos

$$\nabla f(x^k)^T d^k(\alpha) = \frac{1}{\alpha} x^{kT} d^k(\alpha) - \frac{1}{\alpha}\big(x^k - \alpha \nabla f(x^k)\big)^T d^k(\alpha).$$

Defina $x = x^k - \alpha \nabla f(x^k)$. Somando e subtraindo o termo

$$\frac{1}{\alpha} P_\Omega(x)^T d^k(\alpha)$$

do lado direito, ficamos com

$$\nabla f(x^k)^T d^k(\alpha) = \frac{1}{\alpha} x^{kT} d^k(\alpha) - \frac{1}{\alpha} x^T d^k(\alpha) + \frac{1}{\alpha} P_\Omega(x)^T d^k(\alpha) - \frac{1}{\alpha} P_\Omega(x)^T d^k(\alpha).$$

Juntando os termos

$$\nabla f(x^k)^T d^k(\alpha) = \frac{1}{\alpha}\big(x^k - P_\Omega(x)\big)^T d^k(\alpha) - \frac{1}{\alpha}\big(x - P_\Omega(x)\big)^T d^k(\alpha).$$

Assim, usando a definição (1.19) e o Teorema 1.6, item (b), temos

$$\nabla f(x^k)^T d^k(\alpha) \leq -\frac{1}{\alpha}\|d^k(\alpha)\|^2. \tag{1.25}$$

Por outro lado, pelo Lema 1.15

$$f(x^k + d^k(\alpha)) - f(x^k) \leq \nabla f(x^k)^T d^k(\alpha) + \frac{L}{2}\|d^k(\alpha)\|^2.$$

Somando e subtraindo o termo $\mu \nabla f(x^k)^T d^k(\alpha)$ no lado direito desta expressão e utilizando (1.25), ficamos com

$$f(x^k + d^k(\alpha)) - f(x^k) \leq \mu \nabla f(x^k)^T d^k(\alpha) + \left(\frac{L}{2} - \frac{1-\mu}{\alpha}\right)\|d^k(\alpha)\|^2. \tag{1.26}$$

Defina

$$\bar{\alpha} = \frac{2(1-\mu)}{L}.$$

Portanto, para $\alpha \in (0, \bar{\alpha}]$ podemos afirmar que

$$\frac{1-\mu}{\alpha} \geq \frac{L}{2}.$$

Substituindo em (1.26), temos

$$f(x^k + d^k(\alpha)) \leq f(x^k) + \mu \nabla f(x^k)^T d^k(\alpha),$$

concluindo a demonstração. □

O teorema que acabamos de demonstrar juntamente com o procedimento de backtracking descrito anteriormente garantem que existe um limitante inferior positivo para os valores de α_k. De fato, de forma semelhante à análise feita em [7, Capítulo 3], se vale (1.21) então

$$\alpha_k = \hat{\alpha}.$$

Caso contrário, o resultado do Teorema 1.16 afirma que $\bar{\alpha} < \hat{\alpha}$, para $\bar{\alpha}$ dado por (1.24) e, além disso, o fator $\theta \in (0,1)$ e o inteiro $s > 0$ da redução (1.22), nos dizem que,

$$\bar{\alpha}\theta < \hat{\alpha}\theta^s = \alpha_k,$$

pois, caso não fosse verdade, teríamos $\hat{\alpha}\theta^{s-1} \leq \bar{\alpha}$. Se $s = 1$, isto contradiz o fato de que $\bar{\alpha} < \hat{\alpha}$ e, caso $s > 1$, temos uma contradição com (1.22) e a afirmação do Teorema 1.16. Sendo assim, vale a limitação inferior

$$0 < \min\{\hat{\alpha}, \bar{\alpha}\theta\} \leq \alpha_k,$$

para toda iteração k.

1.2.1 Algoritmo básico do método do gradiente projetado

Descrevemos agora o algoritmo do método do gradiente projetado proposto em [7], aplicado ao problema (1.1). O critério de parada do algoritmo é motivado pela condição necessária de otimalidade (1.17).

Algoritmo 1 Gradiente projetado

Entrada: $x^0 \in \Omega$, $k = 0$ e as constantes $\mu \in (0,1)$ e $\varepsilon > 0$.

repita

1. Calcule
$$x_{k+1} = P_\Omega(x^k - \alpha_k \nabla f(x^k)),$$
com $\alpha_k > 0$ satisfazendo a condição (1.20).
2. Faça $k = k + 1$.

até $\|P_\Omega\left(x^k - \alpha_k \nabla f(x^k)\right) - x^k\| \leq \varepsilon$;

Para garantir a convergência global do algoritmo acima, é suficiente considerar o requerimento (1.20) (veja [7, Teorema 4.1.1]). Vimos anteriormente, como consequência do Teorema 1.16 e do procedimento de backtracking, que é possível garantir a existência de uma limitação inferior positiva para o passo α_k. No método de região de confiança estudado neste trabalho para o problema (1.1) é necessário que o passo α_k não seja muito pequeno.

Para o esquema descrito no Algoritmo 1 o cálculo da projeção é requerido. Lembre-se que explicitamos este cálculo na expressão (1.7), no caso em que Ω é a caixa (1.2). Note também que não controlamos o tamanho da direção $d^k(\alpha_k)$. Uma maneira de fazer isto é por meio de estratégias de região de confiança.

1.3 Método de região de confiança para problemas irrestritos

Considere o problema irrestrito

$$\begin{aligned} &\text{minimizar} \quad f(x) \\ &\text{sujeito a} \quad x \in \mathbb{R}^n, \end{aligned} \tag{1.27}$$

sendo $f : \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável. Em métodos de região de confiança clássicos para esse problema, aproximamos a função objetivo f por um modelo quadrático e minimizamos o modelo em uma determinada região em torno de um ponto na qual confiamos no modelo. Esta minimização é obtida de forma *aproximada*, no sentido de que buscamos soluções que forneçam um decréscimo suficiente na função objetivo. Caso esse decréscimo seja satisfeito é possível garantir a convergência global do método.

Considere um iterando $x^k \in \mathbb{R}^n$, uma matriz simétrica $B_k \in \mathbb{R}^{n\times n}$ e o modelo quadrático em torno do iterando x^k,

$$q_k(x) = f(x^k) + \nabla f(x^k)^T(x - x^k) + \frac{1}{2}(x - x^k)^T B_k(x - x^k). \tag{1.28}$$

Com base no modelo acima, podemos obter o modelo quadrático em função do deslocamento em uma direção d

$$m_k(d) = q_k(x^k + d) - q_k(x^k) = \nabla f(x^k)^T d + \frac{1}{2} d^T B_k d. \tag{1.29}$$

Dado $\Delta_k > 0$, a cada iteração do método de região de confiança, obtemos uma direção d^k como sendo uma solução (possivelmente aproximada) para o subproblema

$$\begin{aligned} &\text{minimizar} \quad m_k(d) \\ &\text{sujeito a} \quad \|d\| \leq \Delta_k. \end{aligned} \tag{1.30}$$

Então, decidimos se aceitamos a solução d^k, isto é, se atualizamos o ponto corrente como

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

e se aumentamos ou não o valor de Δ_k. Denominamos Δ_k como sendo o raio da região de confiança

$$\{d \in \mathbb{R}^n \mid \|d\| \leq \Delta_k\}.$$

Uma forma de medir se a solução d^k fornece uma redução aceitável para função objetivo é analisar a razão entre o valor da redução real com a redução predita, dada por

$$\rho_k = \frac{f(x^k) - f(x^k + d^k)}{q_k(x^k) - q_k(x^k + d^k)} = \frac{f(x^k) - f(x^k + d^k)}{m_k(0) - m_k(d^k)}. \tag{1.31}$$

Note que, o denominador na expressão acima é sempre positivo quando d^k é não nulo, pois d^k está minimizando o modelo. Portanto, um valor de ρ_k positivo indica que reduzimos o valor da função objetivo. Quanto mais perto este valor estiver de 1, melhor será a redução obtida. Em geral, pede-se que a redução real seja pelo menos uma fração da redução predita.

1.3.1 Estratégia de atualização no método de região de confiança

Descrevemos agora as estratégias de atualização dadas no método de região de confiança, como descritas em [13] e [12]. No método de região de confiança para o problema restrito (1.1), descrito na seção seguinte, utilizamos as mesmas estratégias.

Para a atualização do raio da região de confiança, considere dadas as constantes η_1 e η_2 tais que

$$0 < \eta_1 < \eta_2 < 1, \tag{1.32}$$

e as constantes σ_1, σ_2 e σ_3 tais que

$$0 < \sigma_1 < \sigma_2 < 1 < \sigma_3. \tag{1.33}$$

Então, o raio Δ_{k+1} é atualizado de acordo com:

$$\begin{cases} \text{Caso } \rho_k \leq \eta_1 & \text{então} \quad \Delta_{k+1} \in [\sigma_1 \min\{\|d^k\|, \Delta_k\}, \sigma_2 \Delta_k], \\ \text{Caso } \rho_k \in (\eta_1, \eta_2) & \text{então} \quad \Delta_{k+1} \in [\sigma_2 \Delta_k, \Delta_k], \\ \text{Caso } \rho_k \geq \eta_2 & \text{então} \quad \Delta_{k+1} \in [\Delta_k, \sigma_3 \Delta_k]. \end{cases} \tag{1.34}$$

Na Figura 1.8, ilustramos os intervalos nos quais o raio Δ_{k+1} é menor, maior, ou igual ao raio anterior Δ_k, em função das contantes η_1 e η_2.

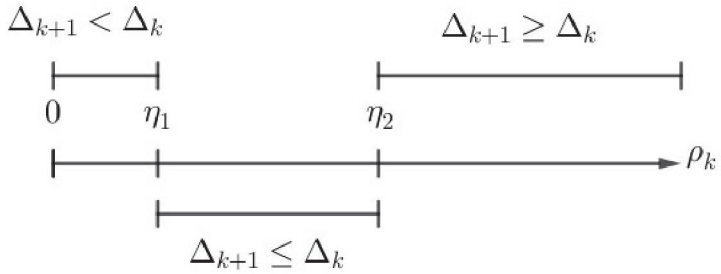

Figura 1.8: Intervalos para ρ_k em (1.34). Fonte: a autora.

Em [11] são sugeridos os valores $\eta_1 = 0.01$, $\eta_2 = 0.9$, $\sigma_1 = \sigma_2 = 0.5$ e $\sigma_3 = \infty$, e em [5], são escolhidos $\eta_1 = 1/4$ e $\eta_2 = 3/4$. Para ambas as referências, as regras de atualização do raio da região de confiança são semelhantes a (1.34).

Para a atualização do iterando x^{k+1} também pedimos que a redução real seja maior do que uma fração da redução predita. Porém, em geral, o critério é menos exigente do que as condições anteriores.

Dado $\eta > 0$ tal que $\eta < \eta_1$:

$$\begin{cases} \text{Caso } \rho_k > \eta & \text{então} \quad x^{k+1} = x^k + d^k, \\ \text{Caso } \rho_k \leq \eta & \text{então} \quad x^{k+1} = x^k. \end{cases} \tag{1.35}$$

Na referência [11] é sugerido $\eta = 0.01$ e em [5], usa-se $\eta \in [0, 1/4)$. Quando $\rho_k > \eta$, vamos considerar que k é uma *iteração de sucesso*.

1.3.2 Algoritmo básico do método de região de confiança irrestrito

O algoritmo do método de região de confiança clássico para o problema (1.27) é descrito nesta seção. O critério de parada neste algoritmo é justificado pela condição necessária de otimalidade para problemas irrestritos $\nabla f(x) = 0$.

Algoritmo 2 Região de confiança

Entrada: $x^0 \in \mathbb{R}^n$, $\Delta_0 > 0$, $k = 0$ e as constantes η_1 e η_2 satisfazendo (1.32), σ_1, σ_2 e σ_3 satisfazendo (1.33), $\eta \in (0, \eta_1)$ e $\varepsilon > 0$.

enquanto $\|\nabla f(x^k)\| > \varepsilon$ **faça**

1. Resolva o subproblema (1.30) (possivelmente de forma aproximada) e obtenha d^k.
2. Calcule ρ_k usando (1.31).
3. Obtenha Δ_{k+1} e x^{k+1} usando (1.34) e (1.35), respectivamente.
4. Faça $k = k + 1$.

fim

Como mencionado anteriormente, desejamos que a direção d^k obtida no Passo 1 do algoritmo acima seja obtida de forma aproximada, fornecendo um decréscimo suficiente para garantir a convergência global do método de região de confiança. Neste trabalho, demonstraremos a obtenção deste decréscimo e a convergência global do algoritmo de região de confiança para o problema restrito (1.1). Para o caso irrestrito, as demonstrações de convergência podem ser encontradas nas referências [5] e [11].

1.4 Método de região de confiança para problemas restritos

Tendo em vista os resultados das seções anteriores, descrevemos o método de região de confiança para resolver o problema restrito

$$\begin{aligned} &\text{minimizar} \quad f(x) \\ &\text{sujeito a} \quad x \in \Omega, \end{aligned} \tag{1.36}$$

em que $f : \mathbb{R}^n \to \mathbb{R}$ é uma função diferenciável e $\Omega \subset \mathbb{R}^n$ um conjunto convexo e fechado. As principais referências para a construção do algoritmo e para a teoria de convergência global são [12], [13] e [15].

Dado um iterando $x^k \in \Omega$, consideramos o mesmo modelo quadrático definido para o caso irrestrito,

$$m_k(d) = \nabla f(x^k)^T d + \frac{1}{2} d^T B_k d, \tag{1.37}$$

em que $B_k \in \mathbb{R}^{n \times n}$ é uma matriz simétrica. Da mesma forma que no método de região de confiança descrito na Seção 1.3, minimizamos o modelo quadrático acima sujeito a uma região de confiança obtendo uma solução aproximada d_k, porém, mantemos a viabilidade

do problema, isto é, dado o raio $\Delta_k > 0$, resolvemos o subproblema

$$\begin{aligned} &\text{minimizar} \quad m_k(d) \\ &\text{sujeito a} \quad \|d\| \leq \Delta_k \\ &\qquad\qquad x^k + d \in \Omega \end{aligned} \quad . \tag{1.38}$$

A atualização do raio da região de confiança e do próximo iterando é feita através das relações (1.34) e (1.35), respectivamente.

Para escrevermos o algoritmo e demonstrarmos o teorema de convergência global do método, precisamos encontrar uma solução aproximada d^k que satisfaça determinadas condições. Essas condições dependem do *passo de Cauchy* para o problema (1.36), que discutiremos na próxima seção.

1.4.1 Passo de Cauchy

Uma das hipóteses para estabelecer a convergência global do método é pedir que a solução aproximada d^k para o subproblema (1.38) forneça um decréscimo no modelo quadrático (1.37). Esse decréscimo é dado em termos da redução obtida pelo *passo de Cauchy*.

O passo de Cauchy é obtido através de uma minimização aproximada do modelo quadrático na direção

$$d^k(\alpha) = P_\Omega\left(x^k - \alpha \nabla f(x^k)\right) - x^k,$$

definida em (1.19) e sujeita a uma região de confiança. Denotamos por α_k uma solução aproximada deste problema de minimização. Neste capítulo, estudamos os requerimentos para esta solução como descritos nas referências [13] e [15].

O valor $\alpha_k > 0$ pode ser obtido de forma iterativa, como sugerido em (1.22), de modo que satisfaça determinadas condições de decréscimo. Considere as constantes

$$0 < \mu_0 < \frac{1}{2}, \quad 0 < \mu_1 \leq \mu_2, \quad \gamma_1 \geq 0 \quad \text{e} \quad \gamma_2 \in (0, 1]. \tag{1.39}$$

Para [15], o valor de α_k deve satisfazer duas condições. A primeira delas é que

$$m_k\left(d^k(\alpha_k)\right) \leq \mu_0 \nabla f(x^k)^T d^k(\alpha_k) \quad \text{e} \quad \|d^k(\alpha_k)\| \leq \mu_2 \Delta_k. \tag{1.40}$$

A segunda condição pede que α_k não seja muito pequeno, isto é,

$$\alpha_k \geq \gamma_1 \tag{1.41}$$

ou

$$\alpha_k \geq \gamma_2 \bar{\alpha}_k, \tag{1.42}$$

em que $\bar{\alpha}_k > 0$ satisfaz a condição

$$m_k\left(d^k(\bar{\alpha}_k)\right) \geq (1 - \mu_0) \nabla f(x^k)^T d^k(\bar{\alpha}_k) \quad \text{ou} \quad \|d^k(\bar{\alpha}_k)\| \geq \mu_1 \Delta_k. \tag{1.43}$$

Podemos interpretar as condições de decréscimo acima para o caso em que $\Omega = \mathbb{R}^n$, de forma semelhante à Figura 1.6. Na Figura 1.9, destacamos um intervalo possível para α_k exemplificando um caso em que $\mu_0 = 1/4$ e $\gamma_2 = 1$.

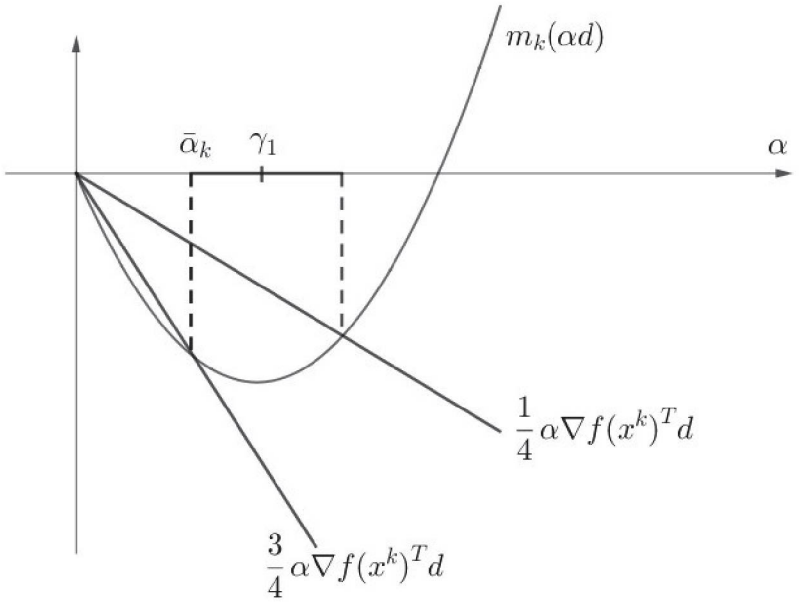

Figura 1.9: Condições de decréscimo (1.40) e (1.43). Fonte: a autora.

Observamos que, em [12], ao invés de (1.43), pede-se que $\bar{\alpha}_k$ satisfaça

$$m_k\left(d^k(\bar{\alpha}_k)\right) > \mu_0 \nabla f(x^k)^T d^k(\bar{\alpha}_k) \quad \text{ou} \quad \|d^k(\bar{\alpha}_k)\| \geq \mu_1 \Delta_k, \tag{1.44}$$

para $\mu_0 \in (0,1)$. Porém, como estamos considerando o modelo quadrático de aproximação, caso $\mu_0 = 1/2$, a condição acima, juntamente com (1.42), poderia rejeitar um minimizador exato deste modelo quadrático. Com base em [12, Teorema 4.2], é possível mostrar que $\alpha_k > 0$ satisfazendo as condições (1.40)-(1.43), é obtido por meio de finitas avaliações do modelo m_k.

Vamos demonstrar agora que o passo de Cauchy fornece um decréscimo suficiente para o modelo quadrático, que será fundamental para o teorema de convergência global do algoritmo de região de confiança. Para isso, precisamos enunciar e demonstrar alguns resultados. O primeiro deles é estabelecido ao longo de [12].

Lema 1.17. *Sejam $x^k \in \Omega$ e $d^k(\alpha)$ definido em (1.19), para todo $k \in \mathbb{N}$. Então, para qualquer $\alpha > 0$, vale que*

$$-\nabla f(x^k)^T d^k(\alpha) \geq \frac{\|d^k(\alpha)\|^2}{\alpha}. \tag{1.45}$$

Demonstração. Como $x^k \in \Omega$, do Teorema 1.6, item (b), vale que

$$\left(P_\Omega\left(x^k - \alpha\nabla f(x^k)\right) - (x^k - \alpha\nabla f(x^k))\right)^T \left(P_\Omega\left(x^k - \alpha\nabla f(x^k)\right) - x^k\right) \leq 0.$$

Então, da definição de $d^k(\alpha)$, ficamos com

$$(d^k(\alpha) + \alpha\nabla f(x^k))^T d^k(\alpha) \leq 0.$$

Distribuindo os termos, segue que

$$\|d^k(\alpha)\|^2 + \alpha\nabla f(x^k)^T d^k(\alpha) \leq 0.$$

Portanto, (1.45) vale para todo $\alpha > 0$. □

Fixado $k \in \mathbb{N}$, para $d \neq 0$, considere a função auxiliar

$$\omega_k(d) = \frac{m_k(d) - \nabla f(x^k)^T d}{\|d\|^2}. \tag{1.46}$$

Para [12], esta função é definida de forma mais genérica, em termos de um modelo de aproximação não necessariamente quadrático.

No próximo lema, demonstramos uma relação entre a função (1.46) e o modelo quadrático m_k. Nesta demonstração, fazemos uma adaptação do resultado demonstrado em [12, Lema 4.3], na qual a condição (1.44) foi considerada.

Lema 1.18. *Sejam $x^k \in \Omega$ e $d^k(\alpha)$ definido em (1.19), para todo $k \in \mathbb{N}$. Suponha que x^k não é estacionário para o problema (1.36) e que para $\alpha > 0$ vale a relação*

$$m_k\left(d^k(\alpha)\right) \geq (1 - \mu_0)\nabla f(x^k)^T d^k(\alpha),$$

para $0 < \mu_0 < 1$. Então $\omega\left(d^k(\alpha)\right)$ é positivo e

$$\alpha \geq \frac{\mu_0}{\omega\left(d^k(\alpha)\right)}. \tag{1.47}$$

Demonstração. Como x^k é não estacionário, então $d^k(\alpha)$ é não nulo para qualquer $\alpha > 0$. Logo, da definição (1.46), temos

$$\omega\left(d^k(\alpha)\right) = \frac{m_k\left(d^k(\alpha)\right)}{\|d^k(\alpha)\|^2} - \frac{\nabla f(x^k)^T d^k(\alpha)}{\|d^k(\alpha)\|^2}.$$

Usando a hipótese do teorema, ficamos com

$$\omega\left(d^k(\alpha)\right) \geq (1 - \mu_0)\frac{\nabla f(x^k)^T d^k(\alpha)}{\|d^k(\alpha)\|^2} - \frac{\nabla f(x^k)^T d^k(\alpha)}{\|d^k(\alpha)\|^2} = -\mu_0 \frac{\nabla f(x^k)^T d^k(\alpha)}{\|d^k(\alpha)\|^2}. \tag{1.48}$$

Além disso, como $\mu_0 > 0$, do Lema 1.17,

$$-\mu_0 \frac{\nabla f(x^k)^T d^k(\alpha)}{\|d^k(\alpha)\|^2} \geq \frac{\mu_0}{\alpha}.$$

Daí, substituindo em (1.48), concluímos que

$$\omega\left(d^k(\alpha)\right) \geq \frac{\mu_0}{\alpha}.$$

Isso mostra que $\omega\left(d^k(\alpha)\right)$ é positivo e vale a desigualdade (1.47). □

Observamos que, sempre que a sequência $(\|B_k\|)_{k\in\mathbb{N}}$ for limitada, teremos que a família das funções $\{\omega_k\}_{k\in\mathbb{N}}$ será uniformemente limitada, isto é, existe uma constante $M > 0$ tal que

$$|\omega_k(d)| \leq M,$$

para qualquer $k \in \mathbb{N}$ e para qualquer d não nulo. Daí, pelo lema anterior, temos

$$\alpha \geq \frac{\mu_0}{M},$$

e o processo de backtracking para determinar o valor do passo α_k termina em um número finito de passos.

Definimos, para μ_2 descrito em (1.39) e $\Delta_k > 0$, a sequência

$$\beta_k = 1 + \sup\{|\omega_k(d)| \mid 0 < \|d\| \leq \mu_2 \Delta_k\} \tag{1.49}$$

e estabelecemos o decréscimo gerado pelo passo de Cauchy. Fazemos uma adaptação do resultado de [12, Teorema 4.1], na qual é considerada a condição (1.44) ao invés de (1.43).

Teorema 1.19. *Suponha que $\alpha_k > 0$ satisfaz as condições (1.40)-(1.43). Então existe $\mu_3 > 0$ tal que*

$$-\nabla f(x^k)^T d^k(\alpha_k) \geq \mu_3 \frac{\|d^k(\alpha_k)\|}{\alpha_k} \min\left\{\Delta_k, \frac{1}{\beta_k}\left(\frac{\|d^k(\alpha_k)\|}{\alpha_k}\right)\right\}. \tag{1.50}$$

Demonstração. Se x^k é estacionário, pela condição de otimalidade (1.16) temos $d^k(\alpha_k) = 0$ e a desigualdade (1.50) é satisfeita trivialmente. Portanto, considere $d^k(\alpha_k)$ não nulo. Do Lema 1.17, temos

$$-\nabla f(x^k)^T d^k(\alpha_k) \geq \frac{\|d^k(\alpha_k)\|^2}{\alpha_k} = \left(\frac{\|d^k(\alpha_k)\|}{\alpha_k}\right)^2 \alpha_k. \tag{1.51}$$

De acordo com as condições (1.40)-(1.43), vamos dividir esta demonstração em três casos.

Caso 1. Suponha que valem as condições (1.40) e (1.41). Como $\beta_k \geq 1$, então, na desigualdade anterior, temos

$$-\nabla f(x^k)^T d^k(\alpha_k) \geq \gamma_1 \left(\frac{\|d^k(\alpha_k)\|}{\alpha_k}\right)^2 \geq \gamma_1 \left(\frac{\|d^k(\alpha_k)\|}{\alpha_k}\right)^2 \frac{1}{\beta_k}. \tag{1.52}$$

Caso 2. Suponha agora que valem as condições (1.40) e (1.42), sendo que para $\bar{\alpha}_k > 0$ vale somente a condição

$$m_k\left(d^k(\bar{\alpha}_k)\right) \geq (1 - \mu_0)\nabla f(x^k)^T d^k(\bar{\alpha}_k).$$

Sendo assim, podemos afirmar que

$$\|d^k(\bar{\alpha}_k)\| < \mu_1 \Delta_k \leq \mu_2 \Delta_k. \tag{1.53}$$

A primeira desigualdade segue do fato de que podemos negar a segunda condição em (1.43), pois esta será analisada no próximo caso. A segunda desigualdade é válida pois $\mu_1 \leq \mu_2$.

Usando o Lema 1.18, temos

$$\omega\left(d^k(\bar{\alpha}_k)\right) > 0 \quad \text{e} \quad \bar{\alpha}_k \geq \frac{\mu_0}{\omega\left(d^k(\bar{\alpha}_k)\right)}.$$

Utilizando as desigualdades (1.42) e (1.51), segue que

$$-\nabla f(x^k)^T d^k(\alpha_k) \geq \left(\frac{\|d^k(\alpha_k)\|}{\alpha_k}\right)^2 (\gamma_2 \bar{\alpha}_k) \geq \left(\frac{\|d^k(\alpha_k)\|}{\alpha_k}\right)^2 \left(\gamma_2 \frac{\mu_0}{\omega\left(d^k(\bar{\alpha}_k)\right)}\right). \tag{1.54}$$

Note que, como vale (1.53), então, da definição de β_k em (1.49)

$$\omega\left(d^k(\bar{\alpha}_k)\right) \leq \beta_k.$$

Logo,

$$\frac{1}{\omega\left(d^k(\bar{\alpha}_k)\right)} \geq \frac{1}{\beta_k}.$$

Substituindo em (1.54), concluímos para este caso a desigualdade

$$-\nabla f(x^k)^T d^k(\alpha_k) \geq (\gamma_2\mu_0)\left(\frac{\|d^k(\alpha_k)\|}{\alpha_k}\right)^2 \frac{1}{\beta_k}. \tag{1.55}$$

Caso 3. Suponha que valem as condições (1.40), (1.42) e agora

$$\|d^k(\bar{\alpha}_k)\| \geq \mu_1\Delta_k. \tag{1.56}$$

Como $\gamma_2 \leq 1$, então $\gamma_2\bar{\alpha}_k \leq \bar{\alpha}_k$. Sendo assim, podemos aplicar o item (e) do Teorema 1.6 e obter

$$\frac{\|d^k(\gamma_2\bar{\alpha}_k)\|}{\gamma_2\bar{\alpha}_k} \geq \frac{\|d^k(\bar{\alpha}_k)\|}{\bar{\alpha}_k}. \tag{1.57}$$

Além disso, da condição (1.42), utilizamos o item (d) do Teorema 1.6 e temos a desigualdade

$$\|d^k(\alpha_k)\| \geq \|d^k(\gamma_2\bar{\alpha}_k)\|. \tag{1.58}$$

Unindo (1.57) e (1.58) e usando a hipótese (1.56), segue que

$$\|d^k(\alpha_k)\| \geq \gamma_2\mu_1\Delta_k.$$

Novamente, do Lema 1.17 e da desigualdade anterior,

$$-\nabla f(x^k)^T d^k(\alpha_k) \geq \frac{\|d^k(\alpha_k)\|^2}{\alpha_k} \geq (\gamma_2\mu_1)\Delta_k\frac{\|d^k(\alpha_k)\|}{\alpha_k}. \tag{1.59}$$

Considere $\mu_3 > 0$ tal que $\mu_3 \leq \min\{\gamma_1, \gamma_2\mu_0, \gamma_2\mu_1\}$. Então, das conclusões (1.52), (1.55) e (1.59) dos casos anteriores, concluímos a demonstração. □

Note que, da condição (1.40) e de $m_k(0) = 0$, podemos escrever

$$m_k(0) - m_k\left(d^k(\alpha_k)\right) \geq -\mu_0\nabla f(x^k)^T d^k(\alpha_k).$$

Então α_k satisfaz

$$m_k(0) - m_k\left(d^k(\alpha_k)\right) \geq \mu_0\mu_3\frac{\|d^k(\alpha_k)\|}{\alpha_k}\min\left\{\Delta_k, \frac{1}{\beta_k}\left(\frac{\|d^k(\alpha_k)\|}{\alpha_k}\right)\right\}.$$

1.4.2 Algoritmo básico do método de região de confiança restrito

O algoritmo do método de região de confiança para o problema (1.36) pede que a solução aproximada d^k do problema de região de confiança (1.38) forneça uma redução no modelo quadrático que depende do passo de Cauchy, $d^k(\alpha_k)$, descrito anteriormente. Para [15], o passo d^k deve satisfazer as condições

$$m_k(d^k) \leq \mu_0 m_k\left(d^k(\alpha_k)\right), \quad \|d^k\| \leq \mu_2 \Delta_k \quad \text{e} \quad x^k + d^k \in \Omega. \tag{1.60}$$

Descrevemos agora o algoritmo do método de região de confiança estudado.

Algoritmo 3 Região de confiança restrito

Entrada: $x^0 \in \Omega$, $\Delta_0 > 0$, $k = 0$ e as constantes η_1 e η_2 satisfazendo (1.32), σ_1, σ_2 e σ_3 satisfazendo (1.33), μ_0, μ_1, μ_2, γ_1 e γ_2 satisfazendo (1.39), $\eta \in (0, \eta_1)$ e $\varepsilon > 0$.

repita

1. Construa o modelo quadrático (1.37).
2. Encontre o passo $\alpha_k > 0$ satisfazendo as condições (1.40)-(1.43).
3. Encontre d^k satisfazendo a condição (1.60).
4. Calcule ρ_k usando (1.31).
5. Obtenha Δ_{k+1} e x^{k+1} usando (1.34) e (1.35), respectivamente.
6. Faça $k = k + 1$.

até $\|x^k - P_\Omega\left(x^k - \nabla f(x^k)\right)\| \leq \varepsilon$;

Assim como no método de gradiente projetado, o critério de parada do algoritmo anterior é justificado pela condição necessária de otimalidade demonstrada no Teorema 1.13, com $\alpha = 1$. Para [13], o passo d^k também deve satisfazer as condições em (1.60), porém, além das condições (1.40)-(1.43), o valor de α_k está limitado superiormente, isto é, pede-se que

$$\alpha_k \leq \gamma_3, \tag{1.61}$$

para alguma constante $\gamma_3 > 0$ fixada. No Capítulo 3 deste trabalho, descrevemos os detalhes computacionais da abordagem utilizada em [13] para obter o valores de α_k e de d^k requeridos nos Passos 2 e 3, respectivamente.

1.4.3 Convergência global

Nesta seção vamos estudar a teoria de convergência global do Algoritmo 3, aplicado ao problema restrito (1.36). Vamos mostrar que, se a sequência (x^k) gerada pelo algoritmo possui um ponto limite então este ponto é estacionário. Primeiramente, definimos a sequência $(\hat{\alpha}_k)$ sendo

$$\hat{\alpha}_k = \min\{\alpha_k, \gamma_3\}, \tag{1.62}$$

para uma constante fixada $\gamma_3 > 0$ e $\alpha_k > 0$ gerado pelo algoritmo. Esta sequência é definida para os casos em que o valor de α_k torna-se muito grande. Note que, ao contrário da sequência α_k do Algoritmo 3, esta sequência é limitada superiormente, pois,

$$\hat{\alpha}_k \leq \gamma_3, \tag{1.63}$$

para todo $k \in \mathbb{N}$. Esta informação será útil para demonstrarmos os resultados de convergência global, estabelecidos em [15]. Primeiramente, vamos mostrar que a relação de decréscimo satisfeita pelo passo de Cauchy, dada pelo Teorema 1.19, pode ser expressa em termos de $\hat{\alpha}_k$. Para isto, demonstramos um resultado auxiliar na forma de um lema.

Lema 1.20. *Considere o caminho $d^k : \mathbb{R} \to \mathbb{R}^n$ definido em (1.19). Se $\alpha \geq \beta > 0$ então*

$$-\nabla f(x^k)^T d^k(\alpha) \geq -\nabla f(x^k)^T d^k(\beta).$$

Demonstração. Denote $x = x^k$ e $g = \nabla f(x^k)$. Da propriedade da projeção, descrita no Teorema 1.6, item (b), temos

$$\left(P_\Omega(x - \alpha g) - x + \alpha g\right)^T \left(P_\Omega(x - \alpha g) - z\right) \leq 0, \tag{1.64}$$

para qualquer $z \in \Omega$. Escolhendo $z = P_\Omega(x - \beta g)$ e substituindo na expressão (1.64), ficamos com

$$\left(P_\Omega(x - \alpha g) - x + \alpha g\right)^T \left(P_\Omega(x - \alpha g) - P_\Omega(x - \beta g)\right) \leq 0,$$

que é o mesmo que

$$\left(P_\Omega(x - \alpha g) - x + \alpha g\right)^T \left(P_\Omega(x - \alpha g) - P_\Omega(x - \beta g) + x - x\right) \leq 0.$$

Portanto, da definição de d^k,

$$\left(d^k(\alpha) + \alpha g\right)^T \left(d^k(\alpha) - d^k(\beta)\right) \leq 0.$$

Distribuindo os termos desta desigualdade, segue que

$$\alpha g^T \left(d^k(\alpha) - d^k(\beta)\right) \leq d^k(\alpha)^T d^k(\beta) - \|d^k(\alpha)\|^2. \tag{1.65}$$

Por outro lado, da desigualdade de Cauchy-Schwarz e do Teorema 1.6, item (d),

$$d^k(\alpha)^T d^k(\beta) \leq \|d^k(\alpha)\|^2.$$

Substituindo esta expressão em (1.65), temos

$$\alpha g^T \left(d^k(\alpha) - d^k(\beta)\right) \leq 0.$$

Como $\alpha > 0$, concluímos que

$$g^T d^k(\alpha) \leq g^T d^k(\beta),$$

como queríamos demonstrar. □

Podemos expressar o decréscimo no modelo quadrático fornecido pelo passo de Cauchy, $d^k(\alpha_k)$, em termos da sequência $(\hat{\alpha}_k)$, definida anteriormente. Este resultado também é justificado ao longo de [12] com a utilização da condição (1.44) em vez de (1.43).

Lema 1.21. *Suponha que $\alpha_k > 0$ satisfaz as condições (1.40)-(1.43). Considere $\hat{\alpha}_k$ definido em (1.62). Então existe uma constante $c > 0$ tal que*

$$-\nabla f(x^k)^T d^k(\alpha_k) \geq c \frac{\|d^k(\hat{\alpha}_k)\|}{\hat{\alpha}_k} \min\left\{\Delta_k, \frac{1}{\beta_k}\left(\frac{\|d^k(\hat{\alpha}_k)\|}{\hat{\alpha}_k}\right)\right\}. \tag{1.66}$$

Demonstração. Vamos analisar a sequência $(\hat{\alpha}_k)_{k\in\mathbb{N}}$, dada por (1.62). Caso $\alpha_k < \gamma_3$ então $\hat{\alpha}_k = \alpha_k$. Logo, do Teorema 1.19, existe $\mu_3 > 0$ tal que

$$-\nabla f(x^k)^T d^k(\alpha_k) \geq \mu_3 \left(\frac{\|d^k(\hat{\alpha}_k)\|}{\hat{\alpha}_k}\right)^2 \frac{1}{\beta_k} \tag{1.67}$$

ou

$$-\nabla f(x^k)^T d^k(\alpha_k) \geq \mu_3 \Delta_k \frac{\|d^k(\hat{\alpha}_k)\|}{\hat{\alpha}_k}. \tag{1.68}$$

Caso $\alpha_k \geq \gamma_3$ então $\hat{\alpha}_k = \gamma_3$. Daí

$$-\nabla f(x^k)^T d^k(\alpha_k) \geq -\nabla f(x^k)^T d^k(\hat{\alpha}_k) \geq \left(\frac{\|d^k(\hat{\alpha}_k)\|}{\hat{\alpha}_k}\right)^2 \frac{\gamma_3}{\beta_k}. \tag{1.69}$$

A primeira desigualdade segue do Lema 1.20 e a segunda desigualdade é justificada pelo Lema 1.17. Seja $c > 0$ tal que $c \leq \min\{\mu_3, \gamma_3\}$. Portanto, das relações (1.67), (1.68) e (1.69), concluímos a demonstração. □

Para demonstrar a convergência global do Algoritmo 3 consideramos a sequência (x_c^k) dada por

$$x_c^k = x^k + d^k(\hat{\alpha}_k), \tag{1.70}$$

em que x^k é gerado pelo algoritmo e $\hat{\alpha}_k$ é definido em (1.62). Recorde que, da definição (1.19),

$$d^k(\hat{\alpha}_k) = P_\Omega\left(x^k - \hat{\alpha}_k \nabla f(x^k)\right) - x^k.$$

Logo $x_c^k \in \Omega$, para todo $k \in \mathbb{N}$. Além disso, vamos considerar a medida de estacionariedade para $x \in \Omega$ em termos do vetor chamado de *gradiente projetado*, definido nas referências [12], [13] e [15]. Este vetor é a projeção no cone tangente a partir de x, do vetor oposto ao gradiente, isto é,

$$\nabla_\Omega f(x) \equiv P_{T_\Omega(x)}\left(-\nabla f(x)\right).$$

Note que este vetor é a projeção no cone tangente da direção de máxima descida. Nesta seção vamos estudar as principais propriedades da função

$$x \mapsto \|\nabla_\Omega f(x)\|, \tag{1.71}$$

para $x \in \Omega$. Veremos como obter a convergência global do Algoritmo 3 utilizando esta medida e a sequência (x_c^k).

Quando o conjunto Ω é a caixa definida em (1.2), temos a expressão para o gradiente projetado

$$\left(\nabla_\Omega f(x)\right)_i = \begin{cases} -\partial_i f(x) & \text{se} \quad x_i \in (\ell_i, u_i), \\ \max\{-\partial_i f(x), 0\} & \text{se} \quad x_i = \ell_i, \\ \min\{-\partial_i f(x), 0\} & \text{se} \quad x_i = u_i, \end{cases} \tag{1.72}$$

para cada $i = 1, \ldots, n$. A Figura 1.10 ilustra o caso $n = 2$ para uma função cuja matriz Hessiana é definida positiva.

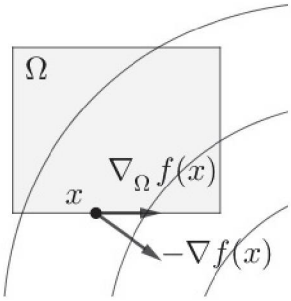

Figura 1.10: Cálculo do vetor (1.72). Fonte: a autora.

Temos que $\nabla_\Omega f(x) = 0$ se, e somente se, x satisfaz uma condição necessária de primeira ordem para o problema (1.36). Para justificarmos essa afirmação, primeiramente, demonstramos duas propriedades do vetor $\nabla_\Omega f(x)$. Em seguida, a equivalência mencionada. A referência para estes resultados é [16, Lema 3.1].

Lema 1.22. *Seja Ω um conjunto convexo e fechado. Para qualquer $x \in \Omega$ valem as seguintes relações*

(a) $\|\nabla_\Omega f(x)\|^2 = -\nabla f(x)^T \nabla_\Omega f(x)$.

(b) $\min\{\nabla f(x)^T v \mid v \in T_\Omega(x), \|v\| \leq 1\} = -\|\nabla_\Omega f(x)\|$.

Demonstração. Como $T_\Omega(x)$ é um cone e, por definição, $\nabla_\Omega f(x) \in T_\Omega(x)$, então, para qualquer $\lambda \geq 0$, vale que $\lambda \nabla_\Omega f(x) \in T_\Omega(x)$. Do Teorema 1.6, item (a),

$$(\nabla_\Omega f(x) - (-\nabla f(x)))^T (\lambda \nabla_\Omega f(x) - \nabla_\Omega f(x)) \leq 0.$$

Fazendo $\lambda = 0$ e $\lambda = 2$, concluímos que

$$(\nabla_\Omega f(x) + \nabla f(x))^T \nabla_\Omega f(x) = 0.$$

Portanto, distribuindo os termos nesta última igualdade, provamos o item (a). Da definição de projeção, temos

$$\|\nabla_\Omega f(x) + \nabla f(x)\|^2 \leq \|u + \nabla f(x)\|^2,$$

para qualquer $u \in T_\Omega(x)$. Caso $\|u\| \leq \|\nabla_\Omega f(x)\|$, desenvolvendo o lado esquerdo e o lado direito da desigualdade acima, ficamos com

$$\nabla f(x)^T \nabla_\Omega f(x) \leq \nabla f(x)^T u.$$

Assim, usando o item (a)

$$-\|\nabla_\Omega f(x)\|^2 \leq \nabla f(x)^T u. \tag{1.73}$$

Seja $v \in T_\Omega(x)$ tal que $\|v\| \leq 1$. Então

$$\|\nabla_\Omega f(x)\| \|v\| \leq \|\nabla_\Omega f(x)\|.$$

Defina $u = \lambda v$, sendo $\lambda = \|\nabla_\Omega f(x)\|$. Substituindo em (1.73), obtemos

$$-\|\nabla_\Omega f(x)\| \leq \nabla f(x)^T v. \tag{1.74}$$

Note que, a igualdade em (1.74) é válida quando

$$v = \frac{\nabla_\Omega f(x)}{\|\nabla_\Omega f(x)\|},$$

o que prova o item (b). □

Teorema 1.23. *Considere o problema restrito (1.36) e $\bar{x} \in \Omega$. Então $\nabla_\Omega f(\bar{x}) = 0$ se, e somente se, $\bar{x}$ é estacionário para este problema.*

Demonstração. Suponha que $\bar{x}$ é estacionário, isto é, satisfaz a condição necessária de primeira ordem

$$\nabla f(\bar{x})^T (x - \bar{x}) \geq 0, \tag{1.75}$$

para todo $x \in \Omega$. Seja v uma direção viável a partir de $\bar{x}$, ou seja, $v \in V_\Omega(\bar{x})$. Então, de acordo com a Definição 1.1, é imediato que $\nabla f(\bar{x})^T v \geq 0$. Como vale (1.5), por continuidade, segue que

$$\nabla f(\bar{x})^T u \geq 0, \ \forall \ u \in T_\Omega(\bar{x}).$$

Sendo assim, comparando esta desigualdade com o lema anterior, item (b), concluímos que

$$-\|\nabla_\Omega f(\bar{x})\| \geq 0,$$

ou seja $\|\nabla_\Omega f(\bar{x})\| = 0$. Reciprocamente, se $\nabla_\Omega f(\bar{x}) = 0$ então, novamente, do lema anterior, item (b), vale que

$$\nabla f(\bar{x})^T v \geq 0, \ \forall \ v \in T_\Omega(\bar{x}) \text{ com } \|v\| \leq 1.$$

Logo, é fácil ver que, esta relação vale caso $\|v\| > 1$. Considere $x \in \Omega$ qualquer. Do Lema 1.11 e da relação (1.4), temos $x - \bar{x} \in T_\Omega(\bar{x})$. Portanto, (1.75) é satisfeito, isto é, $\bar{x}$ é estacionário. □

O objetivo desta seção é mostrar que, se x^* é limite de uma subsequência de (x^k), gerada pelo Algoritmo 3, digamos (x^{k_i}), então a subsequência $(x_c^{k_i})$ de (1.70) também converge para x^* e vale que

$$\lim_{i \to \infty} \|\nabla_\Omega f(x_c^{k_i})\| = 0. \tag{1.76}$$

A medida de estacionariedade (1.71) pode ser descontínua. Isto significa que, para x próximo de um ponto estacionário, o valor de $\|\nabla_\Omega f(x)\|$ pode se tornar grande. A Figura 1.11 exemplifica este comportamento para três pontos de um conjunto viável $\mathcal{C}$ em que considera-se uma função objetivo linear. Os vetores representados pelas linhas grossas ilustram o vetor $\nabla_\Omega f(x)$ para algum ponto x.

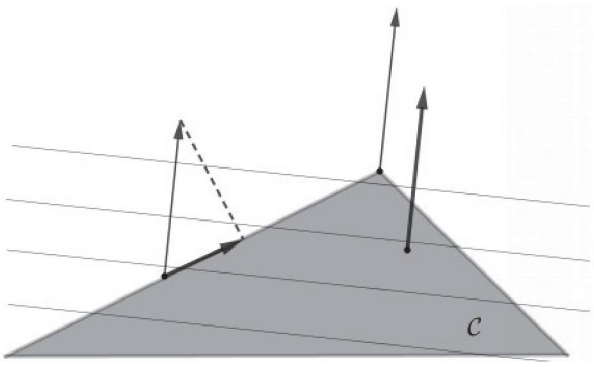

Figura 1.11: Descontinuidade de (1.71). Fonte: [11, Figura 12.1.6].

No entanto, demonstramos no lema seguinte uma propriedade desta medida que será suficiente para garantirmos a convergência (1.76) nos próximos teoremas.

Lema 1.24. [16, Lema 3.3] *Seja $(x^k) \subset \Omega$ uma sequência convergente para $x \in \Omega$. Se $f : \mathbb{R}^n \to \mathbb{R}$ é continuamente diferenciável no conjunto Ω então*

$$\|\nabla_\Omega f(x)\| \leq \liminf_k \|\nabla_\Omega f(x^k)\|.$$

Demonstração. Do Lema 1.11 e da inclusão (1.4) temos, para qualquer $z \in \Omega$ tal que $z \neq x^k$, que

$$\frac{z - x^k}{\|z - x^k\|} \in T_\Omega(x^k).$$

Então, do Lema 1.22, item (b), vale que

$$\nabla f(x^k)^T(x^k - z) \leq \|\nabla_\Omega f(x^k)\|\|z - x^k\|.$$

Como x^k converge para x e f é continuamente diferenciável,

$$\nabla f(x)^T(x - z) \leq \liminf_k \|\nabla_\Omega f(x^k)\|\|z - x\|.$$

Seja $v \in V_\Omega(x)$ tal que $\|v\| \leq 1$. Assim, por definição, existe $\varepsilon > 0$ tal que $z = x + \varepsilon v \in \Omega$. Substituindo z na desigualdade acima ficamos com

$$-\nabla f(x)^T v \leq \liminf_k \|\nabla_\Omega f(x^k)\|.$$

Como $T_\Omega(x) = \overline{V_\Omega(x)}$, a desigualdade acima é válida se considerarmos $v \in T_\Omega(x)$ tal que $\|v\| \leq 1$. Novamente, do Lema 1.22, item (b), podemos afirmar que

$$\|\nabla_\Omega f(x)\| \leq \liminf_k \|\nabla_\Omega f(x^k)\|,$$

concluindo a demonstração. □

Como mencionamos anteriormente, objetivo desta seção é mostrar a convergência (1.76). Primeiramente, recorde que

$$d^k(\hat{\alpha}_k) = x_c^k - x^k.$$

O lema a seguir e o próximo teorema estabelecem comportamentos desta sequência que serão utilizados na teoria de convergência global. Estes resultados foram demonstrados em [12, Teorema 4.4] com a condição (1.44) em vez de (1.43) e a hipótese de que a função f é limitada por baixo em vez de somente a sequência $(f(x^k))$ limitada inferiormente.

Lema 1.25. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ continuamente diferenciável e (x^k) uma sequência gerada pelo Algoritmo 3. Suponha que $(f(x^k))$ é limitada inferiormente e que a sequência (β_k), definida em (1.49), é limitada. Se existir $\varepsilon_0 > 0$ tal que*

$$\frac{\|d^k(\hat{\alpha}_k)\|}{\hat{\alpha}_k} \geq \varepsilon_0, \tag{1.77}$$

para todo k suficientemente grande, então a série

$$\sum_{k=1}^{\infty} \Delta_k \tag{1.78}$$

é convergente, sendo Δ_k o raio da região de confiança.

Demonstração. Vamos analisar o conjunto

$$J = \{k \in \mathbb{N} \mid \rho_k > \eta_1\},$$

em que η_1 é definido no Algoritmo 3. Se J é finito então seu complementar é infinito. Logo, da regra de atualização (1.34) para o raio da região de confiança, temos

$$\Delta_{k+1} \leq \sigma_2 \Delta_k,$$

para k suficientemente grande. Como $\sigma_2 < 1$, a série cujo termo geral é σ_2^k é convergente. Usando o critério de d'Alembert para convergência de séries, descrito em [21], concluímos que a série (1.78) converge. Caso J seja infinito, considere uma subsequência indexada por (k_i) tal que $\rho_{k_i} > \eta_1$. Isso implica que $\rho_{k_i} > \eta$. Dessa forma a direção d^{k_i} obtida no algoritmo é aceita. Então, como $\eta_1 > \eta$

$$x^{k_i+1} = x^{k_i} + d^{k_i},$$

para todo $i \in \mathbb{N}$. Da definição de ρ_{k_i} em (1.31) e usando as condições (1.60) e (1.40), segue que

$$f(x^{k_i}) - f(x^{k_i+1}) > \eta_1 \left(-\mu_0^2 \nabla f(x^{k_i})^T d^{k_i}(\alpha_{k_i})\right).$$

Além disso, recorde que $d^{k_i}(\alpha_{k_i})$ é o passo de Cauchy. Então, do Lema 1.21 e usando a hipótese (1.77), existe $c > 0$ tal que

$$f(x^{k_i}) - f(x^{k_i+1}) > c\eta_1 \mu_0^2 \varepsilon_0 \min\left\{\Delta_{k_i}, \frac{\varepsilon_0}{\beta_{max}}\right\},$$

em que $\beta_{max} > 0$ é tal que $\beta_k \leq \beta_{max}$, pois estamos supondo a sequência (β_k) limitada. Note que, o mínimo no lado direito da desigualdade acima não pode ser $\varepsilon_0/\beta_{max}$ para infinitos valores de $k_i \in J$, pois $(f(x^k))$ é limitada inferiormente e não crescente. Dessa forma, podemos afirmar que existe uma subsequência indexada por (k_i) tal que $\rho_{k_i} > \eta_1$ e

$$f(x^{k_i}) - f(x^{k_i+1}) > c\eta_1 \mu_0^2 \varepsilon_0 \Delta_{k_i}.$$

Temos $f(x^{k_{i+1}}) \leq f(x^{k_i+1})$, pois $k_i + 1 \leq k_{i+1}$. Consequentemente,

$$\Delta_{k_i} < \frac{f(x^{k_i}) - f(x^{k_{i+1}})}{c\eta_1 \mu_0^2 \varepsilon_0}. \tag{1.79}$$

Note que, as hipóteses sobre f garantem que a série cujo termo geral é $(f(x^{k_i}) - f(x^{k_{i+1}}))$ é convergente. Portanto, a desigualdade (1.79) implica que a série

$$\sum_{i=1}^{\infty} \Delta_{k_i}$$

é convergente. Como $k_i \in J$, da regra de atualização (1.34), temos,

$$\Delta_{k_i+1} \leq \sigma_3 \Delta_{k_i}. \tag{1.80}$$

Sabemos que, entre k_i e k_{i+1} existem no máximo um número finito de índices, digamos j_i índices. Sendo assim, da definição do conjunto J, $\rho_{k_i+\ell} \leq \eta_1$, para todo $\ell = 1, \ldots, j_i$. Novamente, da regra de atualização (1.34)

$$\Delta_{(k_i+\ell)+1} \leq \sigma_2 \Delta_{k_i+\ell}, \tag{1.81}$$

para todo $\ell = 1, \ldots, j_i$. Daí, as desigualdades em (1.81), juntamente com (1.80), implicam que

$$\begin{cases} \Delta_{k_i+2} & \leq & \sigma_2 \Delta_{k_i+1} & \leq & \sigma_2 \left(\sigma_3 \Delta_{k_i}\right), \\ \Delta_{k_i+3} & \leq & \sigma_2 \Delta_{k_i+2} & \leq & \sigma_2^2 \left(\sigma_3 \Delta_{k_i}\right), \\ & & \vdots & & \\ \Delta_{k_i+j_i+1} & \leq & \sigma_2 \Delta_{k_i+j_i} & \leq & \sigma_2^{j_i} \left(\sigma_3 \Delta_{k_i}\right). \end{cases} \tag{1.82}$$

Somando as desigualdades em (1.82) e a desigualdade (1.80), temos

$$\sum_{k=k_i+1}^{k_i+j_i+1} \Delta_k \leq \left(1 + \sigma_2 + \sigma_2^2 + \ldots + \sigma_2^{j_i}\right) \sigma_3 \Delta_{k_i},$$

para todo $i \in \mathbb{N}$. Além disso, da expressão para a soma finita do lado direito da desigualdade acima e, notando que $k_i + j_i + 1 = k_{i+1}$, para todo $i \in \mathbb{N}$, segue que

$$\sum_{k=k_i+1}^{k_{i+1}} \Delta_k \leq \left(\frac{1 - \sigma_2^{j_i+1}}{1 - \sigma_2}\right) \sigma_3 \Delta_{k_i}.$$

Como $\sigma_2 > 0$, então $1 - \sigma_2^{j_i+1} < 1$. Portanto,

$$\sum_{k=k_1+1}^{\infty} \Delta_k \leq \left(\frac{\sigma_3}{1 - \sigma_2}\right) \sum_{i=1}^{\infty} \Delta_{k_i}.$$

Provamos anteriormente que a série do lado direito da desigualdade acima é convergente. Então a série (1.78) é convergente, como queríamos demonstrar. □

Teorema 1.26. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ continuamente diferenciável e (x^k) uma sequência gerada pelo Algoritmo 3. Suponha que $(f(x^k))$ é limitada inferiormente e que a sequência (β_k), definida em (1.49), é limitada. Então*

$$\liminf_k \frac{\|d^k(\hat{\alpha}_k)\|}{\hat{\alpha}_k} = 0. \tag{1.83}$$

Demonstração. Suponha por absurdo que (1.83) é falso. Então, podemos afirmar que existe $\varepsilon_0 > 0$ tal que

$$\frac{\|d^k(\hat{\alpha}_k)\|}{\hat{\alpha}_k} \geq \varepsilon_0, \tag{1.84}$$

para todo k suficientemente grande. Pelo lema anterior, a série

$$\sum_{k=1}^{\infty} \Delta_k \tag{1.85}$$

é convergente. Afirmamos que

$$|\rho_k - 1| \to 0. \tag{1.86}$$

De fato, sendo f continuamente diferenciável, da fórmula de Taylor de primeira ordem (1.14) temos, para todo ponto x^k gerado pelo algoritmo,

$$f(x^k + d) = f(x^k) + \nabla f(x^k)^T d + R(d)\|d\|, \tag{1.87}$$

sendo $R(d) \to 0$ sempre que $\|d\| \to 0$, pois vale (1.15). Note que, a condição (1.60) e a convergência da série (1.85) implicam que $\|d^k\| \to 0$. Logo, fazendo $d = d^k$ em (1.87), concluímos que existe uma sequência (ξ_k) tal que $\xi_k \to 0$ e

$$|f(x^k + d^k) - f(x^k) - \nabla f(x^k)^T d^k| = \xi_k \|d^k\|. \tag{1.88}$$

É fácil ver que a definição de β_k, dada em (1.49), e (1.60), implicam que

$$\beta_k \geq \frac{|m_k(d^k) - \nabla f(x^k)^T d^k|}{\|d^k\|^2}.$$

Daí,

$$|m_k(d^k) - \nabla f(x^k)^T d^k| \leq \beta_k \|d^k\|^2.$$

Esta desigualdade, (1.88) e a desigualdade triangular, implicam que

$$|f(x^k + d^k) - f(x^k) - m_k(d^k)| \leq \beta_k \|d^k\|^2 + \xi_k \|d^k\|. \tag{1.89}$$

Por outro lado, usando a definição de ρ_k

$$|\rho_k - 1| = \frac{|f(x^k + d^k) - f(x^k) - m_k(d^k)|}{|m_k(d^k)|}.$$

Usando (1.89) e o fato de que d^k minimiza o modelo m_k, temos

$$|\rho_k - 1| \leq \frac{\beta_k \|d^k\|^2 + \xi_k \|d^k\|}{-m_k(d^k)}. \tag{1.90}$$

Além disso, da condição (1.40) para α_k

$$-m_k(d^k) \geq -\mu_0 m_k(d^k(\alpha_k)).$$

Usando o Lema 1.21, condição (1.60) e a hipótese (1.84), podemos afirmar que existe $c > 0$ tal que

$$-m_k(d^k) \geq c\mu_2^2 \varepsilon_0 \min\left\{\Delta_k, \frac{\varepsilon_0}{\beta_{max}}\right\},$$

em que $\beta_{max} > 0$ é tal que $\beta_k \leq \beta_{max}$, pois, por hipótese, a sequência (β_k) é limitada. Porém, vale que $\Delta_k \to 0$, pois a série (1.85) é convergente. Então

$$-m_k(d^k) \geq c\mu_2^2 \varepsilon_0 \Delta_k,$$

para todo k suficientemente grande. Substituindo em (1.90), podemos concluir que existe uma constante $c_1 > 0$ tal que

$$|\rho_k - 1| \leq c_1 \xi_k \frac{\|d^k\|}{\Delta_k} + c_1 \beta_{max} \frac{\|d^k\|^2}{\Delta_k},$$

para todo k suficientemente grande. Novamente, usando a condição (1.60), temos

$$|\rho_k - 1| \leq c_1\mu_2\xi_k + c_1\mu_2\beta_{max}\Delta_k.$$

Da convergência das sequências (Δ_k) e (ξ_k), a afirmação (1.86) está provada. Logo, $\rho_k \to 1$. Recorde que $\eta_2 < 1$. Então $\rho_k \geq \eta_2$ para k suficientemente grande. Isso implica que, para este índices,

$$\Delta_{k+1} \geq \Delta_k,$$

ou seja, a sequência (Δ_k) não pode convergir para zero. Isso contradiz a convergência da série (1.85), concluíndo a demonstração. □

O próximo lema estabelece uma relação entre $d^k(\hat{\alpha}_k)$ e a medida de estacionariedade $\nabla_\Omega f(x)$, em termos da sequência x_c^k.

Lema 1.27. [15, Lema 5.1] *Seja $f : \mathbb{R}^n \to \mathbb{R}$ diferenciável no conjunto Ω. Considere a sequência (x_c^k) definida em (1.70). Então*

$$\|\nabla_\Omega f(x_c^k)\| \leq \|\nabla f(x_c^k) - \nabla f(x^k)\| + \frac{\|x_c^k - x^k\|}{\hat{\alpha}_k}.$$

Demonstração. Como $x_c^k = P_\Omega\left(x^k - \hat{\alpha}_k\nabla f(x^k)\right)$, da propriedade da projeção, Teorema 1.6, item (b), temos

$$\left(x_c^k - x^k + \hat{\alpha}_k\nabla f(x^k)\right)^T\left(x_c^k - z^k\right) \leq 0,$$

para todo $z^k \in \Omega$. Distribuindo os termos do produto interno acima e usando a desigualdade de Cauchy-Schwarz

$$\hat{\alpha}_k\nabla f(x^k)^T(x_c^k - z^k) \leq \|x_c^k - x^k\|\|x_c^k - z^k\|.$$

Considere v^k uma direção viável qualquer a partir de x_c^k tal que $\|v^k\| \leq 1$. Então existe $\tau_k > 0$ tal que $z^k = x_c^k + \tau_k v^k \in \Omega$. Substituindo z^k na desigualdade anterior, ficamos com

$$-\nabla f(x^k)^T v^k \leq \frac{\|x_c^k - x^k\|}{\hat{\alpha}_k},$$

que é o mesmo que

$$\nabla f(x_c^k)^T v^k - \nabla f(x_c^k)^T v^k - \nabla f(x^k)^T v^k \leq \frac{\|x_c^k - x^k\|}{\hat{\alpha}_k}.$$

Rearranjando os termos nessa desigualdade e usando novamente a desigualdade de Cauchy-Schwarz, segue que

$$-\|\nabla f(x_c^k) - \nabla f(x^k)\| - \frac{\|x_c^k - x^k\|}{\hat{\alpha}_k} \leq \nabla f(x_c^k)^T v^k. \tag{1.91}$$

Sabemos que $T_\Omega(x_c^k) = \overline{V_\Omega(x_c^k)}$. Então (1.91) implica que

$$-\|\nabla f(x_c^k) - \nabla f(x^k)\| - \frac{\|x_c^k - x^k\|}{\hat{\alpha}_k} \leq \nabla f(x_c^k)^T v,$$

para qualquer $v \in T_\Omega(x_c^k)$ tal que $\|v\| \leq 1$. Portanto, do Lema 1.22, item (b), concluímos a demonstração. □

Finalmente, podemos demonstrar a convergência (1.76).

Teorema 1.28. [15, Teorema 5.4] *Suponha as mesmas hipóteses do Teorema 1.26 e $x^* \in \Omega$ limite de uma subsequência de (x^k). Então existe uma subsequência (x^{k_i}) que converge para x^* satisfazendo*

$$\lim_{i\to\infty} \|\nabla_\Omega f(x_c^{k_i})\| = 0,$$

em que (x_c^k) é a sequência definida por (1.70). Além disso, $x_c^{k_i} \to x^*$ e, portanto, x^* é estacionário para o problema (1.36).

Demonstração. Considere (x^{l_i}) uma subsequência de (x^k) convergindo para x^*. O caso mais simples para a demonstração deste teorema é supor, primeiramente, que

$$\frac{\|d^{l_i}(\hat{\alpha}_{l_i})\|}{\hat{\alpha}_{l_i}} \to 0. \tag{1.92}$$

Recorde que a sequência $(\hat{\alpha}_k)$ é limitada superiormente por uma constante não nula e que $\hat{\alpha}_k > 0$. Então, (1.92) implica que $\|d^{l_i}(\hat{\alpha}_{l_i})\| \to 0$. Consequentemente, segue da definição de $x_c^{l_i}$ que

$$x_c^{l_i} \to x^*.$$

Pelo Lema 1.27 vale a desigualdade

$$\|\nabla_\Omega f(x_c^{l_i})\| \leq \|\nabla f(x_c^{l_i}) - \nabla f(x^{l_i})\| + \frac{\|x_c^{l_i} - x^{l_i}\|}{\hat{\alpha}_{l_i}} = \|\nabla f(x_c^{l_i}) - \nabla f(x^{l_i})\| + \frac{\|d^{l_i}(\hat{\alpha}_{l_i})\|}{\hat{\alpha}_{l_i}},$$

e, sendo f continuamente diferenciável, por (1.92)

$$\|\nabla_\Omega f(x_c^{l_i})\| \to 0.$$

Pelo Lema 1.24, $\|\nabla_\Omega f(x^*)\| = 0$, ou seja, x^* é estacionário.

Vamos supor agora que o limite (1.92) não é válido. Então existe $\varepsilon_0 > 0$ tal que

$$\frac{\|d^{l_i}(\hat{\alpha}_{l_i})\|}{\hat{\alpha}_{l_i}} \geq \varepsilon_0,$$

para todo l_i suficientemente grande. Seja $\delta \in (0, \varepsilon_0)$. O Teorema 1.26 garante que existe uma sequência de índices (n_i) tal que

$$\frac{\|d^k(\hat{\alpha}_k)\|}{\hat{\alpha}_k} \geq \delta, \quad \text{para} \quad l_i \leq k < n_i \quad \text{e} \quad \frac{\|d^{n_i}(\hat{\alpha}_{n_i})\|}{\hat{\alpha}_{n_i}} < \delta. \tag{1.93}$$

Seja k uma iteração de sucesso, isto é, $\rho_k > \eta$. Então, da definição de ρ_k, da condição (1.40) sobre o modelo m_k e da estimativa de redução obtida pelo Lema 1.21, podemos concluir que existe uma contante $c > 0$ tal que

$$f(x^k) - f(x^{k+1}) > \eta c \mu_0^2 \delta \min\left\{\Delta_k, \frac{\delta}{\beta_k}\right\},$$

sempre que k é uma iteração de sucesso e $l_i \leq k < n_i$. Da condição (1.60)

$$\|x^{k+1} - x^k\| = \|d^k\| \leq \mu_2 \Delta_k.$$

Além disso, $(f(x^k))$ é não crescente e limitada inferiormente. Portanto, convergente. Ainda, estamos supondo (β_k) limitada superiormente. Sendo assim, estas hipóteses implicam que existe $c_1 > 0$ tal que

$$f(x^k) - f(x^{k+1}) > c_1\|x^{k+1} - x^k\|,$$

para k suficientemente grande. Porém, se k não é uma iteração de sucesso, isto é, $\rho_k \leq \eta$, então $x^{k+1} = x^k$. Logo

$$f(x^k) - f(x^{k+1}) \geq c_1\|x^{k+1} - x^k\|,$$

para todo $l_i \leq k < n_i$. Fazendo $k = l_i, l_i + 1, \ldots, n_i - 1$ na desigualdade acima, é fácil ver que

$$f(x^{l_i}) - f(x^{n_i}) \geq c_1\|x^{l_i} - x^{n_i}\|.$$

Novamente, usando a convergência da sequência $(f(x^k))$ e, além disso, a convergência de (x^{l_i}), concluímos que $x^{n_i} \to x^*$.

A convergência da sequência (x^{n_i}) e a condição (1.93) nos permite concluir que, para cada $\delta \in (0, \varepsilon_0)$, existe um conjunto infinito $\mathcal{H} \subset \mathbb{N}$, tal que

$$\frac{\|d^k(\hat{\alpha}_k)\|}{\hat{\alpha}_k} < \delta \quad \text{e} \quad \|x^k - x^*\| < \delta, \tag{1.94}$$

para todo $k \in \mathcal{H}$. Caso $\delta > \varepsilon_0$, existe $r > 0$ tal que $r\delta < \varepsilon_0$. Então, (1.94) também é válido. Sendo assim, existe uma subsequência (x^{k_i}) convergente para x^* e

$$\frac{\|d^{k_i}(\hat{\alpha}_{k_i})\|}{\hat{\alpha}_{k_i}} \to 0.$$

Este limite implica que podemos utilizar os mesmos argumentos do primeiro caso desta demonstração para a sequência (x^{k_i}) (veja a hipótese (1.92)), e concluir que x^* é estacionário. □

O teorema anterior mostrou que existe uma subsequência (x^{k_i}) convergente para x^*, ponto estacionário. No entanto, esta sequência não é, necessariamente, indexada por iterações de sucesso, isto é, não garantimos que $\rho_{k_i} > \eta$, para todo k_i. Para garantir esse resultado, primeiramente, em [15, Teorema 5.3], é demonstrado que (1.83) também vale para iterações de sucesso. Mais precisamente, se

$$\mathcal{J} = \{k \in \mathbb{N} \mid \rho_k > \eta\}$$

então

$$\liminf_{k \in \mathcal{J}} \frac{\|d^k(\hat{\alpha}_k)\|}{\hat{\alpha}_k} = 0. \tag{1.95}$$

Para isso, além das hipóteses do Teorema 1.26, considera-se a hipótese de que a convergência de série (1.78) implica que a sequência de modelos (m_k) satisfaz

$$\lim_{k \to \infty} \frac{f(x^k + d^k) - f(x^k) - m_k(d^k)}{\|d^k\|^2} = 0. \tag{1.96}$$

Note que, quando $B_k = \nabla^2 f(x^k)$, esta hipótese é satisfeita, uma vez que a convergência da série (1.78) implica que $\|d^k\| \to 0$. Sendo assim, da fórmula de Taylor de segunda ordem (1.10), o limite (1.96) é válido. A convergência global para esse caso é estabelecida pelo teorema seguinte.

Teorema 1.29. *[13, Teorema 2.1] Suponha as mesmas hipóteses do Teorema 1.28. Então existe uma subsequência (x^{k_i}), indexada por iterações de sucesso, que converge para $x^* \in \Omega$ satisfazendo*

$$\lim_{i\to\infty} \|\nabla_\Omega f(x_c^{k_i})\| = 0,$$

em que (x_c^k) é a sequência definida por (1.70). Além disso, $x_c^{k_i} \to x^*$ e, portanto, x^* é estacionário para o problema (1.36).

Demonstração. A demonstração é semelhante à demonstração do teorema anterior. Basta supor a subsequência (x^{l_i}) indexada por iterações de sucesso convergindo para x^*. Então, definindo $\mathcal{J} = \{k \in \mathbb{N} \mid \rho_k > \eta\}$, utilizamos (1.95) para garantir que os índices (n_i) podem ser escolhidos como iterações de sucesso. □

Na última seção deste capítulo, apresentamos algumas propriedades geométricas do Algoritmo 3.

1.4.4 Propriedades geométricas do algoritmo de região de confiança restrito

Para encerrar este capítulo, estudamos propriedades geométricas do Algoritmo 3 que motivam alguns procedimentos computacionais para obter o valor do passo d^k, que serão apresentados no Capítulo 3. Todos os teoremas desta seção são baseados nos resultados de [13] e [14], e são válidos para o caso em que Ω é o conjunto definido pelas m restrições lineares

$$\Omega = \{x \in \mathbb{R}^n \mid c_j^T x \geq \delta_j,\ 1 \leq j \leq m\}, \tag{1.97}$$

para quaisquer vetores $c_j \in \mathbb{R}^n$ e escalares δ_j. Note que a caixa (1.2) é um caso particular de (1.97). De fato, basta considerar os vetores canônicos $e_1, \ldots, e_n$ de $\mathbb{R}^n$ e fazer

$$\begin{cases} c_j = e_j, \\ \delta_j = \ell_j, \end{cases} \tag{1.98}$$

caso $j = 1, \ldots, n$, e

$$\begin{cases} c_j = -e_{j-n}, \\ \delta_j = -u_{j-n}, \end{cases} \tag{1.99}$$

caso $j = n+1, \ldots, 2n$.

Definimos, primeiramente, um conjunto importante para o desenvolvimento das propriedades geométricas do Algoritmo 3.

Definição 1.30. *Considere o conjunto Ω descrito em (1.97) e $x \in \Omega$. O conjunto ativo em x é denotado por $A(x)$ e é dado por*

$$A(x) = \{j \in \{1, \ldots, m\} \mid c_j^T x = \delta_j\}.$$

Definimos também o conceito de *face*.

Definição 1.31. *Considere o conjunto Ω descrito em (1.97). Uma face F de Ω é o conjunto*

$$F = \{x \in \Omega \mid c_j^T x = \delta_j,\ j \in A(F)\}, \tag{1.100}$$

para algum conjunto de índices $A(F) \subset \{1, \ldots, m\}$.

Em [14], a definição de face é apresentada também para conjunto mais gerais, isto é, para o caso em que Ω é um conjunto convexo qualquer. Na Figura 1.12, exemplificamos uma face F para o caso em que Ω é da forma (1.2), com $n = 2$.

Figura 1.12: Exemplo de uma face da forma (1.100). Fonte: a autora.

Nesta seção, relacionamos a condição necessária de primeira ordem dada pelo Teorema 1.12 e a convergência da sequência (x^k) gerada pelo Algoritmo 3, com o conceito de *face exposta*.

Definição 1.32. *Sejam $\Omega \subset \mathbb{R}^n$ um conjunto convexo e $d \in \mathbb{R}^n$. A face exposta por d é denotada por $E(d)$ e é dado pelo conjunto*

$$E[d] = \arg\max\{d^T x \mid x \in \Omega\}.$$

Este conjunto esta relacionado fortemente com o cone normal, definido na Seção 1.1. Note que, por definição, $d \in N_\Omega(x)$ se, e somente se, $d^T y \leq d^T x$, para qualquer $y \in \Omega$. Consequentemente,

$$d \in N_\Omega(x) \Leftrightarrow x \in E[d]. \tag{1.101}$$

Podemos mostrar como a condição necessária de primeira ordem (1.12) para $x^* \in \Omega$ está relacionada com a face exposta $E[-\nabla f(x^*)]$, utilizando (1.101) diretamente.

Teorema 1.33. [14, Lema 2.2] *Considere o problema (1.1) e $x^* \in \Omega$. Então $x^* \in E[-\nabla f(x^*)]$ se, e somente se, x^* é estacionário para este problema.*

Demonstração. Recorde que, se x^* é estacionário então $\nabla f(x^*)^T(x - x^*) \geq 0$. Logo, basta aplicar (1.101) para $d = -\nabla f(x^*)$. □

É possível mostrar que toda face é uma face exposta e vice-versa (veja [22, Teorema 2.4.12]), sempre que Ω é o conjunto de m restrições (1.97). Quando o conjunto Ω é a caixa (1.2) e $d = -\nabla f(x^*)$, para $x^* \in \Omega$ fixado, temos

$$E[-\nabla f(x^*)] = \{x \in \Omega \mid x_i = l_i, \text{ caso } \partial_i f(x^*) > 0 \text{ e } x_i = u_i, \text{ caso } \partial_i f(x^*) < 0\}. \tag{1.102}$$

A Figura 1.13 destaca o conjunto $E[-\nabla f(x^*)]$ para este caso com $n = 2$.

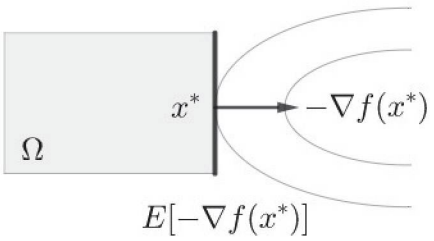

Figura 1.13: Exemplo de uma face exposta da forma (1.102). Fonte: a autora.

Nos próximos teoremas, vamos mostrar que, sob determinadas hipóteses, podemos garantir que a sequência (x^k) gerada pelo Algoritmo 3 permanece em uma face exposta. Vamos ver como isso se dá em termos de restrições ativas. O primeiro resultado que enunciamos é demonstrado em [14].

Teorema 1.34. [14, Teorema 4.2] *Considere o problema (1.1), com Ω dado por (1.97). Suponha que $f : \mathbb{R}^n \to \mathbb{R}$ é continuamente diferenciável em Ω e que $(x^k) \subset \Omega$ é uma sequência convergente para um ponto estacionário x^*. Então*

$$\lim_{k\to\infty} \|\nabla_\Omega f(x^k)\| = 0,$$

se, e somente se, existe um índice k_0 tal que

$$x^k \in E[-\nabla f(x^*)],$$

para todo $k \geq k_0$.

Utilizando esse resultado, conseguimos garantir que a sequência (x_c^k) dada por

$$x_c^k = x^k + d^k(\hat{\alpha}_k),$$

estudada na seção anterior, permanece na face exposta $E[-\nabla f(x^*)]$.

Teorema 1.35. [13, Teorema 3.2] *Considere o problema (1.1), com Ω dado por (1.97). Suponha que $f : \mathbb{R}^n \to \mathbb{R}$ é continuamente diferenciável em Ω e que $(x^k) \subset \Omega$ é uma sequência convergente para um ponto estacionário x^*. Se existe um índice k_0 tal que*

$$x^k \in E[-\nabla f(x^*)],$$

para todo $k \geq k_0$, então

$$x_c^k = P_\Omega\left(x^k - \hat{\alpha}_k \nabla f(x^k)\right) \in E[-\nabla f(x^*)],$$

para todo k suficientemente grande.

Demonstração. Como consequência da definição do cone normal e da projeção, temos

$$x^k - \hat{\alpha}_k \nabla f(x^k) - P_\Omega\left(x^k - \hat{\alpha}_k \nabla f(x^k)\right) \in N_\Omega\left(P_\Omega\left(x^k - \hat{\alpha}_k \nabla f(x^k)\right)\right).$$

Sendo o cone normal um cone, então, definindo

$$v^k = -\nabla f(x^k) + \frac{x^k - P_\Omega\left(x^k - \hat{\alpha}_k \nabla f(x^k)\right)}{\hat{\alpha}_k},$$

segue que

$$P_\Omega\left(x^k - \hat{\alpha}_k \nabla f(x^k)\right) \in E[v^k], \tag{1.103}$$

pois vale a equivalência (1.101).

Por outro lado, se $v^k \to v^*$ então $E[v^k] \subset E[v^*]$, para todo k suficientemente grande (veja [14, Teorema 3.1]). Vamos mostrar que

$$v^k \to -\nabla f(x^*). \tag{1.104}$$

Como $\hat{\alpha}_k > 0$, usando o Lema 1.17 e o Lema 1.22, item (b), temos a desigualdade

$$\frac{\|P_\Omega\left(x^k - \hat{\alpha}_k \nabla f(x^k)\right) - x^k\|}{\hat{\alpha}_k} \leq \|\nabla_\Omega f(x^k)\|.$$

Das hipóteses de que $x^k \to x^*$ e $x^k \in E[-\nabla f(x^*)]$, para todo $k \geq k_0$, segue do teorema anterior que o lado direito da desigualdade acima converge para zero. Logo, da definição de v^k e sendo f continuamente diferenciável, concluímos que vale a convergência (1.104). Sendo assim, (1.103) mostra que

$$P_\Omega\left(x^k - \hat{\alpha}_k \nabla f(x^k)\right) \ \in \ E[v^*],$$

para todo k suficientemente grande, concluíndo a demonstração. □

Supondo x^* estacionário, podemos caracterizar o conjunto $E[-\nabla f(x^*)]$ por meio das condições de otimalidade de Karush-Kuhn-Tucker. No Capítulo 2 deste trabalho, essas condições são descritas com mais detalhes para problemas de minimização mais gerais. Quando o conjunto Ω é dado por (1.97), essas condições implicam que existem escalares $\lambda_1, \ldots, \lambda_m$ não negativos, tais que

$$\nabla f(x^*) = \sum_{j \in A(x^*)} \lambda_j^* c_j.$$

Então, em [14, Teorema 4.4], mostra-se que

$$E[-\nabla f(x^*)] = \{x \in \Omega \ \mid \ c_j^T x = \delta_j, \text{ caso } \lambda_j^* > 0\}.$$

Portanto, dados $x, y \in \Omega$,

$$x \in E[-\nabla f(x^*)] \quad \text{e} \quad A(x) \subset A(y) \quad \Rightarrow \quad y \in E[-\nabla f(x^*)]. \tag{1.105}$$

O teorema enunciado anteriormente, juntamente com o resultado de convergência global do Algoritmo 3 (Teorema 1.29) diz que, se (x^k) converge para x^*, então $x_c^k \in E[-\nabla f(x^*)]$, para alguma subsequência de iterações de sucesso. Podemos supor condições sobre o conjunto ativo $A(x^k + d^k)$ para garantir que $x^{k+1} \in E[-\nabla f(x^*)]$, sempre que (x^k) é convergente para x^*.

Teorema 1.36. [13, Teorema 3.3] *Considere as mesmas hipóteses do Teorema 1.29 e o conjunto Ω dado por (1.97). Suponha que as sequências (x^k) e (d^k) geradas pelo Algoritmo 3 satisfazem*

$$A(x_c^k) \subset A(x^k + d^k), \tag{1.106}$$

para todo $k \geq 0$, sendo (x_c^k) definida em (1.70). Se $x^k \to x^*$ então existe um índice k_0 tal que

$$x^k \in E[-\nabla f(x^*)] \quad e \quad x^k + d^k \in E[-\nabla f(x^*)],$$

para todo $k \geq k_0$.

Demonstração. Primeiramente, segue do resultado de convergência global (Teorema 1.29) que existe um subconjunto infinito $\mathcal{K}$ de iterações de sucesso tal que $x_c^k \to x^*$ e $\nabla_\Omega f(x_c^k) \to 0$, para $k \in \mathcal{K}$. Além disso, x^* é estacionário. Sendo assim, do Teorema 1.34, podemos considerar que $x_c^k \in E[-\nabla f(x^*)]$, para $k \in \mathcal{K}$. Da hipótese (1.106) e da relação (1.105), concluímos que

$$x^{k+1} \in E[-\nabla f(x^*)], \ \forall \ k \in \mathcal{K},$$

pois toda iteração $k \in \mathcal{K}$ é de sucesso. Em particular, existe um índice k_0 tal que $x^{k_0} \in E[-\nabla f(x^*)]$. Note que $x^{k+1} = x^k$ ou $x^{k+1} = x^k + d^k$. Por indução em $k \geq k_0$, podemos utilizar o Teorema 1.35 e, novamente a hipótese (1.106) e a relação (1.105), para obter

$$x^k \in E[-\nabla f(x^*)], \ \forall \ k \geq k_0. \tag{1.107}$$

Para concluir a demonstração, queremos garantir que

$$x^k + d^k \in E[-\nabla f(x^*)], \ \forall \ k \geq k_0. \tag{1.108}$$

Para isso, basta utilizar o que acabamos de provar em (1.107), e, novamente, o Teorema 1.35 e a hipótese (1.106). □

O conjunto $E[-\nabla f(x^*)]$ é utilizado no desenvolvimento da teoria de convergência do método quando a matriz B_k é a Hessiana $\nabla^2 f(x^k)$. Nesse caso, em [13, Seção 5] o método é considerado uma versão de Newton para região de confiança. Nesta teoria de convergência é considerada a hipótese de que a matriz Hessiana de f seja limitada no conjunto de nível $\{x \in \Omega \ \mid \ f(x) \leq f(x^0)\}$ e que satisfaça uma condição chamada de *condição suficiente de segunda ordem forte*. Esta condição pede que a matriz Hessiana de f em um ponto limite x^* da sequência (x^k) gerada pelo Algoritmo 3, seja definida positiva no menor subespaço afim que contenha $E[-\nabla f(x^*)] - x^*$. Daí, mostra-se que (x^k) converge para x^* (veja [13, Teorema 5.1]).

No Capítulo 3 descrevemos os detalhes de implementação do Algoritmo 3 para o caso em que o conjunto Ω é a caixa (1.2). Descrevemos como gerar o passo d^k garantindo que a hipótese (1.106) do teorema anterior seja satisfeita.

Capítulo 2

Método de Lagrangiano aumentado para problemas com restrições de igualdade e de caixa

No capítulo anterior estudamos um método de otimização para problemas restritos, supondo o conjunto viável convexo e fechado. Em particular, quando o conjunto formado pelas restrições do problema é da forma

$$\Omega = \{x \in \mathbb{R}^n \mid \ \ell \leq x \leq u\}, \tag{2.1}$$

para quaisquer vetores $\ell \in \mathbb{R}^n$ e $u \in \mathbb{R}^n$, vimos que, alguns cálculos requeridos pelos algoritmos apresentados tornam-se mais simples e estabelecemos boas propriedades geométricas.

Neste capítulo, apresentamos o método de Lagrangiano aumentado para resolver um problema de otimização bastante comum em muitas aplicações, cujo formato é

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && h(x) = 0, \\ &&& x \in \Omega, \end{aligned} \tag{2.2}$$

supondo $f : \mathbb{R}^n \to \mathbb{R}$ e $h : \mathbb{R}^n \to \mathbb{R}^p$ funções diferenciáveis e Ω a caixa (2.1). Note que, quando um problema possui restrições de desigualdade, este pode ser escrito na forma (2.2) através do uso de variáveis de folga. No entanto, nem sempre temos a garantia de que o conjunto Ω com as variáveis de folga inclusas seja um conjunto fechado. No método de Lagrangiano aumentado, o problema (2.2) é resolvido de forma iterativa, sendo que a cada iteração um subproblema de otimização mais simples é resolvido, onde são consideradas somente as restrições de caixa. O tratamento das restrições de igualdade é feito através de estratégias de métodos de penalização quadrática ao problema

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && h(x) = 0. \end{aligned} \tag{2.3}$$

Dividimos este capítulo em duas partes. Na Seção 2.1 fazemos um breve resumo das condições de otimalidade de primeira e segunda ordem para problemas de otimização mais gerais, estudadas em [7] e [6]. Vemos também como as condições de primeira ordem são

reescritas para o problema de nosso interesse (2.2). Na Seção 2.2 estudamos o método de Lagrangiano aumentado clássico para o problema com restrições de igualdade (2.3) e em seguida apresentamos uma versão do método, proposta em [3], para resolver o problema (2.2).

2.1 Condições de otimalidade

Considere $f : \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável e o problema de otimização na forma padrão

$$\begin{aligned} &\text{minimizar} \quad f(x) \\ &\text{sujeito a} \quad x \in D, \end{aligned} \tag{2.4}$$

em que $D \in \mathbb{R}^n$ é o conjunto formado por restrições de igualdade e desigualdade

$$D = \{x \in \mathbb{R}^n \mid h(x) = 0 \text{ e } g(x) \leq 0\}, \tag{2.5}$$

com $h : \mathbb{R}^n \to \mathbb{R}^p$ e $g : \mathbb{R}^n \to \mathbb{R}^m$ funções dadas. As condições de otimalidade de primeira e segunda ordem serão enunciadas nesta seção em termos de alguns conceitos importantes na teoria de otimização não linear. Primeiramente, definimos a *função Lagrangiana* e o *conjunto ativo* para restrições de desigualdade, respectivamente.

Definição 2.1. *A função Lagrangiana associada ao problema (2.4) é a função* $L : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \to \mathbb{R}$, *dada por*

$$L(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x). \tag{2.6}$$

Definição 2.2. *Seja* $x^* \in D$, *sendo* D *o conjunto definido em (2.5). O conjunto ativo em* x^*, *com relação à função* g, *é o conjunto de índices*

$$I(x^*) = \{i \in \{1, \ldots, m\} \mid g_i(x^*) = 0\}. \tag{2.7}$$

Quando $i \in I(x^*)$, dizemos que a restrição de desigualdade g_i é *ativa* em x^*. Além disso, sendo $x^* \in D$, também dizemos que a restrição h_i é ativa em x^*, para $i = 1, \ldots, p$. Para as condições de otimalidade de segunda ordem, definimos o *cone crítico* em x^*.

Definição 2.3. *Considere o problema (2.4) e o conjunto* D *definido em (2.5) com* h *e* g *funções diferenciáveis em* $x^* \in D$. *O cone crítico em* x^* *é o conjunto de direções*

$$C(x^*) = \left\{ d \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \nabla f(x^*)^T d \leq 0, \\ \nabla g_i(x^*)^T d \leq 0, \ i \in I(x^*), \\ \nabla h_i(x^*)^T d = 0, \ i = 1, \ldots, p. \end{array} \right\}. \tag{2.8}$$

Uma condição de qualificação é uma propriedade desejável para pontos viáveis de um problema de minimização. Quando esta propriedade é satisfeita por um minimizador local, garantimos o cumprimento das *condições de otimalidade de Lagrange*, também conhecidas como as *condições de Karush-Kuhn-Tucker (KKT)*, que veremos mais adiante. Uma condição de qualificação bastante utilizada para o problema (2.4) é a *condição de qualificação de independência linear (LICQ)*. Nesta condição, os gradientes das restrições

ativas em um ponto viável são linearmente independentes, isto é, $x^* \in D$, com D definido em (2.5), satisfaz LICQ se

$$\{\nabla h_i(x^*), i = 1, \ldots, p\} \cup \{\nabla g_i(x^*), i \in I(x^*)\} \text{ é linearmente independente,} \tag{2.9}$$

em que $I(x^*)$ é o conjunto ativo definido em (2.7). Outras condições de qualificação mais fracas do que a condição acima podem ser consideradas no desenvolvimento das condições de otimalidade para o problema (2.4). Estudos mais detalhados a respeito de condições de qualificação são encontrados em [23] e [24].

2.1.1 Condições necessárias de primeira ordem

Podemos enunciar agora as condições necessárias de otimalidade de primeira ordem de Karush-Kuhn-Tucker para o problema na forma padrão (2.4).

Teorema 2.4. [6, Teorema 4.2.2] *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ e $g : \mathbb{R}^n \to \mathbb{R}^m$ funções diferenciáveis e a função $h : \mathbb{R}^n \to \mathbb{R}^p$ diferenciável em uma vizinhança de $x^* \in D$, conjunto definido por (2.5). Suponha que h possui derivada contínua em x^* e que este é um minimizador local para o problema (2.4). Se vale uma condição de qualificação em x^*, então existem $\lambda^* \in \mathbb{R}^p$ e $\mu^* \in \mathbb{R}^m_+$ tais que*

$$\nabla_x L(x^*, \lambda^*, \mu^*) = 0 \tag{2.10}$$

e

$$\mu_i^* g_i(x^*) = 0, \ i = 1, \ldots, m, \tag{2.11}$$

sendo L a função Lagrangiana definida em (2.6). Além disso, se vale (2.9) em x^* então o par (λ^*, μ^*) satisfazendo as condições acima é único.

Quando existem $x^* \in D$, $\lambda^* \in \mathbb{R}^p$ e $\mu^* \in \mathbb{R}^m_+$ satisfazendo as igualdades (2.10) e (2.11), o ponto x^* é chamado de *ponto de Karush-Kuhn-Tucker* (ou ponto KKT), ou ainda, um ponto *estacionário*, e os vetores λ^* e μ^* são chamados de *multiplicadores de Lagrange*.

Grande parte dos algoritmos que buscam minimizadores locais de problemas restritos procuram pontos viáveis que satisfaçam as condições de KKT. Para o problema de nosso interesse (2.2), estas condições podem ser reescritas em termos da projeção no conjunto Ω. Considere a função Lagrangiana associada somente às restrições de igualdade, isto é, $L_h : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$, dada por

$$L_h(x, \lambda) = f(x) + \lambda^T h(x). \tag{2.12}$$

Então, utilizando as condições (2.10) e (2.11) do teorema anterior para o problema (2.2) e a expressão da projeção (1.7), mostra-se que as condições de KKT para este problema são equivalentes a

$$x^* = P_\Omega\big(x^* - \nabla_x L_h(x^*, \lambda^*)\big) \quad \text{e} \quad h(x^*) = 0, \tag{2.13}$$

com $\lambda^* \in \mathbb{R}^p$ multiplicador de Lagrange associado às restrições de igualdade. Demonstramos este fato no teorema a seguir.

Teorema 2.5. *Suponha x^* um ponto viável para o problema (2.2). Então, x^* satisfaz as condições de KKT para este problema, com λ^* um multiplicador de Lagrange associado às restrições de igualdade, se, e somente se, vale (2.13).*

Demonstração. Primeiramente, suponha que x^* e λ^* satisfazem as condições de KKT. Pelo Teorema 2.4, existem vetores $\mu^* \in \mathbb{R}^n$ e $\sigma^* \in \mathbb{R}^n$ tais que

$$\nabla_x L_h(x^*, \lambda^*) + \mu^* - \sigma^* = 0 \tag{2.14a}$$
$$(x_i^* - u_i)\mu_i^* = 0, \ i = 1, \ldots, n \tag{2.14b}$$
$$(\ell_i - x_i^*)\sigma_i^* = 0, \ i = 1, \ldots, n \tag{2.14c}$$
$$\mu_i^* \geq 0, \ i = 1, \ldots, n \tag{2.14d}$$
$$\sigma_i^* \geq 0, \ i = 1, \ldots, n \tag{2.14e}$$

onde L_h é definida em (2.12). Seja $i \in \{1, \ldots, n\}$. Caso $x_i^* = \ell_i$ então $x_i^* < u_i$. Logo, por (2.14b), $\mu_i^* = 0$. Substituindo em (2.14a) e usando (2.14e), segue que

$$\big(\nabla_x L_h(x^*, \lambda^*)\big)_i = \sigma_i^* \geq 0.$$

Multiplicando esta desigualdade por -1 e somando x_i^* em ambos os lados, ficamos com

$$x_i^* - \big(\nabla_x L_h(x^*, \lambda^*)\big)_i \leq x_i^* = \ell_i.$$

Da expressão (1.7) para o cálculo da projeção, segue que

$$\big(P_\Omega\big(x^* - \nabla_x L_h(x^*, \lambda^*)\big)\big)_i = \ell_i. \tag{2.15}$$

Caso $x_i^* = u_i$, mostra-se de forma análoga ao caso anterior, que

$$\big(P_\Omega\big(x^* - \nabla_x L_h(x^*, \lambda^*)\big)\big)_i = u_i. \tag{2.16}$$

Por outro lado, se $\ell_i < x_i^* < u_i$, de (2.14b) e (2.14c), concluímos que $\mu_i^* = \sigma_i^* = 0$ e, substituindo em (2.14a), temos

$$\big(\nabla_x L_h(x^*, \lambda^*)\big)_i = 0.$$

Logo,

$$\ell_i < x_i^* - \big(\nabla_x L_h(x^*, \lambda^*)\big)_i < u_i.$$

Isso implica que

$$\big(P_\Omega\big(x^* - \nabla_x L_h(x^*, \lambda^*)\big)\big)_i = x_i^*. \tag{2.17}$$

Portanto, da viabilidade de x^* e das possibilidades para x_i^*, (2.15), (2.16) e (2.17), vale (2.13). Suponha agora que vale (2.13) para algum λ^*. Vamos mostrar que existem μ^* e σ^* tais que valem as condições (2.14a)-(2.14e). Seja $i \in \{1, \ldots, n\}$. Se

$$x_i^* - \big(\nabla_x L_h(x^*, \lambda^*)\big)_i \leq \ell_i$$

então, da hipótese (2.13), vale que $x_i^* = \ell_i$. Daí, $\big(\nabla_x L_h(x^*, \lambda^*)\big)_i \geq 0$. Sendo assim, defina

$$\mu_i^* = 0 \quad \text{e} \quad \sigma_i^* = \big(\nabla_x L_h(x^*, \lambda^*)\big)_i.$$

Se, do contrário, for verdade que

$$x_i^* - \big(\nabla_x L_h(x^*, \lambda^*)\big)_i \geq u_i,$$

temos $-\big(\nabla_x L_h(x^*, \lambda^*)\big)_i \geq 0$. Daí, defina

$$\sigma_i^* = 0 \quad \text{e} \quad \mu_i^* = -\big(\nabla_x L_h(x^*, \lambda^*)\big)_i.$$

Por fim, se

$$\ell_i < x_i^* - \big(\nabla_x L_h(x^*, \lambda^*)\big)_i < u_i$$

defina

$$\mu_i^* = 0 \quad \text{e} \quad \sigma_i^* = 0.$$

Como i é qualquer, segue das conclusões anteriores que as condições de KKT são satisfeitas para x^* e λ^*. □

2.1.2 Condições necessárias e suficientes de segunda ordem

Nesta seção apresentamos as condições de otimalidade de segunda ordem para o problema (2.4). Os estudos detalhados destas condições são encontrados nas referências [7] e [6]. Nos próximos teoremas enunciamos as condições necessárias e as condições suficientes, respectivamente.

Teorema 2.6. [6, Teorema 4.3.2] *Sejam $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ e $h : \mathbb{R}^n \to \mathbb{R}^p$ funções duas vezes diferenciáveis em x^*, um minimizador local para o problema (2.4). Suponha que LICQ (2.9) é satisfeita em x^*. Então*

$$d^T \nabla_x^2 L(x^*, \lambda^*, \mu^*) d \geq 0, \ \forall\ d \in C(x^*),$$

em que $C(x^*)$ é o cone crítico e L é a função Lagrangiana, definidos em (2.8) e (2.6), respectivamente, e os vetores $\lambda^* \in \mathbb{R}^p$ e $\mu^* \in \mathbb{R}^m_+$ são os multiplicadores de Lagrange associados.

Teorema 2.7. [7, Teorema 1.2.3 (b)] *Sejam $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ e $h : \mathbb{R}^n \to \mathbb{R}^p$ funções duas vezes diferenciáveis em x^*, um ponto viável para o problema (2.4), e a função Lagrangiana associada, dada em (2.6). Se existem vetores $\lambda^* \in \mathbb{R}^p$ e $\mu^* \in \mathbb{R}^m_+$ satisfazendo as condições do Teorema 2.4, e $\gamma > 0$ tais que*

$$d^T \nabla_x^2 L(x^*, \lambda^*, \mu^*) d \geq \gamma \|d\|^2, \ \forall\ d \in C(x^*), \tag{2.18}$$

em que $C(x^*)$ é o cone crítico definido em (2.8), então x^* é um minimizador local estrito.

Observamos que para o problema com restrições somente de igualdade (2.3), em [6, Teorema 2.3.2] mostra-se que a condição suficiente de segunda ordem (2.18) é reescrita como

$$d^T \nabla_x^2 L_h(x^*, \lambda^*) d \geq \gamma \|d\|^2, \ \forall\ d \in \ker\ J_h(x^*) \setminus \{0\}, \tag{2.19}$$

sendo L_h a função Lagrangiana (2.12) e $J_h(x^*)$ a matriz Jacobiana com relação às restrições de igualdade. Esta condição é utilizada para justificar propriedades boas do método de Lagrangiano aumentado, descrito nas seções seguintes.

2.2 Método de Lagrangiano aumentado

Nesta seção estudamos o método de Lagrangiano aumentado para resolver o problema (2.2) com restrições de igualdade e de caixa. Inicialmente, este método foi desenvolvido considerando-se o problema com apenas restrições de igualdade (2.3), em que uma estratégia de penalização quadrática dessas restrições é utilizada. Descrevemos a seguir esta penalização.

2.2.1 Penalização quadrática para problemas com restrições de igualdade

Considere o problema com restrições de igualdade (2.3). No método de penalização quadrática, este problema é resolvido através de uma sequência de subproblemas de minimização irrestrita. Seja $k \geq 1$. Na k-ésima iteração do método, dado $c_k > 0$, a função objetivo do subproblema é dada por $Q(\cdot, c_k) : \mathbb{R}^n \to \mathbb{R}$, sendo

$$Q(x, c_k) = f(x) + \frac{c_k}{2}\|h(x)\|^2. \tag{2.20}$$

O valor c_k é chamado de parâmetro de penalidade. Note que, para pontos viáveis, a função acima coincide com a função objetivo do problema original.

No método de penalização quadrática, a violação na restrição de igualdade é penalizada por meio do crescimento do parâmetro de penalidade. Sendo assim, o segundo termo da função de penalidade Q aumenta com o crescimento de c_k para pontos não viáveis. Apresentamos o algoritmo básico do método de penalização quadrática.

Algoritmo 4 Penalização quadrática

Entrada: $c_1 > 0$, $k = 1$ e $x^0 \in \mathbb{R}^n$.

repita

1. Resolva (possivelmente de forma aproximada) o subproblema irrestrito

$$\begin{array}{ll} \text{minimizar} & Q(x, c_k) \\ \text{sujeito a} & x \in \mathbb{R}^n, \end{array} \tag{2.21}$$

para Q, função definida em (2.20), e obtenha x^k.
2. Escolha $c_{k+1} \geq c_k$.
3. Faça $k = k + 1$.

até x^k estacionário para o problema (2.3);

Em geral, no Passo 1 do algoritmo anterior, usa-se x^{k-1} como ponto inicial na resolução do subproblema. A atualização do parâmetro de penalidade pode ser feita com a utilização da regra $c_{k+1} = 10c_k$, (veja [8]). A seguir, enunciamos o teorema de convergência para pontos estacionários.

Teorema 2.8. [7, Teorema 4.4.2] *Considere o problema com restrições de igualdade (2.3), com f e h continuamente diferenciáveis e as sequências $(x^k) \subset \mathbb{R}^n$ e $(c_k) \subset \mathbb{R}_+$ geradas pelo Algoritmo 4. Suponha que x^* é um ponto de acumulação de (x^k) satisfazendo LICQ*

(2.9) e que $c_k \to \infty$. Então x^* satisfaz as condições de otimalidade para o problema (2.3), dadas pelo Teorema 2.4. Além disso, para qualquer subsequência (x^{k_j}) que converge para x^*, vale que

$$c_{k_j} h(x^{k_j}) \xrightarrow{j} \lambda^*, \tag{2.22}$$

sendo $\lambda^* \in \mathbb{R}^p$ o (único) multiplicador de Lagrange associado.

O crescimento ilimitado do parâmetro de penalidade pode tornar a Hessiana da função Q mal condicionada, dificultando a resolução do subproblema (2.21), como ilustra [25, Exemplo 1.5]. O método de Lagrangiano aumentado busca contornar este problema por meio da adição de mais um termo à função Q, que depende de uma aproximação do multiplicador de Lagrange. Na seção seguinte apresentamos como funciona o método de Lagrangiano aumentado.

2.2.2 Método clássico de Lagrangiano aumentado

Apresentamos agora o método de Lagrangiano aumentado para o problema com restrições de igualdade (2.3). A ideia deste método é adicionar o termo de penalização das restrições à função Lagrangiana associada ao problema. Recorde que para este caso a função Lagrangiana é dada pela fórmula (2.12). Inicialmente, para entendermos o funcionamento do método, definimos a função *Lagrangiana aumentada*

$$\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}_+ \to \mathbb{R},$$

como

$$\mathcal{L}(x, \lambda, c) = f(x) + \lambda^T h(x) + \frac{c}{2}\|h(x)\|^2. \tag{2.23}$$

Note que, a função Lagrangiana aumentada é a Lagrangiana associada ao problema

$$\begin{aligned} \text{minimizar} \quad & f(x) + \frac{c}{2}\|h(x)\|^2 \\ \text{sujeito a} \quad & h(x) = 0. \end{aligned}$$

Além disso, o problema acima é equivalente ao problema original (2.3), uma vez que ambas as funções objetivo coincidem no conjunto viável.

Uma justificativa para a construção do método de Lagrangiano aumentado é apresentada em [7] da seguinte maneira. O Teorema 2.4 implica que as condições de KKT para o problema (2.3) são

$$\nabla_x L_h(x, \lambda) = \nabla f(x) + J_h(x)^T \lambda = 0 \tag{2.24}$$

e

$$h(x) = 0. \tag{2.25}$$

Portanto, gostaríamos de resolver as equações acima e, se conhecermos um multiplicador de Lagrange ótimo, esta tarefa pode se tornar mais fácil. Porém, nem sempre um ponto estacionário x^* do problema original é um minimizador da função $L_h(\cdot, \lambda^*)$, com λ^* sendo o multiplicador de Lagrange associado (veja [7, Exemplo 4.4.1]). Sob determinadas hipóteses, a busca por minimizadores da função Lagrangiana aumentada pode ser feita. Isso pode ser justificado no próximo teorema.

Teorema 2.9. [7, Teorema 4.4.5] *Considere o problema com restrições de igualdade (2.3), com f e h continuamente diferenciáveis no ponto estacionário $x^* \in \mathbb{R}^n$. Suponha que x^* satisfaz a condição suficiente de segunda ordem (2.19), para algum $\lambda^* \in \mathbb{R}^p$. Então para todo $c > 0$ suficientemente grande, a matriz Hessiana em relação a x*

$$\nabla_x^2 \mathcal{L}(x^*, \lambda^*, c) = \nabla_x^2 L_h(x^*, \lambda^*) + cJ_h(x^*)^T J_h(x^*),$$

da função Lagrangiana aumentada é definida positiva. Em particular, o ponto estacionário x^* do problema

$$\begin{aligned} &\textit{minimizar} && \mathcal{L}(x, \lambda^*, c) \\ &\textit{sujeito a} && x \in \mathbb{R}^n, \end{aligned}$$

satisfaz a condição suficiente de segunda ordem para problemas irrestritos.

A construção do método de Lagrangiano aumentado também pode ser justificada através do estudo de um *método de penalidade deslocada* (veja [26, Capítulo 3]).

De forma semelhante ao método de penalidade quadrática, na k-ésima iteração do método de Lagrangiano aumentado, é fixado um parâmetro de penalidade $c_k > 0$. Além disso, é fixado também um vetor $\lambda^k \in \mathbb{R}^p$ como uma estimativa do multiplicador de Lagrange. Esta estimativa é motivada pela comparação da condição em (2.24) com a equação

$$\nabla_x \mathcal{L}(x, \lambda, c) = \nabla f(x) + J_h(x)^T (\lambda + ch(x)) = 0.$$

Apresentamos um algoritmo básico, adaptado de [8, Algoritmo 10.3.3].

Algoritmo 5 Lagrangiano aumentado clássico

Entrada: $c_0 > 0$, $\lambda^0 \in \mathbb{R}^p$, $k = 0$ e $x^0 \in \mathbb{R}^n$.

enquanto x^k não é estacionário para o problema (2.3) **faça**

1. Usando x^k como ponto inicial, resolva (possivelmente de forma aproximada) o subproblema irrestrito

$$\begin{aligned} &\text{minimizar} && \mathcal{L}(x, \lambda^k, c_k) \\ &\text{sujeito a} && x \in \mathbb{R}^n \end{aligned} \tag{2.26}$$

e obtenha x^{k+1}, para $\mathcal{L}$, a função Lagrangiana aumentada definida em (2.23).

2. Atualize o multiplicador

$$\lambda^{k+1} = \lambda^k + c_k h(x^{k+1}).$$

3. Se

$$\|h(x^{k+1})\| \leq 0.1 \|h(x^k)\|$$

atualize o parâmetro de penalidade fazendo $c_{k+1} = 10c_k$, caso contrário, faça $c_{k+1} = c_k$.

4. Faça $k = k + 1$.

fim

O teorema de convergência do método de Lagrangiano aumentado clássico é enunciado em seguida e é apresentado em [7, Teorema 4.4.6]. A demonstração é omitida neste trabalho e pode ser encontrada de forma detalhada em [4, Proposição 2.4].

Teorema 2.10. *Considere o problema com restrições de igualdade (2.3), com f e h duas vezes continuamente diferenciáveis em uma vizinhança do ponto $x^* \in \mathbb{R}^n$. Suponha que x^* é estacionário com $\lambda^* \in \mathbb{R}^p$ multiplicador de Lagrange associado, e que x^* satisfaz as condições LICQ (2.9) e a condição suficiente de segunda ordem (2.19). Então existem uma vizinhança U de x^* e os escalares $\bar{c} \geq 0$, $\delta > 0$ e $M > 0$ tais que*

(a) Para todo par $(\lambda, c) \in V(\bar{c}, \delta)$, em que

$$V(\bar{c}, \delta) = \{(\lambda, c) \in \mathrm{R}^p \times \mathbb{R} \mid \|\lambda - \lambda^*\| < \delta c \text{ e } c > \bar{c}\},$$

o problema

$$\begin{aligned} &\textit{minimizar} \quad \mathcal{L}(x, \lambda, c) \\ &\textit{sujeito a} \quad x \in \mathbb{R}^n, \end{aligned}$$

tem em U um único ponto estacionário, que denotamos por $x(\lambda, c)$, satisfazendo

$$\|x(\lambda, c) - x^*\| \leq M \frac{\|\lambda - \lambda^*\|}{c} \tag{2.27}$$

e

$$\|\lambda + ch\,(x(\lambda, c)) - \lambda^*\| \leq M \frac{\|\lambda - \lambda^*\|}{c}. \tag{2.28}$$

(b) Existe $\bar{\delta} \in (0, \delta]$ tal que para qualquer $\lambda^0 \in \mathbb{R}^p$ e qualquer sequência $(c_k) \in \mathbb{R}_+$ não decrescente, satisfazendo

$$(\lambda^0, c_0) \in V(\bar{c}, \bar{\delta}),$$

a sequência definida por

$$\lambda^{k+1} = \lambda^k + c_k h\left(x(\lambda^k, c_k)\right),$$

com $k \geq 0$, é tal que

$$(\lambda^k, c_k) \in V(\bar{c}, \bar{\delta}), \tag{2.29}$$

para todo k. Além disso, valem as convergências $\lambda^k \to \lambda^*$ e $x(\lambda^k, c_k) \to x^*$. Para a sequência (λ^k) a convergência é linear e, se $c_k \to \infty$, a convergência é superlinear.

Como observado em [7], a dificuldade da demonstração do teorema anterior é provar a existência do valor $\bar{\delta}$ tal que (2.29) seja satisfeito. No entanto, provado este fato, tem-se como consequência das relações (2.27) e (2.28), que

$$\|x(\lambda^k, c_k) - x^*\| \leq M \frac{\|\lambda^k - \lambda^*\|}{\bar{c}}$$

e

$$\|\lambda^{k+1} - \lambda^*\| \leq M \frac{\|\lambda^k - \lambda^*\|}{\bar{c}}.$$

Estas desigualdades são fundamentais para a prova de convergência. Também é observado em [7] que, se for possível obter tais valores de λ^0 e c_0, quando x^k é escolhido como ponto estacionário de problema (2.26) que está mais próximo de x^*, então o par (x^k, λ^k) gerado pelo Algoritmo 5 está bem definido e converge para (x^*, λ^*), com a taxa de convergência mencionada anteriormente. Estas e outras observações também são feitas em [4].

2.2.3 Tratamento das restrições de caixa

Para finalizar este capítulo, descrevemos o método de Lagrangiano aumentado para o problema de nosso interesse (2.2), o qual possui restrições mistas: de igualdade e de caixa. As restrições de igualdade são tratadas da mesma forma que no método de Lagrangiano aumentado clássico, isto é, por meio da penalização quadrática, e as restrições de caixa são mantidas a cada iteração. Isto significa que, para $k \geq 0$, dados $\lambda^k \in \mathbb{R}^p$ e um parâmetro de penalidade $c_k > 0$, resolvemos o subproblema

$$\begin{aligned} &\text{minimizar} \quad \mathcal{L}(x, \lambda^k, c_k) \\ &\text{sujeito a} \quad \ell \leq x \leq u, \end{aligned} \tag{2.30}$$

sendo $\mathcal{L}$ a função Lagrangiana aumentada definida em (2.23). Em seguida, as atualizações do multiplicador de Lagrange e do parâmetro de penalidade podem ser feitas de forma semelhante ao método clássico.

Sob determinadas hipóteses, no caso em que o subproblema (2.30) é resolvido de forma exata, o método de Lagrangiano aumentado é globalmente convergente para uma solução do problema original (2.2). Este resultado é enunciado no próximo teorema.

Teorema 2.11. [4, Proposição 2.1] *Considere o problema (2.2) com f e h continuamente diferenciáveis. Suponha que, para $k \geq 0$, x^{k+1} é um minimizador global do subproblema (2.30). Se $(\lambda^k) \subset \mathbb{R}^p$ é limitada e $(c_k) \subset \mathbb{R}_+$ é crescente, com $c_k \to \infty$, então todo ponto de acumulação de $(x^k)_{k>0}$ é um minimizador global do problema original (2.2).*

Na prática, o subproblema (2.30) é resolvido de forma aproximada. Note que, este problema possui o formato (1.1), estudado no capítulo anterior deste trabalho. Sendo assim, uma estratégia para a resolução dos subproblemas do método de Lagrangiano aumentado com restrições de caixa, é utilizar o Algoritmo 3. Além disso, recorde que, se x^* é um ponto estacionário do problema original (2.2), as condições de KKT (2.13) são satisfeitas. Então estas condições motivam o critério de parada para x^k

$$\|x^k - P_\Omega\big(x^k - \nabla_x L_h(x^k, \lambda^k)\big)\| \leq \varepsilon_g \tag{2.31}$$

e

$$\|h(x^k)\| \leq \varepsilon_h, \tag{2.32}$$

em que $\varepsilon_g > 0$ e $\varepsilon_h > 0$ são tolerâncias fixadas. Escrevemos o algoritmo do método de Lagrangiano aumentado, adaptado de [1, Algoritmo 1] e [3, Algoritmo 17.4].

Algoritmo 6 Lagrangiano aumentado com caixa

Entrada: $k = 0$, $c_0 > 0$, $\lambda^0 \in \mathbb{R}^p$, $\eta_0 > 0$, $x^0 \in \Omega$, com Ω a caixa (1.2), e as constantes $\varepsilon_g > 0$ e $\varepsilon_h > 0$.

enquanto x^k não satisfaz (2.31) e (2.32) **faça**

1. Resolva o subproblema (2.30) (possivelmente de forma aproximada) utilizando o Algoritmo 3 com x^k como ponto inicial e obtenha x^{k+1}.
2. Se $\|h(x^{k+1})\| \leq \eta_k$ atualize o multiplicador

$$\lambda^{k+1} = \lambda^k + c_k h(x^{k+1}),$$

o parâmetro de penalidade

$$c_{k+1} = c_k$$

e a tolerância na viabilidade

$$\eta_{k+1} = \eta_k / (c_{k+1})^{0.9}.$$

Caso contrário, faça

$$\lambda^{k+1} = \lambda^k, \quad c_{k+1} = 10 c_k \quad \text{e} \quad \eta_{k+1} = 0.1/(c_{k+1})^{0.1}.$$

3. Faça $k = k + 1$.

fim

Observamos que a atualização da aproximação λ^{k+1} do multiplicador de Lagrange pode ser feita para toda iteração do método de Lagrangiano aumentado, como é feito em [27, Algoritmo 4.1]. Da mesma forma que o Teorema 2.10, a teoria de convergência do método de Lagrangiano aumentado para o problema com restrições de igualdade e de caixa também utiliza como hipótese uma condição de qualificação das restrições. Na referência [24], é estudada a condição de *dependência linear positiva constante (CPLD)* e é feito um estudo da teoria de convergência. Uma outra abordagem é feita em [28]. O próximo capítulo deste trabalho é dedicado aos detalhes computacionais e os testes numéricos dos Algoritmos 3 e 6.

Capítulo 3

Experimentos numéricos

Neste capítulo, apresentamos os detalhes da implementação computacional do Algoritmo 6 e os testes numéricos realizados. Todas os algoritmos implementados foram desenvolvidos em linguagem de programação Julia, com base na infraestrutura JuliaSmoothOptimizers [17] para otimização não linear. Denominamos nossa implementação computacional de Percival e a disponibilizamos em [18]. Este capítulo está dividido em três seções. Na Seção 3.1 detalhamos alguns procedimentos computacionais dos algoritmos implementados. Na Seção 3.2, vemos como o subproblema do método de Lagrangiano aumentado pode ser resolvido. Na Seção 3.3, apresentamos os resultados numéricos.

3.1 Detalhes de implementação

Nesta seção apresentamos alguns detalhes da implementação computacional do Algoritmo 6 que foram utilizados na construção do algoritmo Percival para a resolução de problemas da forma

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && h(x) = 0, \\ & && x \in \Omega, \end{aligned} \tag{3.1}$$

supondo $f : \mathbb{R}^n \to \mathbb{R}$ e $h : \mathbb{R}^n \to \mathbb{R}^p$ funções diferenciáveis e Ω a caixa (2.1).

Consideramos a tolerância ε_g exigida no critério de estacionariedade (2.31), como sendo

$$\varepsilon_g = \varepsilon_g^{(a)} + \varepsilon_g^{(r)} \|x^0 - P_\Omega\big(x^0 - \nabla_x L_h(x^0, \lambda^0)\big)\|,$$

em que $\varepsilon_g^{(a)} > 0$ e $\varepsilon_g^{(r)} > 0$ são dados e chamados de tolerância absoluta e relativa, respectivamente, L_h é a função definida em (2.12), x^0 é um ponto inicial fornecido para o algoritmo e λ_0 é o multiplicador de Lagrange inicial. Este multiplicador é calculado internamente na implementação do algoritmo, através da solução de quadrados mínimos do problema

$$\begin{aligned} &\text{minimizar} && \frac{1}{2}\|J_h(x^0)^T\lambda + \nabla f(x^0)\|^2 \\ &\text{sujeito a} && \lambda \in \mathbb{R}^p, \end{aligned} \tag{3.2}$$

em que J_h é a matriz Jacobiana em relação às restrições de igualdade do problema (3.1).

O subproblema (2.30) do método de Lagrangiano aumentado pode ser modelado computacionalmente para uma iteração $k \geq 0$ do algoritmo, através da aplicação de

uma estrutura que denominamos de AugLagModel. A implementação desta estrutura também está disponível em [18]. Nesta implementação, consideramos como dados de entrada: o problema (3.1), o iterando x^k, a aproximação do multiplicador de Lagrange λ^k, o parâmetro de penalidade c_k e h^k como sendo o valor da restrição h em x^k. Então, um modelo computacional que armazena os valores de x^k, h^k, λ^k, c_k e $\lambda^k + c_k h^k$ é gerado e funções que possibilitam suas atualizações para $k \geq 1$ são construídas.

Na implementação dos algoritmos, a função Lagrangiana aumentada $\mathcal{L}(x, \lambda^k, c_k)$, o gradiente $\nabla_x \mathcal{L}(x, \lambda^k, c_k)$ e o produto $\nabla_x^2 \mathcal{L}(x, \lambda^k, c_k)w$ podem ser atualizados de forma otimizada para quaisquer $x \in \mathbb{R}^n$, $w \in \mathbb{R}^n$ e $k \geq 0$. Perceba que, como

$$\nabla_x \mathcal{L}(x, \lambda^k, c_k) = \nabla f(x) + J_h(x)^T \left(\lambda^k + c_k h(x)\right)$$

e

$$\nabla_x^2 \mathcal{L}(x, \lambda^k, c_k)w = \nabla^2 f(x)w + \sum_{i=1}^{p} \left(\lambda_i^k + c_k h_i(x)\right) \nabla^2 h_i(x)w + c_k J_h(x)^T J_h(x)w,$$

o vetor $\lambda^k + c_k h(x)$ pode ser calculado previamente e o produto $J_h(x)^T J_h(x)$ pode ser evitado, o que conduz uma economia nos cálculos.

Observamos ainda que, a estrutura AugLagModel é criada somente na primeira iteração da implementação do Algoritmo 6, isto é, para $k = 0$, uma vez que, as funções de atualização são geradas nesta primeira chamada. Isto também possibilita uma economia de cálculos e de memória, acelerando o trabalho computacional da aplicação do Algoritmo 6.

3.2 Resolução do subproblema do método de Lagrangiano aumentado

Cada iteração do Algoritmo 6 busca uma solução aproximada para o subproblema (2.30). Utilizamos como subrotina para resolver este subproblema a função TRON, disponível no pacote JSOSolvers.jl [1]. A implementação computacional desta função baseia-se na teoria desenvolvida em [13], a qual estudamos no Capítulo 1 desde trabalho. Nesta seção, vemos como o Algoritmo 3 de região de confiança executa os Passos 2 e 3.

3.2.1 Cálculo do passo de Cauchy

Recorde que, no Algoritmo 3, o cálculo do passo de Cauchy $d^k(\alpha_k)$ depende do valor obtido para $\alpha_k > 0$. Na função TRON, um esquema iterativo de busca é feito para obter o valor de α_k satisfazendo as condições

$$m_k\left(d^k(\alpha_k)\right) \leq \mu_0 \nabla f(x^k)^T d^k(\alpha_k) \quad \text{e} \quad \|d^k(\alpha_k)\| \leq \mu_2 \Delta_k, \tag{3.3}$$

estudadas no Capítulo 1. Este esquema é descrito em [13] sendo que, para cada $k \geq 0$, uma sequência de tentativas

$$\left(\alpha_k^j\right)_{j \in \mathbb{N}}$$

é gerada. Dado $\sigma > 0$, esta sequência obedece a seguinte regra:

[1] https://github.com/JuliaSmoothOptimizers/JSOSolvers.jl

1. Se α_k^0 falhar para a condição (3.3) então diminua o valor da tentativa de comprimento do passo, isto é, faça
$$\alpha_k^{(j+1)} = \frac{1}{\sigma}\alpha_k^j,$$
para todo $j \geq 0$, até que esta condição seja satisfeita. Daí, faça α_k como sendo o último valor de tentativa.

2. Se α_k^0 satisfazer a condição (3.3) então aumente o valor da tentativa de comprimento do passo, isto é, faça
$$\alpha_k^{(j+1)} = \sigma\alpha_k^j,$$
para todo $j \geq 0$, até que esta condição falhe. Daí, faça α_k como sendo o último valor de tentativa.

Para $k = 0$, o primeiro valor de comprimento do passo é $\alpha_0^0 = 1$ e para os demais valores de k é definido $\alpha_{k+1}^0 = \alpha_k$. Outros detalhes sobre este esquema, inclusive os que garantem a limitação superior (1.61), estão descritos em [13]. Para a próxima seção denotamos o passo de Cauchy por d_c^k.

3.2.2 Cálculo da solução aproximada para o subproblema

Descrevemos agora os procedimentos sugeridos em [13] para o cálculo da direção d^k após a obtenção do passo de Cauchy, isto é, os procedimentos práticos que podem ser realizados no Passo 3 do Algoritmo 3. Primeiramente, definimos, para cada $k \geq 0$, o iterando obtido pelo passo de Cauchy, isto é,
$$x_c^k = x^k + d_c^k.$$
Em [13], pede-se que a direção d^k satisfaça as condições em (1.60) do Algoritmo 3 e a condição
$$A\left(x_c^k\right) \subset A\left(x^k + d^k\right), \tag{3.4}$$
sendo A o conjunto ativo em relação às restrições de caixa. Recorde que, a inclusão acima é motivada pelo Teorema 1.36. A ideia para satisfazer estas condições é realizar um esquema iterativo no qual uma sequência de pontos
$$\left(x^{(k,j)}\right)_{j\in\mathbb{N}} \subset \Omega$$
é gerada. Esta sequência de pontos satisfaz as condições
$$A\left(x_c^k\right) \subset A\left(x^{(k,j)}\right) \quad \text{e} \quad \|x^{(k,j)} - x^k\| \leq \mu_2\Delta_k \tag{3.5}$$
através do seguinte procedimento:

1. Faça $x^{(k,0)} = x_c^k$.

2. Para $j \geq 1$, encontre uma direção de descida $w^{(k,j-1)}$ resolvendo o problema
$$\begin{aligned} &\text{minimizar} && q_k\left(x^{(k,j-1)} + w\right) \\ &\text{sujeito a} && w_i = 0,\ i \in A(x^{(k,j-1)}) \\ &&& \|w\| \leq \mu_2\Delta_k, \end{aligned} \tag{3.6}$$

e obtenha $x^{(k,j)}$ fazendo

$$x^{(k,j)} = P_\Omega\left(x^{(k,j-1)} + \beta w^{(k,j-1)}\right), \tag{3.7}$$

em que $\beta > 0$ e q_k é o modelo quadrático (1.28).

Este processo termina quando $d^k = x^{(k,j)} - x^k$ satisfaz as condições em (1.60) e (3.4), para algum $j \in \mathbb{N}$.

Na função TRON, o problema (3.6) é resolvido da seguinte maneira. Primeiramente, é construído um operador linear de projeção $Z \in \mathbb{R}^{n\times m}$ no subespaço

$$\{w \in \mathbb{R}^n \;\mid\; w_i = 0,\; i \in A(x^{(k,j-1)})\},$$

cujo número m de colunas é a cardinalidade de $A(x^{(k,j-1)})$. A construção deste operador linear é feita utilizando o pacote LinearOperators.jl [29]. Em seguida, é resolvido o problema

$$\begin{aligned} &\text{minimizar} && q_k\left(x^{(k,j-1)} + Zv\right) \\ &\text{sujeito a} && \|Zv\| \leq \mu_2 \Delta_k, \end{aligned} \tag{3.8}$$

por meio de uma variação do método de Steihaug para minimização de funções quadráticas com região de confiança, descrito em [30] e implementado no pacote Krylov.jl [31]. Este método é livre de fatoração matricial. O valor de β em (3.7) é obtido por meio de uma busca linear projetada, tal que uma condição de decréscimo na função q_k seja satisfeita. Mais detalhes sobre esta busca e esta condição de decréscimo podem ser encontrados em [13, Seção 6].

3.3 Resultados numéricos

Os resultados numéricos para o algoritmo Percival foram obtidos para uma coleção de problemas da biblioteca CUTEst [19], coletados por meio do pacote CUTEst.jl [32], e foram comparados com o algoritmo para problemas de otimização com restrições IPOPT. O desenvolvimento teórico deste algoritmo pode ser encontrado em [20] e a implementação computacional está disponibilizada em https://github.com/coin-or/Ipopt. Para acessá-lo e integrá-lo com a linguagem de programação Julia, utilizamos o pacote NLPModelsIpopt.jl [33]. Consideramos todos os 384 problemas da coleção da forma

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && h(x) = 0, \\ &&& g(x) \leq 0, \\ &&& \ell \leq x \leq u, \end{aligned} \tag{3.9}$$

em que $f : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}^p$ e $g : \mathbb{R}^n \to \mathbb{R}^m$ são funções diferenciáveis com $1 \leq n \leq 1000$ e $1 \leq m + p \leq 1000$, o vetor $\ell \in (\mathbb{R} \cup \{-\infty\})^n$ e o vetor $u \in (\mathbb{R} \cup \{+\infty\})^n$. Quando não há restrições de caixa, utilizamos o método de Lagrangiano aumentado clássico. As restrições de desigualdade foram transformadas em igualdades por meio da adição de variáveis auxiliares de folga.

Para o algoritmo Percival, consideramos os parâmetros de entrada: `max_iter` = 1000, que é o número máximo de iterações; `max_time` = 60, que é tempo máximo em

segundos; `max_eval` $= 10^5$, que é o número máximo de avaliações da função objetivo; $\varepsilon_g^{(a)} = 10^{-6}$; $\varepsilon_g^{(r)} = 10^{-6}$; $\varepsilon_h = 10^{-6}$ e $c_0 = 10$. Além disso, para cada iteração $k \geq 0$ da resolução do subproblema do método de Lagrangiano aumentado, utilizamos a matriz B_k do modelo quadrático (1.37) como sendo a matriz Hessiana associada ao subproblema.

Já no algoritmo IPOPT, consideramos os parâmetros de entrada: `max_cpu_time` $=$ 60, que é o tempo máximo em segundos; `dual_inf_tol` $= 10^{-6}$, que é a tolerância absoluta para a estacionariedade; `tol` $= 10^{-6}$, que é a tolerância relativa para a estacionariedade; `constr_viol_tol` $= 10^{-6}$, que é a tolerância para a viabilidade; `acceptable_iter` $= 0$, que desativa a heurística de parada antecipada que verifica se o algoritmo parece estar preso e `nlp_scaling_method = 'none'`, que significa que não é considerada qualquer mudança de escala para o problema. Todas as informações sobre estes e outros parâmetros de entrada estão em https://projects.coin-or.org/Ipopt.

Após executarmos os dois algoritmos para resolver cada um dos 384 problemas, consideramos para a análise dos resultados, um critério comparativo entre suas soluções. Podemos nos referir aos algoritmos Percival e IPOPT pelos índices $i = 1$ e $i = 2$, respectivamente. Dado um problema da forma (3.1), consideramos que o algoritmo i encontrou uma solução para este problema se duas condições são satisfeitas. A primeira verifica se a viabilidade do problema foi atingida, isto é, é verificado se o ponto obtido como solução do problema é viável com tolerância 10^{-6} para as restrições de igualdade. A segunda condição, sugerida em [34], verifica se

$$\frac{f_i^* - f_{\min}}{\max\{1, |f_{\min}|\}} \leq 0.01, \tag{3.10}$$

em que f_i^* é o valor de função encontrado pelo algoritmo i e $f_{\min} = \min\{F_1^*, F_2^*\}$, sendo

$$F_i^* = \begin{cases} f_i^*, & \text{se a solução encontrada é viável,} \\ +\infty, & \text{caso contrário,} \end{cases}$$

para o algoritmo i.

Na análise de perfil de desempenho [35] dos algoritmos, avaliamos o tempo de execução na resolução dos problemas e o custo de avaliações de função objetivo e de gradiente utilizando uma medida de trabalho análoga à sugerida em [36]. Esta medida foi calculada como o valor de $2ng + nf$ para os algoritmos em cada problema, sendo que nf é o número de avaliações de função objetivo e ng é o número de avaliações de gradiente.

No Apêndice A, apresentamos no formato de tabela os resultados numéricos que foram utilizados na geração do perfil de desempenho. Listamos os problemas em blocos, em que cada bloco contém os resultados de cada algoritmo em uma linha. As colunas `nvar` e `ncon` representam os números de variáveis e restrições, respectivamente. A coluna `status` indica a *flag* de saída de cada algoritmo e mostra se o mesmo encontrou uma solução aproximada para o problema ou não. Isto significa que, caso a solução satisfaça o critério de estacionariedade e viabilidade requeridos, então `status = 'first_order'`. Se o número de iterações do algoritmo exceder o valor máximo ou o tempo de execução do algoritmo exceder o tempo máximo ou, ainda, o número de avaliações de função objetivo exceder o número máximo permitido, então `status = 'max_iter'`, `status = 'max_time'` ou `status = 'max_eval'`, respectivamente. Para o algoritmo Percival, se o parâmetro de penalidade c_k torna-se muito grande e se

$$\|J_h(x^k)^T h(x^k)\| < \sqrt{\varepsilon_h}\|h(x^k)\|, \tag{3.11}$$

fazemos `status = 'infeasible'`. A expressão (3.11) quer dizer que encontramos um ponto estacionário para a medida de viabilidade das restrições de igualdade com uma tolerância relativa e, nesse caso, como estudado em [37] e [38], x^k é chamado de *ponto estacionário inviável*, o que pode indicar que o problema original é possivelmente inviável. Além disso, no caso em que o parâmetro de penalidade c_k tornou-se muito grande, fazemos `status = 'stalled'`. Para outros tipos de erro, como, por exemplo, erros numéricos, `status` recebe o valor `'exception'` ou `'unknown'`. O valor da função objetivo na solução obtida é exibido na coluna `objective`. A coluna `primal_feas` mostra a norma da medida de viabilidade das restrições na solução obtida. O valor de $2ng + nf$ é exibido na coluna `evals`. Se a viabilidade e o critério (3.10) são satisfeitos, os resultados são destacados em negrito.

Os perfis de desempenho foram gerados utilizando o pacote SolverBenchmark.jl [39] e as rotinas para a execução de todos os testes estão disponíveis em [40]. Apresentamos os resultados nas figuras deste capítulo. Consideramos duas metodologias para verificar que o problema foi resolvido. Primeiramente, geramos os perfis de desempenho considerando o critério de sucesso (viabilidade e (3.10)) e comparamos os algoritmos em tempo de execução e em avaliações de função (gráficos superiores esquerdo e direito, respectivamente). Em seguida, consideramos que o problema foi resolvido se a flag de saída `status` do algoritmo é igual a `'first_order'` e comparamos os algoritmos em tempo de execução e em avaliações de função (gráficos inferiores esquerdo e direito, respectivamente). Utilizamos um critério de empate para comparações dos métodos, sugerido em [41]. Consideramos que um algoritmo é o mais eficiente se seu tempo de execução não ultrapassar em 5% o tempo do mais rápido.

Na Figura 3.1 exibimos os resultados obtidos para todos os 384 problemas considerados. No gráfico superior esquerdo, vemos que o algoritmo Percival resolve 74% dos problemas, já o algoritmo IPOPT resolve 92% dos problemas. Além disso, o algoritmo Percival é mais rápido em 33% dos problemas e o algoritmo IPOPT em 63% dos problemas. Quando consideramos que um problema é resolvido utilizando a flag de saída dos algoritmos, vemos no gráfico inferior esquerdo que o algoritmo Percival também resolve 74% dos problemas e o algoritmo IPOPT, 92% dos problemas, sendo que o algoritmo Percival é mais rápido em 34% dos problemas e o algoritmo IPOPT em 60% dos problemas. Quando consideramos novamente o critério de sucesso, a comparação de avaliações de função no gráfico superior direito mostra que o algoritmo Percival faz menos avaliações de função somente em 7% dos problemas, já o algoritmo IPOPT faz menos avaliações de função em 90% dos problemas. Utilizando a flag de saída, vemos que estes valores são 6% e 89%, respectivamente. Sendo assim, podemos concluir que o algoritmo IPOPT é superior ao algoritmo Percival tanto na eficiência, isto é, é mais rápido e faz menos avaliações de função, quanto na robustez, ou seja, resolve um número maior de problemas.

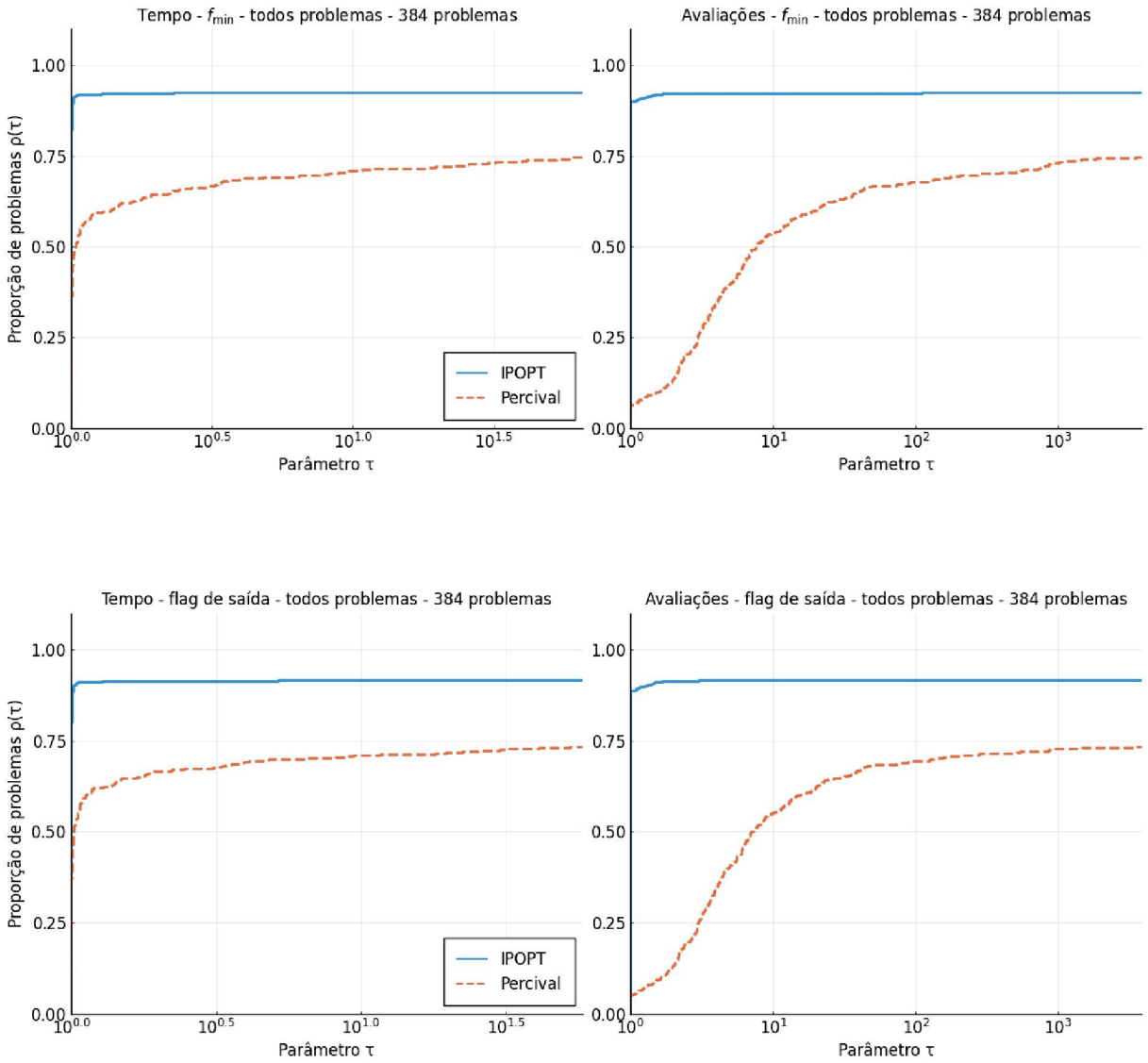

Figura 3.1: Resultados obtidos para todos os problemas.

Os resultados para os 51 problemas com somente restrições de igualdade são mostrados na Figura 3.2. Nesse caso, o algoritmo Percival resolve 90% e 94% dos problemas, considerando o critério de sucesso e a flag de saída do algoritmo (gráficos superiores e inferiores), respectivamente. Já o algoritmo IPOPT resolve 86% e 90% destes problemas, como mostram os gráficos superiores e inferiores, respectivamente. Em relação ao tempo de execução dos algoritmos, os gráficos superior esquerdo e inferior esquerdo indicam que ambos os algoritmos possuem um desempenho semelhante. No entanto, notamos nos gráficos superior direito e inferior direito que o algoritmo IPOPT é mais eficiente em termos de avaliações de função. Sendo assim, podemos concluir que para problemas com somente restrições de igualdade, o algoritmo Percival resolve uma quantidade um pouco maior de problemas, no entanto, no custo de avaliações de função, o algoritmo IPOPT é mais eficiente.

Os problemas com somente restrições de desigualdade somam 63 problemas. Com

base nos gráficos da Figura 3.3, podemos concluir que o algoritmo IPOPT é mais robusto e eficiente em qualquer caso. O mesmo ocorre para os 73 problemas com restrições de igualdade e caixa e para os 78 problemas com restrições de desigualdade e caixa, como mostram as Figuras 3.4 e 3.5, respectivamente.

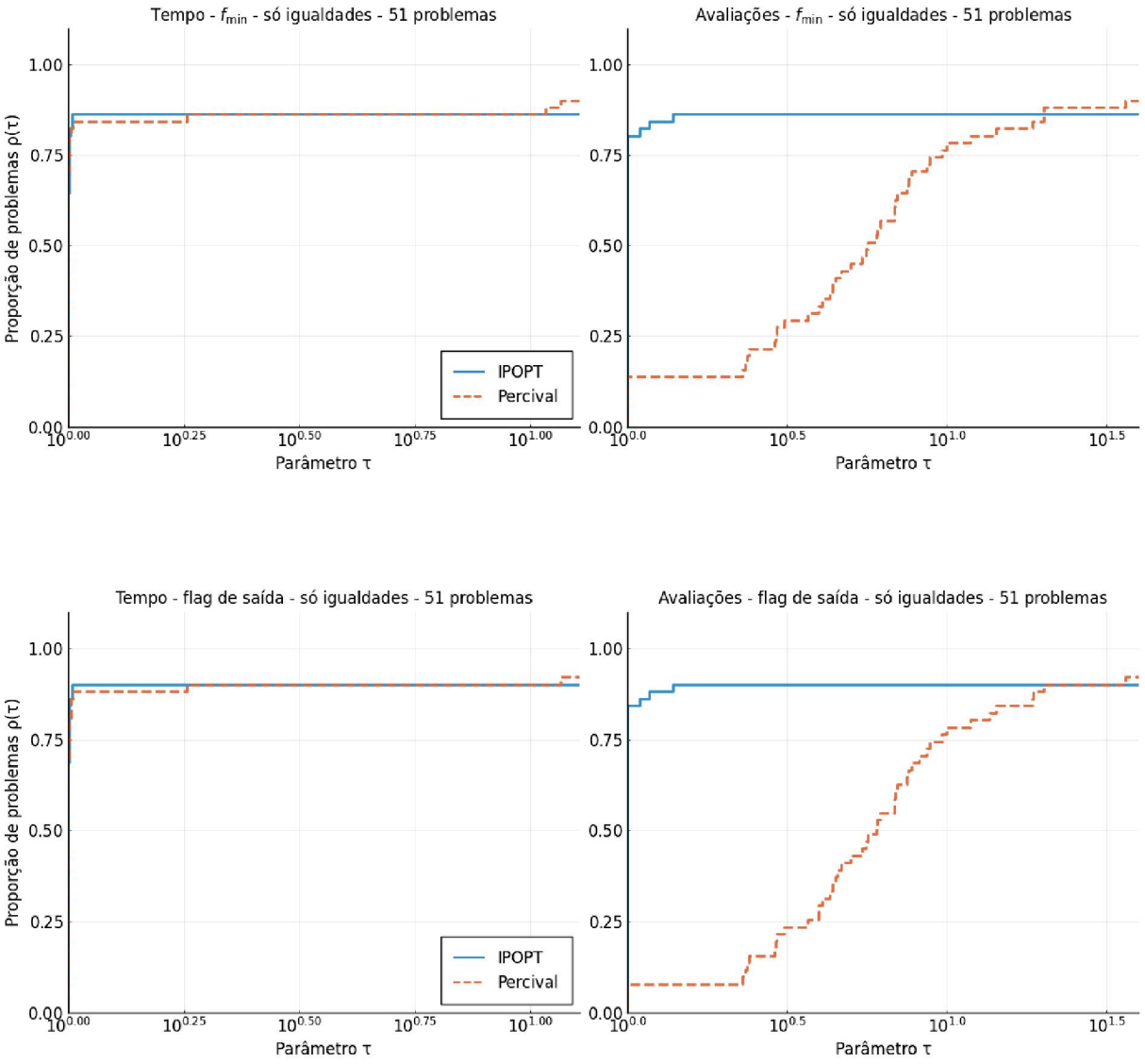

Figura 3.2: Resultados obtidos para problemas com restrições de igualdade.

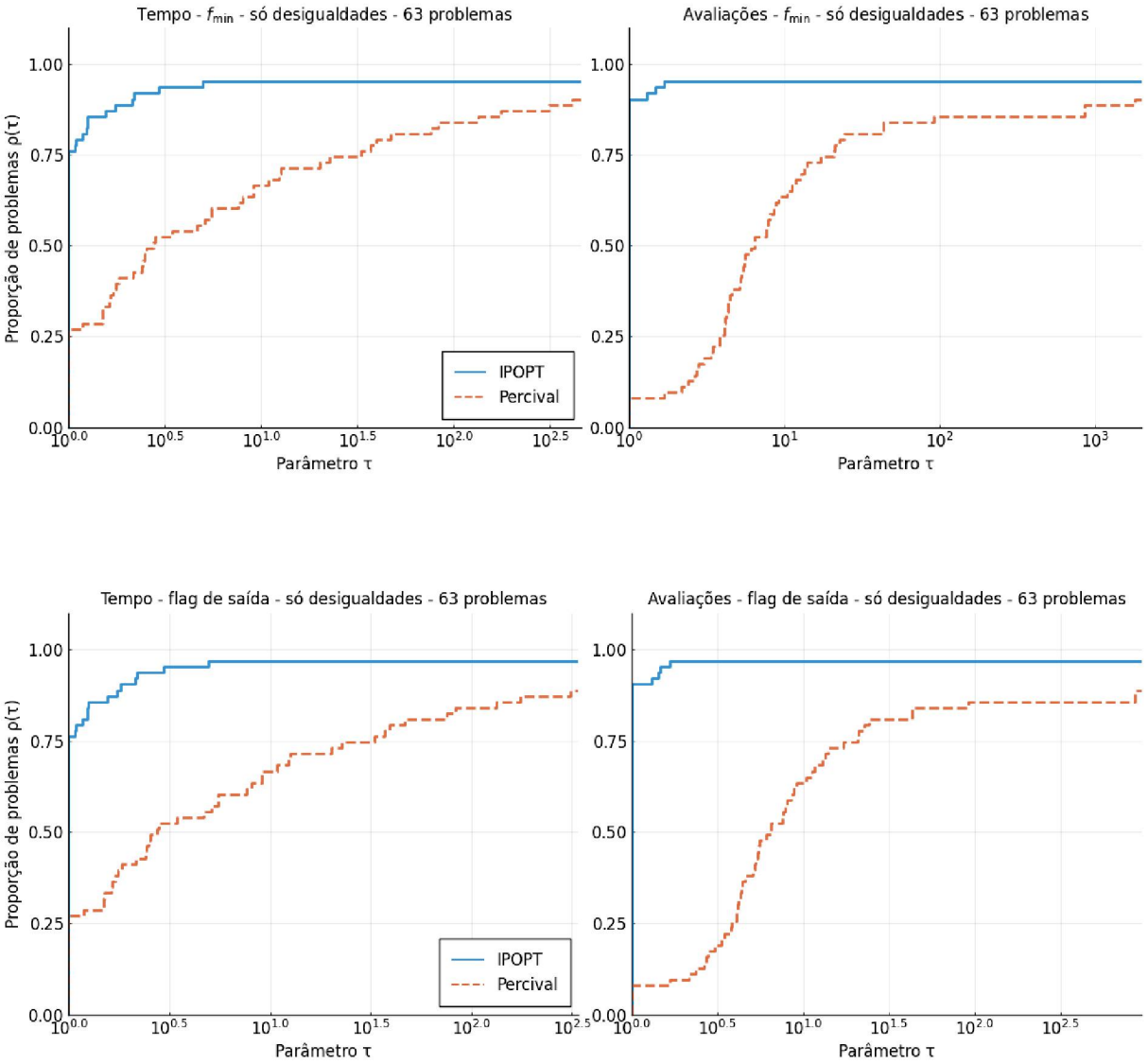

Figura 3.3: Resultados obtidos para problemas com restrições de desigualdade.

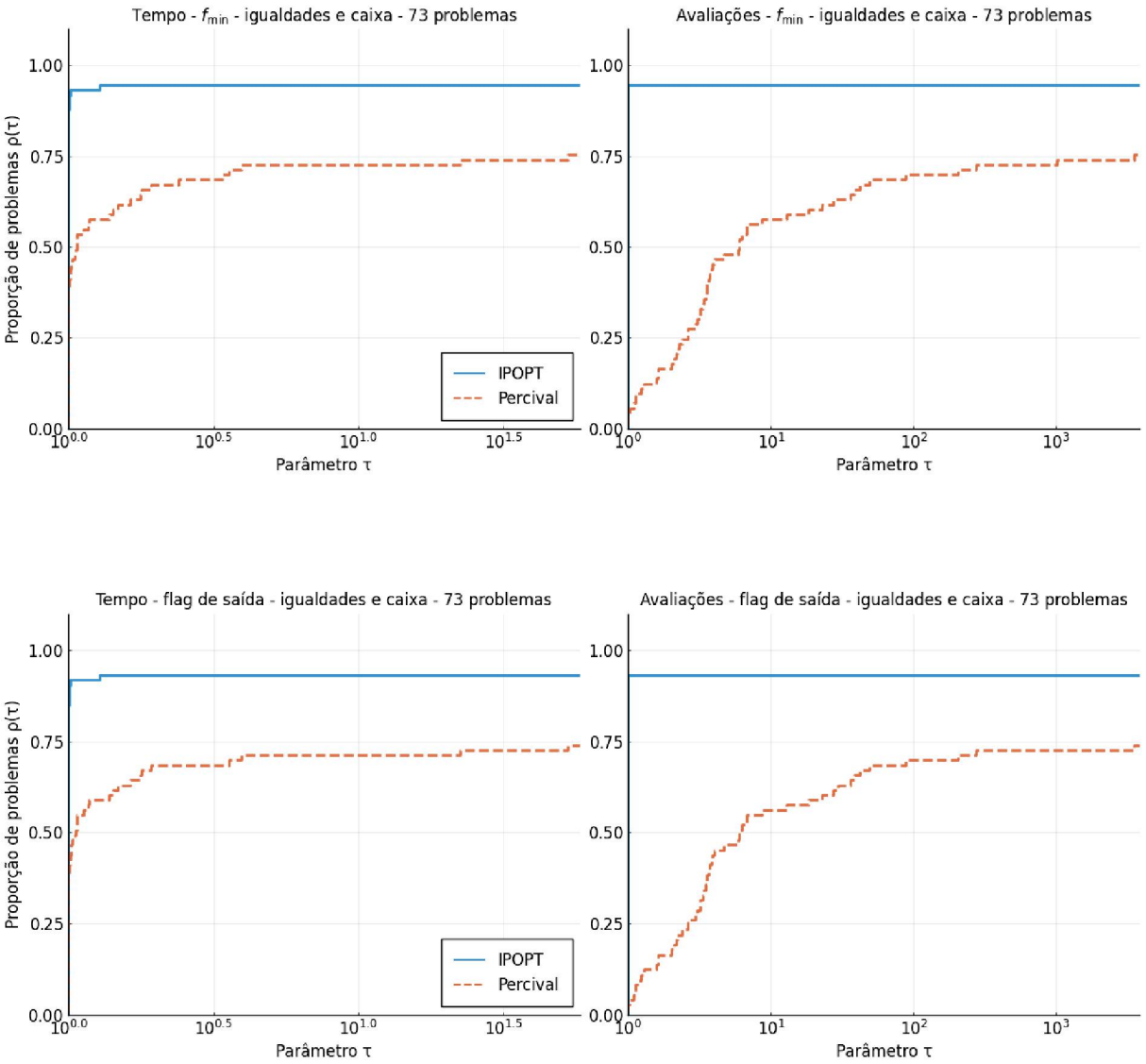

Figura 3.4: Resultados obtidos para problemas com restrições de igualdade e caixa.

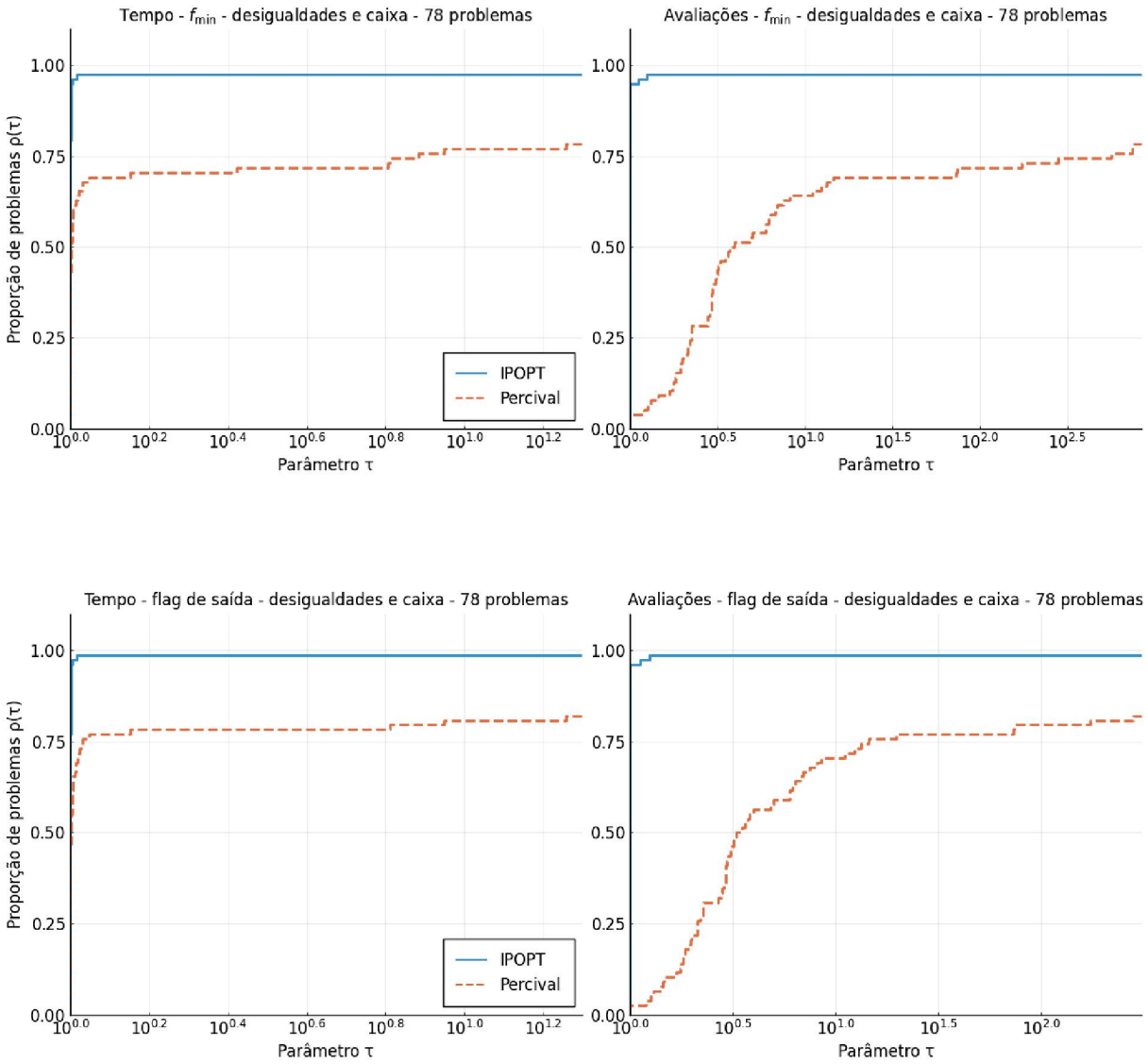

Figura 3.5: Resultados obtidos para problemas com restrições de desigualdade e caixa.

Com base nos resultados numéricos obtidos neste trabalho, notamos que o parâmetro de penalidade c_k para o algoritmo Percival torna-se muito grande em alguns problemas. Além disso, em muitos problemas o tempo de execução, o número de iterações ou o número de avaliações de função excedem os valores permitidos. Sendo assim, o desenvolvimento de novas metodologias para a atualização do parâmetro de penalidade e o aperfeiçoamento da implementação do algoritmo Percival podem fazer com que melhores resultados sejam alcançados. Para os problemas com restrições de desigualdade, podemos implementar estratégias estudadas em [1] para a atualização de variáveis de folga. Em [20], o escalamento da função objetivo e das restrições do problema é sugerido. Portanto, a implementação destas estratégias pode melhorar o desempenho do algoritmo Percival.

Conclusão

Neste trabalho estudamos e implementamos um método de Lagrangiano aumentado para resolver problemas de minimização com restrições de igualdade e caixa. Vimos que, a cada iteração do método, um subproblema de minimização com somente as restrições de caixa é considerado. A busca por uma solução aproximada destes subproblemas foi estudada no Capítulo 1 com a utilização de um método de região de confiança, proposto por Lin e Moré em [13]. O desenvolvimento teórico da abordagem de Lin e Moré foi apresentada em que demonstramos um resultado de convergência global já estabelecido. Utilizando alguns resultados de [12], [14]–[16] e adaptando algumas condições a respeito do passo de Cauchy e da direção de busca na obtenção da solução dos subproblemas, conseguimos provar que qualquer ponto limite da sequência gerada pelo método de região de confiança é estacionário. Além desta teoria de convergência, estudamos quais propriedades geométricas do algoritmo de Lin e Moré motivam alguns procedimentos práticos que podem ser realizados computacionalmente.

O método de Lagrangiano aumentado foi apresentado no Capítulo 2. Implementamos o algoritmo em linguagem de programação Julia, com base no desenvolvimento prático proposto em [1] e utilizando a infraestrutura computacional JuliaSmoothOptimizers [17] para otimização não linear. Denominamos nossa implementação de Percival e a disponibilizamos em [18]. Realizamos testes numéricos do algoritmo em 384 problemas da coleção CUTEst [19] e comparamos os resultados com o algoritmo IPOPT [20].

Através das comparações apresentadas no Capítulo 3, concluímos que o desempenho do algoritmo IPOPT é superior ao desempenho do algoritmo Percival, tanto no tempo de execução quanto no número de avaliações de função objetivo e de gradiente se considerarmos todos os 384 problemas. Observamos que, para problemas com somente restrições de igualdade, o algoritmo Percival resolve uma quantidade maior de problemas. Observamos que, em ralação ao tempo de execução para os problemas com somente restrições de igualdade, os algoritmos obtiveram resultados muito próximos. No entanto, o algoritmo IPOPT é mais eficiente em avaliações de função. Para os problemas com restrições de desigualdade e caixa, igualdade e caixa e somente de desigualdade, concluímos que o algoritmo IPOPT é mais eficiente e mais robusto em qualquer caso.

Em muitos problemas o parâmetro de penalidade utilizado no método de Lagrangiano aumentado implementado torna-se muito grande, ocasionando a falha do algoritmo. Sendo assim, destacamos como trabalhos futuros, o aperfeiçoamento da implementação do algoritmo Percival e o estudo de novas abordagens para serem aplicadas na atualização do parâmetro de penalidade. Além disso, para os problemas com restrições de desigualdade, o algoritmo Percival pode ser melhorado através da implementação de estratégias estudadas em [1] para a atualização de variáveis de folga. Estratégias de escalamento da função objetivo e das restrições do problema são discutidas em [20] e também podem ser

aplicadas para melhorar o desempenho do algoritmo.

Além do aperfeiçoamento da implementação do algoritmo Percival, podemos estudar a resolução de problemas de quadrados mínimos não lineares com restrições de caixa. Nesse caso, a função objetivo nestes problemas é da forma $f : \mathbb{R}^n \to \mathbb{R}$, com

$$f(x) = \frac{1}{2}\|F(x)\|^2,$$

sendo $F : \mathbb{R}^n \to \mathbb{R}^m$. Uma aproximação linear para F pode ser feita, aproximando o problema original por um problema de quadrados mínimos linear com restrições de caixa. A adaptação do algoritmo de Lin e Moré para resolver este problema aproximado pode ser estudada. Sendo assim, destacamos como um possível trabalho futuro, a implementação computacional desta adaptação.

Apêndice A

Anexos

name	solver	nvar	ncon	status	objective	primal_feas	evals
A4X12	**IPOPT**	**130**	**385**	**first_order**	**6.81e−01**	**0.00e+00**	**609**
	Percival	130	385	max_time	−2.53e+01	4.33e−02	2149
ACOPP118	**IPOPT**	**344**	**608**	**first_order**	**1.30e+05**	**6.19e−12**	**57**
	Percival	344	608	max_time	2.11e+05	1.61e+05	2163
ACOPP14	**IPOPT**	**38**	**68**	**first_order**	**8.08e+03**	**1.21e−14**	**36**
	Percival	**38**	**68**	**first_order**	**8.08e+03**	**9.52e−08**	**52135**
ACOPP30	**IPOPT**	**72**	**142**	**first_order**	**5.77e+02**	**1.78e−14**	**66**
	Percival	72	142	max_time	5.77e+02	7.04e−04	5686
ACOPP57	**IPOPT**	**128**	**274**	**first_order**	**4.17e+04**	**2.04e−14**	**56**
	Percival	128	274	max_time	5.64e+04	1.08e+05	13531
ACOPR118	**IPOPT**	**344**	**726**	**first_order**	**1.30e+05**	**7.49e−13**	**365**
	Percival	344	726	max_time	1.32e+05	2.82e+04	4901
ACOPR14	**IPOPT**	**38**	**82**	**first_order**	**8.08e+03**	**1.02e−14**	**561**
	Percival	38	82	max_time	8.20e+03	1.23e+02	24373
ACOPR30	**IPOPT**	**72**	**172**	**first_order**	**5.77e+02**	**1.69e−14**	**67**
	Percival	72	172	max_time	5.77e+02	1.06e−04	5617
ACOPR57	**IPOPT**	**128**	**331**	**first_order**	**4.17e+04**	**2.44e−14**	**180**
	Percival	128	331	max_time	6.04e+04	1.53e+04	3030
AGG	**IPOPT**	**163**	**488**	**first_order**	**−3.60e+07**	**1.16e−10**	**578**
	Percival	163	488	max_time	−7.83e+05	9.53e+02	15121
AIRPORT	**IPOPT**	**84**	**42**	**first_order**	**4.80e+04**	**0.00e+00**	**48**
	Percival	**84**	**42**	**first_order**	**4.80e+04**	**1.85e−11**	**286**
ALJAZZAF	**IPOPT**	**1000**	**1**	**first_order**	**3.74e+04**	**6.71e−10**	**180**
	Percival	**1000**	**1**	**first_order**	**3.74e+04**	**2.08e−08**	**281**
ALLINITA	**IPOPT**	**4**	**4**	**first_order**	**3.33e+01**	**1.11e−16**	**73**
	Percival	**4**	**4**	**first_order**	**3.33e+01**	**6.35e−07**	**497**
ALLINITC	**IPOPT**	**4**	**1**	**first_order**	**3.05e+01**	**1.11e−16**	**81**
	Percival	**4**	**1**	**first_order**	**3.05e+01**	**8.24e−07**	**468**
ALSOTAME	**IPOPT**	**2**	**1**	**first_order**	**8.21e−02**	**0.00e+00**	**27**
	Percival	**2**	**1**	**first_order**	**8.21e−02**	**2.39e−08**	**47**
ANTWERP	**IPOPT**	**27**	**10**	**first_order**	**3.25e+03**	**1.10e−16**	**326**
	Percival	27	10	max_time	5.49e+05	3.58e−01	369686
AVGASA	**IPOPT**	**8**	**10**	**first_order**	**−4.63e+00**	**0.00e+00**	**33**
	Percival	**8**	**10**	**first_order**	**−4.63e+00**	**3.16e−07**	**56**
AVGASB	**IPOPT**	**8**	**10**	**first_order**	**−4.48e+00**	**0.00e+00**	**36**
	Percival	**8**	**10**	**first_order**	**−4.48e+00**	**1.07e−08**	**116**
AVION2	**IPOPT**	**49**	**15**	**first_order**	**9.47e+07**	**1.36e−12**	**179**
	Percival	49	15	max_time	9.47e+07	5.12e−02	702233
BATCH	**IPOPT**	**48**	**73**	**first_order**	**2.59e+05**	**1.11e−16**	**249**
	Percival	**48**	**73**	**max_iter**	**2.59e+05**	**1.71e−12**	**42264**
BIGGSC4	**IPOPT**	**4**	**7**	**first_order**	**−2.45e+01**	**0.00e+00**	**76**
	Percival	**4**	**7**	**first_order**	**−2.45e+01**	**1.10e−07**	**63**
BRITGAS	IPOPT	450	360	max_iter	−3.80e−04	1.07e−04	9812
	Percival	450	360	max_time	1.62e+02	9.15e+00	19
BT1	**IPOPT**	**2**	**1**	**first_order**	**−1.00e+00**	**3.17e−07**	**24**
	Percival	**2**	**1**	**first_order**	**−1.00e+00**	**5.75e−07**	**80**
BT10	**IPOPT**	**2**	**2**	**first_order**	**−1.00e+00**	**4.18e−09**	**21**

Continua na próxima página

name	solver	nvar	ncon	status	objective	primal_feas	evals
	Percival	**2**	**2**	**first_order**	**−1.00e+00**	**8.41e−07**	**104**
BT11	**IPOPT**	**5**	**3**	**first_order**	**8.25e−01**	**1.19e−08**	**24**
	Percival	**5**	**3**	**first_order**	**8.25e−01**	**2.50e−07**	**137**
BT12	**IPOPT**	**5**	**3**	**first_order**	**6.19e+00**	**3.23e−13**	**15**
	Percival	**5**	**3**	**first_order**	**6.19e+00**	**8.73e−11**	**150**
BT13	**IPOPT**	**5**	**1**	**first_order**	**8.09e−08**	**6.13e−07**	**69**
	Percival	**5**	**1**	**first_order**	**−7.70e−82**	**1.24e−42**	**222**
BT2	**IPOPT**	**3**	**1**	**first_order**	**3.26e−02**	**2.19e−12**	**39**
	Percival	**3**	**1**	**first_order**	**3.26e−02**	**7.89e−12**	**296**
BT3	**IPOPT**	**5**	**3**	**first_order**	**4.09e+00**	**1.07e−14**	**6**
	Percival	**5**	**3**	**first_order**	**4.09e+00**	**3.73e−07**	**89**
BT4	IPOPT	3	2	first_order	−3.70e+00	1.03e−09	30
	Percival	**3**	**2**	**first_order**	**−4.55e+01**	**1.57e−08**	**126**
BT5	**IPOPT**	**3**	**2**	**first_order**	**9.62e+02**	**4.19e−07**	**21**
	Percival	**3**	**2**	**first_order**	**9.62e+02**	**9.86e−07**	**84**
BT6	**IPOPT**	**5**	**2**	**first_order**	**2.77e−01**	**5.31e−11**	**46**
	Percival	5	2	first_order	5.15e+00	2.16e−08	695
BT7	**IPOPT**	**5**	**3**	**first_order**	**3.07e+02**	**8.88e−16**	**60**
	Percival	**5**	**3**	**first_order**	**3.06e+02**	**7.82e−07**	**287**
BT8	**IPOPT**	**5**	**2**	**first_order**	**1.00e+00**	**5.24e−24**	**176**
	Percival	**5**	**2**	**first_order**	**1.00e+00**	**1.61e−08**	**151**
BT9	**IPOPT**	**4**	**2**	**first_order**	**−1.00e+00**	**1.68e−12**	**42**
	Percival	**4**	**2**	**first_order**	**−1.00e+00**	**1.10e−08**	**272**
BURKEHAN	IPOPT	1	1	infeasible	−3.35e−05	1.00e+00	28
	Percival	1	1	stalled	0.00e+00	1.00e+00	3775
BYRDSPHR	**IPOPT**	**3**	**2**	**first_order**	**−4.68e+00**	**2.06e−07**	**54**
	Percival	**3**	**2**	**first_order**	**−4.68e+00**	**2.43e−10**	**126**
C-RELOAD	**IPOPT**	**342**	**284**	**first_order**	**−1.02e+00**	**1.78e−07**	**934**
	Percival	**342**	**284**	**first_order**	**−1.02e+00**	**1.38e−08**	**1682**
CANTILVR	**IPOPT**	**5**	**1**	**first_order**	**1.34e+00**	**0.00e+00**	**36**
	Percival	**5**	**1**	**first_order**	**1.34e+00**	**2.46e−09**	**230**
CB2	**IPOPT**	**3**	**3**	**first_order**	**1.95e+00**	**0.00e+00**	**27**
	Percival	**3**	**3**	**first_order**	**1.95e+00**	**3.18e−09**	**174**
CB3	**IPOPT**	**3**	**3**	**first_order**	**2.00e+00**	**0.00e+00**	**30**
	Percival	**3**	**3**	**first_order**	**2.00e+00**	**4.99e−10**	**158**
CHACONN1	**IPOPT**	**3**	**3**	**first_order**	**1.95e+00**	**0.00e+00**	**21**
	Percival	**3**	**3**	**first_order**	**1.95e+00**	**2.59e−10**	**164**
CHACONN2	**IPOPT**	**3**	**3**	**first_order**	**2.00e+00**	**0.00e+00**	**31**
	Percival	**3**	**3**	**first_order**	**2.00e+00**	**5.34e−08**	**118**
CHAIN	**IPOPT**	**802**	**401**	**first_order**	**5.07e+00**	**7.08e−09**	**24**
	Percival	802	401	max_time	1.53e+00	2.79e−01	2600
CLEUVEN7	**IPOPT**	**360**	**946**	**first_order**	**6.83e+02**	**0.00e+00**	**90**
	Percival	360	946	max_time	6.83e+02	2.65e−04	242
COATINGNE	IPOPT	134	252	unknown	0.00e+00	∞	0
	Percival	134	252	max_time	7.28e−12	2.15e+00	17073
CONCON	**IPOPT**	**15**	**11**	**first_order**	**−6.23e+03**	**1.53e−10**	**30**
	Percival	**15**	**11**	**max_iter**	**−6.23e+03**	**4.17e−10**	**11914**
CONGIGMZ	**IPOPT**	**3**	**5**	**first_order**	**2.80e+01**	**9.99e−08**	**82**
	Percival	**3**	**5**	**first_order**	**2.80e+01**	**6.21e−07**	**143**
CORE1	**IPOPT**	**65**	**59**	**first_order**	**9.11e+01**	**9.40e−10**	**257**
	Percival	65	59	max_time	9.36e+01	2.29e−01	110923
CORE2	**IPOPT**	**157**	**134**	**first_order**	**7.29e+01**	**1.39e−12**	**593**
	Percival	157	134	max_time	7.44e+01	1.77e−01	44655
CRESC100	IPOPT	6	200	infeasible	−1.86e−11	2.91e−01	2429
	Percival	**6**	**200**	**first_order**	**5.68e−01**	**7.43e−11**	**2212**
CRESC4	**IPOPT**	**6**	**8**	**first_order**	**8.72e−01**	**0.00e+00**	**849**
	Percival	**6**	**8**	**first_order**	**8.72e−01**	**2.15e−10**	**912**
CRESC50	IPOPT	6	100	infeasible	−1.32e−10	5.57e−01	1400
	Percival	6	100	max_time	5.95e−01	4.18e−05	45832
CSFI1	**IPOPT**	**5**	**4**	**first_order**	**−4.91e+01**	**8.94e−12**	**49**
	Percival	**5**	**4**	**first_order**	**−4.91e+01**	**3.72e−09**	**308**
CSFI2	**IPOPT**	**5**	**4**	**first_order**	**5.50e+01**	**2.13e−14**	**69**
	Percival	**5**	**4**	**first_order**	**5.50e+01**	**1.74e−08**	**2151**
DALLASL	**IPOPT**	**906**	**667**	**first_order**	**−2.03e+05**	**4.16e−14**	**87**
	Percival	906	667	max_time	−2.01e+05	4.46e+00	334
DALLASM	**IPOPT**	**196**	**151**	**first_order**	**−4.82e+04**	**7.11e−15**	**61**
	Percival	196	151	max_time	−4.82e+04	1.66e−05	7195

Continua na próxima página

name	solver	nvar	ncon	status	objective	primal_feas	evals
DALLASS	**IPOPT**	**46**	**31**	**first_order**	**−3.24e+04**	**1.42e−14**	**87**
	Percival	**46**	**31**	**first_order**	**−3.24e+04**	**4.12e−08**	**934**
DECONVC	**IPOPT**	**63**	**1**	**first_order**	**2.57e−03**	**2.72e−15**	**259**
	Percival	**63**	**1**	**first_order**	**6.13e−12**	**3.19e−07**	**643**
DEGENLPA	**IPOPT**	**20**	**15**	**first_order**	**3.05e+00**	**2.74e−17**	**77**
	Percival	**20**	**15**	**first_order**	**3.06e+00**	**5.64e−09**	**20957**
DEGENLPB	**IPOPT**	**20**	**15**	**first_order**	**−3.08e+01**	**5.55e−17**	**99**
	Percival	**20**	**15**	**first_order**	**−3.07e+01**	**9.11e−13**	**3582**
DEMBO7	**IPOPT**	**16**	**20**	**first_order**	**1.75e+02**	**0.00e+00**	**122**
	Percival	**16**	**20**	**max_iter**	**1.75e+02**	**8.18e−12**	**90916**
DEMYMALO	**IPOPT**	**3**	**3**	**first_order**	**−3.00e+00**	**0.00e+00**	**36**
	Percival	**3**	**3**	**first_order**	**−3.00e+00**	**9.16e−07**	**130**
DEVGLA2NE	IPOPT	5	16	unknown	0.00e+00	∞	0
	Percival	**5**	**16**	**first_order**	**−2.42e−27**	**9.81e−08**	**335**
DIPIGRI	**IPOPT**	**7**	**4**	**first_order**	**6.81e+02**	**0.00e+00**	**46**
	Percival	**7**	**4**	**first_order**	**6.81e+02**	**1.43e−07**	**1897**
DISC2	**IPOPT**	**29**	**23**	**first_order**	**1.56e+00**	**4.12e−09**	**152**
	Percival	**29**	**23**	**first_order**	**1.56e+00**	**1.13e−08**	**344**
DISCS	IPOPT	36	66	infeasible	1.21e+01	4.00e+00	405
	Percival	36	66	max_time	2.05e+01	7.40e−03	43403
DIXCHLNG	IPOPT	10	5	first_order	2.47e+03	2.22e−16	33
	Percival	**10**	**5**	**first_order**	**1.84e+03**	**1.55e−07**	**398**
DIXCHLNV	**IPOPT**	**1000**	**500**	**first_order**	**0.00e+00**	**0.00e+00**	**139**
	Percival	1000	500	max_time	3.90e+07	4.42e+02	1135
DNIEPER	**IPOPT**	**61**	**24**	**first_order**	**1.87e+04**	**1.28e−09**	**93**
	Percival	**61**	**24**	**first_order**	**1.87e+04**	**1.42e−08**	**166**
DRUGDISE	**IPOPT**	**603**	**500**	**small_step**	**4.04e+02**	**1.80e−16**	**721**
	Percival	603	500	max_time	2.40e+00	5.22e−01	1705
DUAL1	**IPOPT**	**85**	**1**	**first_order**	**3.50e−02**	**3.37e−16**	**42**
	Percival	**85**	**1**	**first_order**	**3.50e−02**	**1.30e−10**	**46**
DUAL2	**IPOPT**	**96**	**1**	**first_order**	**3.37e−02**	**1.51e−16**	**36**
	Percival	**96**	**1**	**first_order**	**3.37e−02**	**1.62e−10**	**37**
DUAL3	**IPOPT**	**111**	**1**	**first_order**	**1.36e−01**	**5.80e−16**	**33**
	Percival	**111**	**1**	**first_order**	**1.36e−01**	**8.29e−11**	**37**
DUAL4	**IPOPT**	**75**	**1**	**first_order**	**7.46e−01**	**1.95e−16**	**33**
	Percival	**75**	**1**	**first_order**	**7.46e−01**	**4.65e−09**	**37**
DUALC1	**IPOPT**	**9**	**215**	**first_order**	**6.16e+03**	**2.78e−17**	**96**
	Percival	**9**	**215**	**first_order**	**6.16e+03**	**1.51e−11**	**14989**
DUALC2	**IPOPT**	**7**	**229**	**first_order**	**3.55e+03**	**1.72e−17**	**66**
	Percival	**7**	**229**	**first_order**	**3.55e+03**	**4.78e−08**	**446**
DUALC5	**IPOPT**	**8**	**278**	**first_order**	**4.27e+02**	**0.00e+00**	**39**
	Percival	**8**	**278**	**first_order**	**4.27e+02**	**5.84e−11**	**860**
DUALC8	**IPOPT**	**8**	**503**	**first_order**	**1.83e+04**	**1.25e−16**	**51**
	Percival	**8**	**503**	**max_time**	**1.83e+04**	**3.33e−12**	**8552**
EIGMAXA	**IPOPT**	**101**	**101**	**first_order**	**−1.00e+00**	**2.81e−08**	**44**
	Percival	**101**	**101**	**first_order**	**−1.00e+00**	**5.05e−08**	**68**
EIGMAXB	**IPOPT**	**101**	**101**	**first_order**	**−9.67e−04**	**8.40e−13**	**27**
	Percival	**101**	**101**	**first_order**	**−9.67e−04**	**4.81e−09**	**744**
EIGMAXC	**IPOPT**	**202**	**202**	**first_order**	**−1.00e+00**	**3.03e−09**	**27**
	Percival	**202**	**202**	**first_order**	**−1.00e+00**	**2.27e−08**	**71**
EIGMINA	**IPOPT**	**101**	**101**	**first_order**	**1.00e+00**	**2.81e−08**	**44**
	Percival	**101**	**101**	**first_order**	**1.00e+00**	**1.86e−08**	**68**
EIGMINB	**IPOPT**	**101**	**101**	**first_order**	**9.67e−04**	**8.41e−13**	**27**
	Percival	**101**	**101**	**first_order**	**9.67e−04**	**4.35e−08**	**1095**
EIGMINC	**IPOPT**	**202**	**202**	**first_order**	**1.00e+00**	**3.03e−09**	**27**
	Percival	**202**	**202**	**first_order**	**1.00e+00**	**6.94e−08**	**54**
ELATTAR	**IPOPT**	**7**	**102**	**first_order**	**7.42e+01**	**0.00e+00**	**313**
	Percival	7	102	stalled	0.00e+00	1.74e+02	632
ELEC	**IPOPT**	**600**	**200**	**first_order**	**1.84e+04**	**7.74e−11**	**463**
	Percival	**600**	**200**	**first_order**	**1.84e+04**	**9.80e−08**	**324**
EQC	**IPOPT**	**9**	**3**	**first_order**	**−8.30e+02**	**0.00e+00**	**26**
	Percival	9	3	max_eval	−9.40e+02	2.50e+00	2200634
ERRINBAR	**IPOPT**	**18**	**9**	**first_order**	**2.80e+01**	**7.18e−09**	**150**
	Percival	**18**	**9**	**first_order**	**2.80e+01**	**3.78e−09**	**18687**
EXPFITA	**IPOPT**	**5**	**22**	**first_order**	**1.14e−03**	**0.00e+00**	**109**
	Percival	**5**	**22**	**first_order**	**1.14e−03**	**4.76e−09**	**74**
EXPFITB	**IPOPT**	**5**	**102**	**first_order**	**5.02e−03**	**0.00e+00**	**182**

Continua na próxima página

name	solver	nvar	ncon	status	objective	primal_feas	evals
	Percival	**5**	**102**	**first_order**	**5.02e−03**	**2.64e−10**	**109**
EXPFITC	**IPOPT**	**5**	**502**	**first_order**	**2.33e−02**	**0.00e+00**	**335**
	Percival	**5**	**502**	**first_order**	**2.33e−02**	**1.22e−11**	**257**
EXTRASIM	**IPOPT**	**2**	**1**	**first_order**	**1.00e+00**	**0.00e+00**	**18**
	Percival	**2**	**1**	**first_order**	**1.00e+00**	**0.00e+00**	**37**
FCCU	**IPOPT**	**19**	**8**	**first_order**	**1.11e+01**	**1.78e−15**	**27**
	Percival	**19**	**8**	**first_order**	**1.11e+01**	**5.53e−09**	**70**
FEEDLOC	**IPOPT**	**90**	**259**	**first_order**	**1.77e−07**	**1.20e−07**	**129**
	Percival	90	259	max_time	1.39e−01	4.43e+01	34785
FLETCHER	**IPOPT**	**4**	**4**	**first_order**	**1.95e+01**	**1.33e−15**	**78**
	Percival	**4**	**4**	**first_order**	**1.95e+01**	**1.59e−07**	**89**
FLT	**IPOPT**	**2**	**2**	**first_order**	**0.00e+00**	**3.88e−07**	**24**
	Percival	**2**	**2**	**first_order**	**−4.67e−48**	**3.06e−25**	**55**
GENHS28	**IPOPT**	**10**	**8**	**first_order**	**9.27e−01**	**4.44e−16**	**6**
	Percival	**10**	**8**	**first_order**	**9.27e−01**	**1.68e−08**	**89**
GIGOMEZ1	**IPOPT**	**3**	**3**	**first_order**	**−3.00e+00**	**0.00e+00**	**51**
	Percival	**3**	**3**	**first_order**	**−3.00e+00**	**1.27e−10**	**281**
GIGOMEZ2	**IPOPT**	**3**	**3**	**first_order**	**1.95e+00**	**0.00e+00**	**27**
	Percival	**3**	**3**	**first_order**	**1.95e+00**	**3.09e−08**	**209**
GIGOMEZ3	**IPOPT**	**3**	**3**	**first_order**	**2.00e+00**	**0.00e+00**	**27**
	Percival	**3**	**3**	**first_order**	**2.00e+00**	**9.22e−10**	**150**
GMNCASE1	**IPOPT**	**175**	**300**	**first_order**	**2.67e−01**	**0.00e+00**	**36**
	Percival	**175**	**300**	**first_order**	**2.67e−01**	**1.78e−07**	**243**
GMNCASE4	**IPOPT**	**175**	**350**	**first_order**	**5.95e+03**	**0.00e+00**	**147**
	Percival	**175**	**350**	**first_order**	**5.95e+03**	**1.37e−09**	**1116**
GOFFIN	**IPOPT**	**51**	**50**	**first_order**	**4.54e−06**	**0.00e+00**	**27**
	Percival	**51**	**50**	**first_order**	**3.80e−20**	**7.65e−19**	**117**
GOULDQP1	**IPOPT**	**32**	**17**	**first_order**	**−3.49e+03**	**1.78e−15**	**100**
	Percival	**32**	**17**	**first_order**	**−3.49e+03**	**7.83e−08**	**325**
GROUPING	IPOPT	100	125	unknown	0.00e+00	∞	0
	Percival	**100**	**125**	**first_order**	**1.39e+01**	**0.00e+00**	**5**
HAIFAM	**IPOPT**	**99**	**150**	**first_order**	**−4.50e+01**	**0.00e+00**	**99**
	Percival	99	150	max_time	−4.37e+01	2.38e−03	6882
HAIFAS	**IPOPT**	**13**	**9**	**first_order**	**−4.50e−01**	**0.00e+00**	**28**
	Percival	**13**	**9**	**first_order**	**−4.50e−01**	**2.57e−08**	**92**
HALDMADS	IPOPT	6	42	first_order	2.45e+00	0.00e+00	485
	Percival	**6**	**42**	**first_order**	**1.22e−04**	**6.24e−11**	**338**
HATFLDH	**IPOPT**	**4**	**7**	**first_order**	**−2.45e+01**	**0.00e+00**	**45**
	Percival	**4**	**7**	**first_order**	**−2.45e+01**	**5.35e−09**	**80**
HIE1372D	**IPOPT**	**637**	**525**	**first_order**	**2.78e+02**	**3.98e−13**	**54**
	Percival	637	525	max_time	2.82e+02	1.48e−03	44080
HIMMELBI	**IPOPT**	**100**	**12**	**first_order**	**−1.74e+03**	**0.00e+00**	**78**
	Percival	**100**	**12**	**first_order**	**−1.74e+03**	**8.88e−08**	**13503**
HIMMELBJ	**IPOPT**	**45**	**14**	**unknown**	**−1.90e+03**	**2.31e−08**	**73570**
	Percival	**45**	**14**	**first_order**	**−1.91e+03**	**1.50e−09**	**150879**
HIMMELBK	**IPOPT**	**24**	**14**	**first_order**	**5.18e−02**	**6.97e−09**	**54**
	Percival	**24**	**14**	**first_order**	**5.18e−02**	**1.66e−07**	**734**
HIMMELP2	**IPOPT**	**2**	**1**	**first_order**	**−8.20e+00**	**0.00e+00**	**57**
	Percival	**2**	**1**	**first_order**	**−8.20e+00**	**3.90e−08**	**4200**
HIMMELP3	**IPOPT**	**2**	**2**	**first_order**	**−5.90e+01**	**0.00e+00**	**39**
	Percival	2	2	first_order	−8.20e+00	1.71e−12	149
HIMMELP4	**IPOPT**	**2**	**3**	**first_order**	**−5.90e+01**	**0.00e+00**	**75**
	Percival	2	3	first_order	−8.20e+00	2.39e−08	194
HIMMELP5	**IPOPT**	**2**	**3**	**first_order**	**−5.90e+01**	**0.00e+00**	**348**
	Percival	**2**	**3**	**first_order**	**−5.90e+01**	**0.00e+00**	**2346**
HIMMELP6	**IPOPT**	**2**	**5**	**first_order**	**−5.90e+01**	**0.00e+00**	**79**
	Percival	2	5	first_order	−8.20e+00	1.71e−12	1556
HONG	**IPOPT**	**4**	**1**	**first_order**	**2.26e+01**	**5.55e−17**	**30**
	Percival	**4**	**1**	**first_order**	**2.26e+01**	**7.85e−08**	**113**
HS10	**IPOPT**	**2**	**1**	**first_order**	**−1.00e+00**	**0.00e+00**	**39**
	Percival	**2**	**1**	**first_order**	**−1.00e+00**	**5.91e−07**	**165**
HS100	**IPOPT**	**7**	**4**	**first_order**	**6.81e+02**	**0.00e+00**	**46**
	Percival	**7**	**4**	**first_order**	**6.81e+02**	**1.73e−07**	**1866**
HS100LNP	**IPOPT**	**7**	**2**	**first_order**	**6.81e+02**	**1.42e−14**	**63**
	Percival	**7**	**2**	**first_order**	**6.81e+02**	**5.69e−09**	**195**
HS100MOD	**IPOPT**	**7**	**4**	**first_order**	**6.79e+02**	**0.00e+00**	**46**
	Percival	7	4	max_eval	6.79e+02	3.12e−02	788436

Continua na próxima página

name	solver	nvar	ncon	status	objective	primal_feas	evals
HS101	**IPOPT**	**7**	**5**	**first_order**	**1.81e+03**	**0.00e+00**	**398**
	Percival	7	5	max_time	2.62e+03	9.15e−04	574238
HS102	**IPOPT**	**7**	**5**	**first_order**	**9.12e+02**	**0.00e+00**	**76**
	Percival	7	5	max_time	6.05e+02	1.46e−01	491549
HS103	**IPOPT**	**7**	**5**	**first_order**	**5.44e+02**	**0.00e+00**	**182**
	Percival	7	5	max_eval	5.08e+02	2.42e−02	256572
HS104	**IPOPT**	**8**	**5**	**first_order**	**3.95e+00**	**0.00e+00**	**27**
	Percival	**8**	**5**	**first_order**	**3.95e+00**	**1.75e−07**	**135**
HS105	**IPOPT**	**8**	**1**	**first_order**	**1.04e+03**	**0.00e+00**	**54**
	Percival	8	1	first_order	1.06e+03	5.81e−11	79
HS106	**IPOPT**	**8**	**6**	**first_order**	**7.05e+03**	**0.00e+00**	**48**
	Percival	8	6	max_eval	6.51e+03	8.02e−02	234474
HS107	**IPOPT**	**9**	**6**	**first_order**	**5.06e+03**	**2.22e−16**	**30**
	Percival	**9**	**6**	**first_order**	**5.06e+03**	**2.68e−07**	**142**
HS108	IPOPT	9	13	first_order	−6.75e−01	0.00e+00	51
	Percival	**9**	**13**	**first_order**	**−8.66e−01**	**6.37e−10**	**219**
HS109	**IPOPT**	**9**	**10**	**first_order**	**5.36e+03**	**2.50e−11**	**96**
	Percival	9	10	max_iter	5.48e+03	3.72e−09	74616
HS11	**IPOPT**	**2**	**1**	**first_order**	**−8.50e+00**	**0.00e+00**	**27**
	Percival	**2**	**1**	**first_order**	**−8.50e+00**	**1.39e−08**	**100**
HS111	**IPOPT**	**10**	**3**	**first_order**	**−4.78e+01**	**9.45e−08**	**45**
	Percival	**10**	**3**	**first_order**	**−4.78e+01**	**5.33e−07**	**116**
HS111LNP	**IPOPT**	**10**	**3**	**first_order**	**−4.78e+01**	**2.72e−07**	**45**
	Percival	**10**	**3**	**first_order**	**−4.78e+01**	**5.33e−07**	**116**
HS112	**IPOPT**	**10**	**3**	**first_order**	**−4.78e+01**	**2.78e−17**	**54**
	Percival	**10**	**3**	**first_order**	**−4.78e+01**	**6.90e−07**	**83**
HS113	**IPOPT**	**10**	**8**	**first_order**	**2.43e+01**	**0.00e+00**	**36**
	Percival	**10**	**8**	**first_order**	**2.43e+01**	**4.92e−08**	**692**
HS114	**IPOPT**	**10**	**11**	**first_order**	**−1.77e+03**	**4.26e−14**	**60**
	Percival	**10**	**11**	**first_order**	**−1.77e+03**	**3.86e−08**	**677**
HS116	**IPOPT**	**13**	**14**	**first_order**	**9.76e+01**	**0.00e+00**	**75**
	Percival	**13**	**14**	**max_iter**	**9.76e+01**	**9.69e−11**	**50584**
HS117	**IPOPT**	**15**	**5**	**first_order**	**3.23e+01**	**0.00e+00**	**69**
	Percival	**15**	**5**	**first_order**	**3.23e+01**	**9.40e−11**	**761**
HS118	**IPOPT**	**15**	**17**	**first_order**	**6.65e+02**	**0.00e+00**	**36**
	Percival	**15**	**17**	**first_order**	**6.65e+02**	**1.64e−08**	**258**
HS119	**IPOPT**	**16**	**8**	**first_order**	**2.45e+02**	**6.66e−16**	**48**
	Percival	**16**	**8**	**first_order**	**2.45e+02**	**1.74e−09**	**225**
HS12	**IPOPT**	**2**	**1**	**first_order**	**−3.00e+01**	**0.00e+00**	**27**
	Percival	**2**	**1**	**first_order**	**−3.00e+01**	**4.60e−08**	**114**
HS13	IPOPT	2	1	first_order	9.95e−01	0.00e+00	188
	Percival	**2**	**1**	**first_order**	**9.83e−01**	**5.76e−07**	**457**
HS14	**IPOPT**	**2**	**2**	**first_order**	**1.39e+00**	**2.22e−16**	**24**
	Percival	**2**	**2**	**first_order**	**1.39e+00**	**9.98e−08**	**79**
HS15	**IPOPT**	**2**	**2**	**first_order**	**3.06e+02**	**0.00e+00**	**55**
	Percival	**2**	**2**	**first_order**	**3.06e+02**	**6.24e−08**	**195**
HS16	**IPOPT**	**2**	**2**	**first_order**	**2.31e+01**	**0.00e+00**	**30**
	Percival	**2**	**2**	**first_order**	**2.31e+01**	**3.18e−07**	**81**
HS17	**IPOPT**	**2**	**2**	**first_order**	**1.00e+00**	**0.00e+00**	**51**
	Percival	**2**	**2**	**first_order**	**1.00e+00**	**1.67e−07**	**229**
HS18	**IPOPT**	**2**	**2**	**first_order**	**5.00e+00**	**0.00e+00**	**51**
	Percival	**2**	**2**	**first_order**	**5.00e+00**	**9.77e−08**	**730**
HS19	**IPOPT**	**2**	**2**	**first_order**	**−6.96e+03**	**0.00e+00**	**39**
	Percival	**2**	**2**	**first_order**	**−6.96e+03**	**7.60e−11**	**253**
HS20	**IPOPT**	**2**	**3**	**first_order**	**4.02e+01**	**0.00e+00**	**27**
	Percival	**2**	**3**	**first_order**	**4.02e+01**	**2.54e−07**	**84**
HS21	**IPOPT**	**2**	**1**	**first_order**	**−1.00e+02**	**0.00e+00**	**27**
	Percival	**2**	**1**	**first_order**	**−1.00e+02**	**0.00e+00**	**59**
HS21MOD	**IPOPT**	**7**	**1**	**first_order**	**−9.60e+01**	**0.00e+00**	**39**
	Percival	**7**	**1**	**first_order**	**−9.60e+01**	**0.00e+00**	**59**
HS22	**IPOPT**	**2**	**2**	**first_order**	**1.00e+00**	**0.00e+00**	**21**
	Percival	**2**	**2**	**first_order**	**1.00e+00**	**1.69e−09**	**107**
HS23	**IPOPT**	**2**	**5**	**first_order**	**2.00e+00**	**0.00e+00**	**34**
	Percival	**2**	**5**	**first_order**	**2.00e+00**	**1.09e−08**	**268**
HS24	**IPOPT**	**2**	**3**	**first_order**	**−1.00e+00**	**0.00e+00**	**40**
	Percival	**2**	**3**	**first_order**	**−1.00e+00**	**3.57e−07**	**59**
HS26	**IPOPT**	**3**	**1**	**first_order**	**2.17e−12**	**8.29e−07**	**60**

Continua na próxima página

name	solver	nvar	ncon	status	objective	primal_feas	evals
	Percival	3	1	first_order	2.08e−13	1.86e−07	156
HS268	IPOPT	5	5	first_order	1.80e−07	0.00e+00	45
	Percival	5	5	first_order	3.64e−11	2.79e−08	38457
HS27	IPOPT	3	1	first_order	4.00e−02	9.59e−08	261
	Percival	3	1	first_order	4.00e−02	3.06e−11	224
HS28	IPOPT	3	1	first_order	3.08e−31	4.44e−16	6
	Percival	3	1	first_order	1.05e−31	1.11e−13	53
HS29	IPOPT	3	1	first_order	−2.26e+01	0.00e+00	27
	Percival	3	1	first_order	−2.26e+01	1.29e−11	174
HS30	IPOPT	3	1	first_order	1.00e+00	0.00e+00	30
	Percival	3	1	first_order	1.00e+00	1.49e−09	88
HS31	IPOPT	3	1	first_order	6.00e+00	0.00e+00	24
	Percival	3	1	first_order	6.00e+00	1.32e−08	66
HS32	IPOPT	3	2	first_order	1.00e+00	0.00e+00	42
	Percival	3	2	first_order	1.00e+00	2.72e−07	98
HS33	IPOPT	3	2	first_order	−4.59e+00	0.00e+00	40
	Percival	3	2	first_order	−4.00e+00	7.82e−07	64
HS34	IPOPT	3	2	first_order	−8.34e−01	0.00e+00	30
	Percival	3	2	first_order	−8.34e−01	6.03e−07	111
HS35	IPOPT	3	1	first_order	1.11e−01	0.00e+00	24
	Percival	3	1	first_order	1.11e−01	5.22e−09	54
HS35I	IPOPT	3	1	first_order	1.11e−01	0.00e+00	24
	Percival	3	1	first_order	1.11e−01	5.22e−09	54
HS35MOD	IPOPT	3	1	first_order	2.50e−01	0.00e+00	36
	Percival	3	1	first_order	2.50e−01	1.27e−08	53
HS36	IPOPT	3	1	first_order	−3.30e+03	0.00e+00	39
	Percival	3	1	first_order	−3.30e+03	0.00e+00	35
HS37	IPOPT	3	2	first_order	−3.46e+03	0.00e+00	36
	Percival	3	2	first_order	−3.46e+03	2.12e−07	108
HS39	IPOPT	4	2	first_order	−1.00e+00	1.68e−12	42
	Percival	4	2	first_order	−1.00e+00	1.10e−08	272
HS40	IPOPT	4	3	first_order	−2.50e−01	1.86e−10	12
	Percival	4	3	first_order	−2.50e−01	1.23e−07	74
HS41	IPOPT	4	1	first_order	1.93e+00	0.00e+00	32
	Percival	4	1	first_order	1.93e+00	8.95e−08	40
HS42	IPOPT	4	2	first_order	1.39e+01	5.21e−09	12
	Percival	4	2	first_order	1.39e+01	1.57e−07	50
HS43	IPOPT	4	3	first_order	−4.40e+01	0.00e+00	30
	Percival	4	3	first_order	−4.40e+01	1.43e−08	155
HS44	IPOPT	4	6	first_order	−1.30e+01	0.00e+00	58
	Percival	4	6	first_order	−1.30e+01	2.06e−07	63
HS44NEW	IPOPT	4	6	first_order	−1.50e+01	0.00e+00	45
	Percival	4	6	first_order	−1.30e+01	1.87e−07	54
HS46	IPOPT	5	2	first_order	1.44e−11	9.47e−07	42
	Percival	5	2	first_order	7.38e−15	2.37e−07	226
HS47	IPOPT	5	3	first_order	3.36e−11	1.04e−07	52
	Percival	5	3	first_order	−2.67e−02	2.90e−08	175
HS48	IPOPT	5	2	first_order	3.94e−31	4.44e−16	6
	Percival	5	2	first_order	0.00e+00	3.17e−09	45
HS49	IPOPT	5	2	first_order	1.38e−09	4.44e−16	51
	Percival	5	2	first_order	6.13e−16	3.13e−10	352
HS50	IPOPT	5	3	first_order	0.00e+00	0.00e+00	30
	Percival	5	3	first_order	1.84e−18	6.16e−09	72
HS51	IPOPT	5	3	first_order	4.93e−32	0.00e+00	6
	Percival	5	3	first_order	9.15e−14	1.82e−07	61
HS52	IPOPT	5	3	first_order	5.33e+00	4.44e−16	6
	Percival	5	3	first_order	5.33e+00	2.10e−07	50
HS53	IPOPT	5	3	first_order	4.09e+00	1.11e−16	21
	Percival	5	3	first_order	4.09e+00	1.77e−07	59
HS54	IPOPT	6	1	first_order	−9.08e−01	9.09e−13	48
	Percival	6	1	first_order	−8.67e−01	2.74e−07	95
HS55	IPOPT	6	6	first_order	6.71e+00	3.55e−08	9
	Percival	6	6	first_order	6.67e+00	5.24e−16	54
HS56	IPOPT	7	4	first_order	−3.46e+00	6.04e−14	33
	Percival	7	4	first_order	−3.46e+00	4.18e−07	70
HS57	IPOPT	2	1	first_order	3.06e−02	0.00e+00	74
	Percival	2	1	first_order	2.85e−02	6.58e−08	97

Continua na próxima página

name	solver	nvar	ncon	status	objective	primal_feas	evals
HS59	**IPOPT**	**2**	**3**	**first_order**	**−7.80e+00**	**0.00e+00**	**160**
	Percival	**2**	**3**	**first_order**	**−7.80e+00**	**1.02e−10**	**492**
HS6	**IPOPT**	**2**	**1**	**first_order**	**0.00e+00**	**0.00e+00**	**19**
	Percival	**2**	**1**	**first_order**	**1.03e−19**	**7.10e−10**	**271**
HS60	**IPOPT**	**3**	**1**	**first_order**	**3.26e−02**	**8.88e−16**	**24**
	Percival	**3**	**1**	**first_order**	**3.26e−02**	**6.40e−08**	**79**
HS61	**IPOPT**	**3**	**2**	**first_order**	**−1.44e+02**	**9.12e−12**	**30**
	Percival	**3**	**2**	**first_order**	**−1.44e+02**	**1.92e−10**	**122**
HS62	**IPOPT**	**3**	**1**	**first_order**	**−2.63e+04**	**6.94e−18**	**29**
	Percival	**3**	**1**	**first_order**	**−2.63e+04**	**3.92e−07**	**207**
HS63	**IPOPT**	**3**	**2**	**first_order**	**9.62e+02**	**8.70e−14**	**24**
	Percival	**3**	**2**	**first_order**	**9.62e+02**	**1.08e−07**	**83**
HS64	**IPOPT**	**3**	**1**	**first_order**	**6.30e+03**	**0.00e+00**	**54**
	Percival	**3**	**1**	**first_order**	**6.30e+03**	**7.21e−07**	**213**
HS65	**IPOPT**	**3**	**1**	**first_order**	**9.54e−01**	**0.00e+00**	**149**
	Percival	**3**	**1**	**first_order**	**9.54e−01**	**1.84e−08**	**264**
HS66	**IPOPT**	**3**	**2**	**first_order**	**5.18e−01**	**0.00e+00**	**24**
	Percival	**3**	**2**	**first_order**	**5.18e−01**	**2.83e−08**	**60**
HS67	**IPOPT**	**3**	**14**	**first_order**	**−1.16e+03**	**0.00e+00**	**33**
	Percival	**3**	**14**	**first_order**	**−1.16e+03**	**7.54e−13**	**440**
HS68	**IPOPT**	**4**	**2**	**first_order**	**−9.20e−01**	**1.37e−09**	**63**
	Percival	**4**	**2**	**first_order**	**−9.20e−01**	**8.05e−08**	**414**
HS69	**IPOPT**	**4**	**2**	**first_order**	**−9.57e+02**	**3.77e−15**	**54**
	Percival	**4**	**2**	**first_order**	**−9.57e+02**	**1.55e−07**	**263**
HS7	**IPOPT**	**2**	**1**	**first_order**	**−1.73e+00**	**2.51e−13**	**84**
	Percival	**2**	**1**	**first_order**	**−1.73e+00**	**8.49e−10**	**244**
HS70	**IPOPT**	**4**	**1**	**first_order**	**7.50e−03**	**0.00e+00**	**104**
	Percival	**4**	**1**	**first_order**	**7.50e−03**	**7.37e−07**	**189**
HS71	**IPOPT**	**4**	**2**	**first_order**	**1.70e+01**	**1.62e−11**	**27**
	Percival	**4**	**2**	**first_order**	**1.70e+01**	**3.58e−10**	**118**
HS72	**IPOPT**	**4**	**2**	**first_order**	**7.28e+02**	**0.00e+00**	**51**
	Percival	**4**	**2**	**first_order**	**7.28e+02**	**3.30e−08**	**613**
HS73	**IPOPT**	**4**	**3**	**first_order**	**2.99e+01**	**6.94e−18**	**27**
	Percival	4	3	exception	∞	∞	0
HS74	**IPOPT**	**4**	**5**	**first_order**	**5.13e+03**	**1.14e−13**	**30**
	Percival	**4**	**5**	**first_order**	**5.13e+03**	**3.40e−10**	**3999**
HS75	**IPOPT**	**4**	**5**	**first_order**	**5.17e+03**	**9.61e−10**	**27**
	Percival	**4**	**5**	**max_eval**	**5.17e+03**	**1.43e−07**	**464072**
HS76	**IPOPT**	**4**	**3**	**first_order**	**−4.68e+00**	**0.00e+00**	**24**
	Percival	**4**	**3**	**first_order**	**−4.68e+00**	**3.38e−09**	**51**
HS76I	**IPOPT**	**4**	**3**	**first_order**	**−4.68e+00**	**0.00e+00**	**24**
	Percival	**4**	**3**	**first_order**	**−4.68e+00**	**3.38e−09**	**51**
HS77	**IPOPT**	**5**	**2**	**first_order**	**2.42e−01**	**2.26e−10**	**37**
	Percival	5	2	first_order	5.53e+00	4.38e−09	304
HS78	**IPOPT**	**5**	**3**	**first_order**	**−2.92e+00**	**5.56e−12**	**15**
	Percival	**5**	**3**	**first_order**	**−2.92e+00**	**9.59e−07**	**62**
HS79	**IPOPT**	**5**	**3**	**first_order**	**7.88e−02**	**3.89e−09**	**15**
	Percival	**5**	**3**	**first_order**	**7.88e−02**	**1.89e−09**	**130**
HS8	**IPOPT**	**2**	**2**	**first_order**	**−1.00e+00**	**1.25e−10**	**18**
	Percival	**2**	**2**	**first_order**	**−1.00e+00**	**4.92e−13**	**52**
HS80	**IPOPT**	**5**	**3**	**first_order**	**5.39e−02**	**7.09e−11**	**21**
	Percival	**5**	**3**	**first_order**	**5.39e−02**	**1.67e−07**	**65**
HS81	**IPOPT**	**5**	**3**	**first_order**	**5.39e−02**	**7.03e−11**	**24**
	Percival	**5**	**3**	**first_order**	**5.39e−02**	**1.16e−08**	**82**
HS83	**IPOPT**	**5**	**3**	**first_order**	**−3.07e+04**	**0.00e+00**	**51**
	Percival	**5**	**3**	**first_order**	**−3.07e+04**	**4.79e−07**	**115**
HS84	**IPOPT**	**5**	**3**	**first_order**	**−5.28e+06**	**0.00e+00**	**57**
	Percival	**5**	**3**	**max_iter**	**−5.28e+06**	**4.13e−10**	**42790**
HS85	**IPOPT**	**5**	**21**	**first_order**	**−2.22e+00**	**0.00e+00**	**66**
	Percival	5	21	max_time	−2.72e+07	1.81e+02	24901
HS86	**IPOPT**	**5**	**10**	**first_order**	**−3.23e+01**	**0.00e+00**	**33**
	Percival	**5**	**10**	**first_order**	**−3.23e+01**	**2.83e−07**	**85**
HS87	IPOPT	6	4	max_iter	8.81e+03	3.28e+01	27292
	Percival	**6**	**4**	**max_iter**	**9.00e+03**	**8.87e−11**	**64819**
HS88	**IPOPT**	**2**	**1**	**first_order**	**1.36e+00**	**0.00e+00**	**52**
	Percival	**2**	**1**	**first_order**	**1.36e+00**	**1.89e−07**	**339**
HS89	**IPOPT**	**3**	**1**	**first_order**	**1.36e+00**	**0.00e+00**	**80**

Continua na próxima página

name	solver	nvar	ncon	status	objective	primal_feas	evals
	Percival	**3**	**1**	**first_order**	**1.36e+00**	**3.95e−07**	**530**
HS9	**IPOPT**	**2**	**1**	**first_order**	**−5.00e−01**	**0.00e+00**	**14**
	Percival	**2**	**1**	**first_order**	**−5.00e−01**	**5.18e−07**	**50**
HS90	**IPOPT**	**4**	**1**	**first_order**	**1.36e+00**	**0.00e+00**	**72**
	Percival	**4**	**1**	**first_order**	**1.36e+00**	**1.89e−07**	**671**
HS91	**IPOPT**	**5**	**1**	**first_order**	**1.36e+00**	**0.00e+00**	**42**
	Percival	**5**	**1**	**first_order**	**1.36e+00**	**1.21e−07**	**705**
HS92	**IPOPT**	**6**	**1**	**first_order**	**1.36e+00**	**0.00e+00**	**65**
	Percival	**6**	**1**	**first_order**	**1.36e+00**	**1.89e−07**	**581**
HS93	**IPOPT**	**6**	**2**	**first_order**	**1.35e+02**	**0.00e+00**	**30**
	Percival	6	2	stalled	0.00e+00	2.07e+00	452
HS95	**IPOPT**	**6**	**4**	**first_order**	**1.56e−02**	**0.00e+00**	**72**
	Percival	**6**	**4**	**first_order**	**1.56e−02**	**1.42e−14**	**130**
HS96	**IPOPT**	**6**	**4**	**first_order**	**1.56e−02**	**0.00e+00**	**58**
	Percival	**6**	**4**	**first_order**	**1.56e−02**	**1.42e−14**	**169**
HS97	**IPOPT**	**6**	**4**	**first_order**	**4.07e+00**	**0.00e+00**	**54**
	Percival	**6**	**4**	**first_order**	**4.07e+00**	**3.91e−12**	**321**
HS98	IPOPT	6	4	first_order	4.07e+00	0.00e+00	63
	Percival	**6**	**4**	**first_order**	**3.14e+00**	**2.78e−14**	**533**
HS99	**IPOPT**	**7**	**2**	**first_order**	**−8.31e+08**	**1.46e−10**	**24**
	Percival	**7**	**2**	**max_iter**	**−8.31e+08**	**5.53e−08**	**27929**
HS99EXP	**IPOPT**	**31**	**21**	**first_order**	**−1.26e+12**	**1.11e−07**	**68**
	Percival	31	21	max_eval	−1.13e+12	6.58e+02	381065
HUBFIT	**IPOPT**	**2**	**1**	**first_order**	**1.69e−02**	**0.00e+00**	**27**
	Percival	**2**	**1**	**first_order**	**1.69e−02**	**4.05e−08**	**48**
HYDROELM	**IPOPT**	**505**	**504**	**first_order**	**−3.58e+06**	**0.00e+00**	**630**
	Percival	505	504	max_time	−3.59e+06	9.94e+00	19925
HYDROELS	**IPOPT**	**169**	**168**	**first_order**	**−3.58e+06**	**0.00e+00**	**363**
	Percival	**169**	**168**	**first_order**	**−3.58e+06**	**4.10e−09**	**26493**
KISSING	**IPOPT**	**127**	**903**	**first_order**	**8.45e−01**	**6.84e−07**	**923**
	Percival	127	903	first_order	1.00e+00	1.92e−10	307
KISSING2	**IPOPT**	**100**	**625**	**first_order**	**6.63e+00**	**0.00e+00**	**367**
	Percival	100	625	max_time	1.72e+01	2.16e−02	2512
KIWCRESC	**IPOPT**	**3**	**2**	**first_order**	**1.72e−07**	**0.00e+00**	**31**
	Percival	**3**	**2**	**first_order**	**3.35e−09**	**9.49e−09**	**169**
LAKES	**IPOPT**	**90**	**78**	**first_order**	**3.51e+05**	**2.80e−11**	**51**
	Percival	90	78	max_time	5.31e+08	8.33e+00	83390
LAUNCH	**IPOPT**	**25**	**28**	**first_order**	**9.00e+00**	**6.75e−11**	**65**
	Percival	25	28	stalled	−4.45e+04	2.86e+01	39832
LEAKNET	**IPOPT**	**156**	**153**	**first_order**	**8.05e+00**	**5.01e−13**	**55**
	Percival	156	153	max_time	7.20e+00	1.15e−02	21341
LEUVEN7	**IPOPT**	**360**	**946**	**first_order**	**6.95e+02**	**0.00e+00**	**90**
	Percival	360	946	max_time	6.95e+02	8.09e−04	207
LEWISPOL	IPOPT	6	9	unknown	0.00e+00	∞	0
	Percival	6	9	infeasible	1.21e+00	3.06e−05	1151
LHAIFAM	IPOPT	99	150	unknown	0.00e+00	∞	0
	Percival	99	150	exception	∞	∞	0
LIN	**IPOPT**	**4**	**2**	**first_order**	**−1.76e−02**	**2.78e−17**	**27**
	Percival	**4**	**2**	**first_order**	**−1.76e−02**	**8.20e−09**	**96**
LINSPANH	**IPOPT**	**97**	**33**	**first_order**	**−7.70e+01**	**5.85e−12**	**48**
	Percival	**97**	**33**	**first_order**	**−7.70e+01**	**3.15e−07**	**52**
LOADBAL	**IPOPT**	**31**	**31**	**first_order**	**4.53e−01**	**3.19e−14**	**43**
	Percival	**31**	**31**	**first_order**	**4.53e−01**	**3.20e−09**	**389**
LOOTSMA	**IPOPT**	**3**	**2**	**first_order**	**1.41e+00**	**0.00e+00**	**43**
	Percival	3	2	stalled	−8.00e+00	2.83e+00	652
LOTSCHD	**IPOPT**	**12**	**7**	**first_order**	**2.40e+03**	**1.14e−14**	**45**
	Percival	**12**	**7**	**first_order**	**2.40e+03**	**6.32e−07**	**45**
LSNNODOC	**IPOPT**	**5**	**4**	**max_iter**	**1.23e+02**	**2.53e−09**	**11089**
	Percival	**5**	**4**	**first_order**	**1.23e+02**	**6.50e−09**	**100**
LSQFIT	**IPOPT**	**2**	**1**	**first_order**	**3.38e−02**	**0.00e+00**	**24**
	Percival	**2**	**1**	**first_order**	**3.38e−02**	**3.23e−07**	**48**
MADSEN	**IPOPT**	**3**	**6**	**first_order**	**6.16e−01**	**0.00e+00**	**67**
	Percival	**3**	**6**	**first_order**	**6.16e−01**	**1.23e−10**	**255**
MADSSCHJ	**IPOPT**	**201**	**398**	**first_order**	**−4.99e+03**	**0.00e+00**	**136**
	Percival	201	398	max_time	−1.17e+03	1.11e−02	4232
MAKELA1	**IPOPT**	**3**	**2**	**first_order**	**−1.41e+00**	**0.00e+00**	**57**
	Percival	**3**	**2**	**first_order**	**−1.41e+00**	**4.09e−07**	**87**

Continua na próxima página

name	solver	nvar	ncon	status	objective	primal_feas	evals
MAKELA2	**IPOPT**	**3**	**3**	**first_order**	**7.20e+00**	**0.00e+00**	**24**
	Percival	**3**	**3**	**first_order**	**7.20e+00**	**3.24e−10**	**250**
MAKELA3	**IPOPT**	**21**	**20**	**first_order**	**1.81e−06**	**0.00e+00**	**51**
	Percival	**21**	**20**	**first_order**	**7.08e−12**	**3.17e−11**	**1233**
MAKELA4	**IPOPT**	**21**	**40**	**first_order**	**3.63e−06**	**0.00e+00**	**24**
	Percival	**21**	**40**	**first_order**	**−1.44e−18**	**9.23e−18**	**52**
MARATOS	**IPOPT**	**2**	**1**	**first_order**	**−1.00e+00**	**3.19e−08**	**12**
	Percival	**2**	**1**	**first_order**	**−1.00e+00**	**7.79e−09**	**68**
MATRIX2	**IPOPT**	**6**	**2**	**first_order**	**3.88e−06**	**0.00e+00**	**39**
	Percival	**6**	**2**	**first_order**	**1.49e−23**	**7.12e−08**	**72**
MCONCON	**IPOPT**	**15**	**11**	**first_order**	**−6.23e+03**	**1.53e−10**	**30**
	Percival	**15**	**11**	**max_iter**	**−6.23e+03**	**2.75e−10**	**9580**
MESH	IPOPT	41	48	unknown	−1.44e+37	7.05e+00	438
	Percival	41	48	max_time	−7.99e+07	4.19e+00	566302
MIFFLIN1	**IPOPT**	**3**	**2**	**first_order**	**−1.00e+00**	**0.00e+00**	**21**
	Percival	**3**	**2**	**first_order**	**−1.00e+00**	**1.06e−07**	**99**
MIFFLIN2	**IPOPT**	**3**	**2**	**first_order**	**−1.00e+00**	**0.00e+00**	**48**
	Percival	**3**	**2**	**first_order**	**−1.00e+00**	**3.92e−12**	**213**
MINMAXBD	**IPOPT**	**5**	**20**	**first_order**	**1.16e+02**	**0.00e+00**	**146**
	Percival	**5**	**20**	**first_order**	**1.16e+02**	**1.35e−11**	**2042**
MINMAXRB	**IPOPT**	**3**	**4**	**first_order**	**3.54e−07**	**0.00e+00**	**31**
	Percival	**3**	**4**	**first_order**	**4.68e−10**	**9.35e−10**	**530**
MISTAKE	**IPOPT**	**9**	**13**	**first_order**	**−1.00e+00**	**3.83e−07**	**45**
	Percival	9	13	first_order	−6.57e−01	1.63e−10	222
MRIBASIS	**IPOPT**	**36**	**55**	**first_order**	**1.82e+01**	**1.62e−12**	**67**
	Percival	**36**	**55**	**first_order**	**1.82e+01**	**2.50e−13**	**1275**
MSS1	**IPOPT**	**90**	**73**	**unknown**	**−1.60e+01**	**1.12e−10**	**1244**
	Percival	**90**	**73**	**first_order**	**−1.60e+01**	**1.35e−12**	**27906**
MSS2	IPOPT	756	703	max_time	0.00e+00	4.13e−04	164
	Percival	756	703	max_time	−1.25e+02	7.73e−04	2818
MWRIGHT	**IPOPT**	**5**	**3**	**first_order**	**2.50e+01**	**4.23e−10**	**30**
	Percival	**5**	**3**	**first_order**	**2.50e+01**	**2.03e−08**	**153**
NASH	IPOPT	72	24	infeasible	4.91e−05	3.37e+01	103
	Percival	72	24	stalled	2.56e+02	1.91e+01	478
NET1	**IPOPT**	**48**	**57**	**first_order**	**9.41e+05**	**1.75e−13**	**126**
	Percival	**48**	**57**	**first_order**	**9.41e+05**	**1.44e−08**	**220**
NET2	**IPOPT**	**144**	**160**	**unknown**	**1.19e+06**	**6.69e−12**	**265**
	Percival	**144**	**160**	**first_order**	**1.19e+06**	**1.59e−09**	**301**
NET3	**IPOPT**	**464**	**521**	**unknown**	**5.86e+06**	**5.46e−11**	**720**
	Percival	464	521	max_time	5.68e+06	4.01e−01	1705
ODFITS	**IPOPT**	**10**	**6**	**first_order**	**−2.38e+03**	**1.14e−13**	**33**
	Percival	**10**	**6**	**first_order**	**−2.38e+03**	**2.36e−07**	**63**
OPTCNTRL	**IPOPT**	**32**	**20**	**first_order**	**5.50e+02**	**1.56e−15**	**103**
	Percival	**32**	**20**	**first_order**	**5.50e+02**	**1.61e−08**	**238**
OPTPRLOC	**IPOPT**	**30**	**30**	**first_order**	**−1.64e+01**	**0.00e+00**	**83**
	Percival	**30**	**30**	**first_order**	**−1.64e+01**	**2.18e−10**	**23317**
ORTHREGB	**IPOPT**	**27**	**6**	**first_order**	**4.52e−20**	**1.97e−10**	**9**
	Percival	**27**	**6**	**first_order**	**3.37e−20**	**1.69e−11**	**327**
PENTAGON	**IPOPT**	**6**	**15**	**first_order**	**1.37e−04**	**0.00e+00**	**51**
	Percival	**6**	**15**	**first_order**	**1.46e−04**	**4.06e−07**	**113**
POLAK1	**IPOPT**	**3**	**2**	**first_order**	**2.72e+00**	**0.00e+00**	**21**
	Percival	**3**	**2**	**first_order**	**2.72e+00**	**2.33e−09**	**476**
POLAK2	**IPOPT**	**11**	**2**	**first_order**	**5.46e+01**	**0.00e+00**	**57**
	Percival	**11**	**2**	**first_order**	**5.46e+01**	**4.74e−09**	**232**
POLAK3	IPOPT	12	10	max_iter	4.44e+03	4.92e+07	4598
	Percival	**12**	**10**	**first_order**	**5.93e+00**	**2.18e−08**	**1489**
POLAK4	**IPOPT**	**3**	**3**	**first_order**	**2.63e−07**	**0.00e+00**	**45**
	Percival	**3**	**3**	**max_iter**	**3.18e−16**	**1.26e−11**	**82102**
POLAK5	**IPOPT**	**3**	**2**	**first_order**	**5.00e+01**	**5.37e−07**	**63**
	Percival	**3**	**2**	**first_order**	**5.00e+01**	**9.15e−09**	**5780**
POLAK6	**IPOPT**	**5**	**4**	**first_order**	**−4.40e+01**	**0.00e+00**	**1989**
	Percival	**5**	**4**	**first_order**	**−4.40e+01**	**1.11e−08**	**6575**
PORTFL1	**IPOPT**	**12**	**1**	**first_order**	**2.05e−02**	**2.78e−17**	**27**
	Percival	**12**	**1**	**first_order**	**2.05e−02**	**2.69e−07**	**69**
PORTFL2	**IPOPT**	**12**	**1**	**first_order**	**2.97e−02**	**1.39e−16**	**24**
	Percival	**12**	**1**	**first_order**	**2.97e−02**	**2.38e−09**	**76**
PORTFL3	**IPOPT**	**12**	**1**	**first_order**	**3.28e−02**	**5.55e−17**	**27**

Continua na próxima página

name	solver	nvar	ncon	status	objective	primal_feas	evals
	Percival	**12**	**1**	**first_order**	**3.27e−02**	**7.68e−09**	**82**
PORTFL4	**IPOPT**	**12**	**1**	**first_order**	**2.63e−02**	**2.78e−17**	**24**
	Percival	**12**	**1**	**first_order**	**2.63e−02**	**6.72e−07**	**63**
PORTFL6	**IPOPT**	**12**	**1**	**first_order**	**2.58e−02**	**1.67e−16**	**24**
	Percival	**12**	**1**	**first_order**	**2.58e−02**	**7.04e−08**	**63**
PRIMAL1	**IPOPT**	**325**	**85**	**first_order**	**−3.50e−02**	**0.00e+00**	**51**
	Percival	**325**	**85**	**first_order**	**−3.50e−02**	**9.71e−09**	**178**
PRIMAL2	**IPOPT**	**649**	**96**	**first_order**	**−3.37e−02**	**0.00e+00**	**48**
	Percival	**649**	**96**	**first_order**	**−3.37e−02**	**2.44e−08**	**164**
PRIMAL3	**IPOPT**	**745**	**111**	**first_order**	**−1.36e−01**	**0.00e+00**	**36**
	Percival	**745**	**111**	**first_order**	**−1.36e−01**	**4.75e−10**	**457**
PRIMALC1	**IPOPT**	**230**	**9**	**first_order**	**−6.16e+03**	**0.00e+00**	**51**
	Percival	**230**	**9**	**max_iter**	**−6.16e+03**	**2.92e−10**	**67421**
PRIMALC2	**IPOPT**	**231**	**7**	**first_order**	**−3.55e+03**	**0.00e+00**	**54**
	Percival	**231**	**7**	**max_iter**	**−3.55e+03**	**7.57e−12**	**39022**
PRIMALC5	**IPOPT**	**287**	**8**	**first_order**	**−4.27e+02**	**0.00e+00**	**45**
	Percival	**287**	**8**	**first_order**	**−4.27e+02**	**2.23e−10**	**98**
PRIMALC8	**IPOPT**	**520**	**8**	**first_order**	**−1.83e+04**	**0.00e+00**	**75**
	Percival	520	8	max_time	6.22e+00	1.77e−04	6924
PRODPL0	**IPOPT**	**60**	**29**	**first_order**	**5.88e+01**	**8.88e−16**	**57**
	Percival	**60**	**29**	**first_order**	**5.88e+01**	**1.46e−07**	**7036**
PRODPL1	**IPOPT**	**60**	**29**	**first_order**	**3.57e+01**	**8.88e−16**	**90**
	Percival	**60**	**29**	**first_order**	**3.57e+01**	**4.45e−08**	**1875**
PT	**IPOPT**	**2**	**501**	**first_order**	**1.78e−01**	**0.00e+00**	**60**
	Percival	**2**	**501**	**first_order**	**1.78e−01**	**1.66e−07**	**232**
QC	**IPOPT**	**9**	**4**	**first_order**	**−9.57e+02**	**0.00e+00**	**100**
	Percival	**9**	**4**	**first_order**	**−9.51e+02**	**0.00e+00**	**80**
QCNEW	**IPOPT**	**9**	**3**	**first_order**	**−8.07e+02**	**0.00e+00**	**26**
	Percival	9	3	max_time	−8.72e+02	1.19e+00	1778641
QPCBLEND	**IPOPT**	**83**	**74**	**first_order**	**−7.84e−03**	**3.47e−18**	**60**
	Percival	**83**	**74**	**first_order**	**−7.84e−03**	**3.02e−11**	**39700**
QPCBOEI1	**IPOPT**	**384**	**351**	**first_order**	**1.15e+07**	**1.55e−14**	**403**
	Percival	384	351	max_time	1.71e+07	1.35e+02	10934
QPCBOEI2	**IPOPT**	**143**	**166**	**first_order**	**8.17e+06**	**4.00e−15**	**357**
	Percival	143	166	max_time	7.99e+06	1.26e+00	10674
QPCSTAIR	**IPOPT**	**467**	**356**	**first_order**	**6.20e+06**	**2.84e−14**	**581**
	Percival	**467**	**356**	**max_time**	**6.20e+06**	**4.74e−10**	**20470**
QPNBLEND	**IPOPT**	**83**	**74**	**first_order**	**−9.13e−03**	**6.94e−18**	**63**
	Percival	**83**	**74**	**first_order**	**−9.14e−03**	**1.05e−11**	**32075**
QPNBOEI1	**IPOPT**	**384**	**351**	**first_order**	**6.75e+06**	**1.47e−14**	**1721**
	Percival	384	351	max_time	9.87e+06	1.06e+02	6242
QPNBOEI2	**IPOPT**	**143**	**166**	**first_order**	**1.37e+06**	**2.31e−14**	**680**
	Percival	143	166	max_time	1.44e+06	8.96e−01	10011
QPNSTAIR	**IPOPT**	**467**	**356**	**first_order**	**5.15e+06**	**2.84e−14**	**671**
	Percival	**467**	**356**	**first_order**	**5.15e+06**	**2.90e−10**	**14631**
READING6	**IPOPT**	**102**	**50**	**first_order**	**−1.45e+02**	**2.36e−07**	**57**
	Percival	**102**	**50**	**first_order**	**−1.45e+02**	**6.33e−07**	**772**
RK23	**IPOPT**	**17**	**11**	**first_order**	**8.33e−02**	**3.11e−12**	**33**
	Percival	**17**	**11**	**first_order**	**8.33e−02**	**9.69e−09**	**373**
ROBOT	**IPOPT**	**14**	**2**	**first_order**	**6.59e+00**	**1.67e−16**	**62**
	Percival	**14**	**2**	**first_order**	**6.59e+00**	**3.06e−08**	**124**
ROSENMMX	**IPOPT**	**5**	**4**	**first_order**	**−4.40e+01**	**0.00e+00**	**60**
	Percival	**5**	**4**	**first_order**	**−4.40e+01**	**1.14e−09**	**705**
S268	**IPOPT**	**5**	**5**	**first_order**	**1.80e−07**	**0.00e+00**	**45**
	Percival	**5**	**5**	**first_order**	**3.64e−11**	**2.79e−08**	**38457**
S277-280	**IPOPT**	**4**	**4**	**first_order**	**5.08e+00**	**0.00e+00**	**39**
	Percival	**4**	**4**	**first_order**	**5.08e+00**	**1.39e−16**	**88**
S308NE	IPOPT	2	3	unknown	0.00e+00	∞	0
	Percival	2	3	infeasible	5.00e−01	8.79e−01	628
S316-322	**IPOPT**	**2**	**1**	**first_order**	**3.34e+02**	**1.11e−16**	**24**
	Percival	**2**	**1**	**first_order**	**3.34e+02**	**3.89e−07**	**87**
S365	IPOPT	7	5	unknown	6.00e+00	1.80e+01	6
	Percival	7	5	exception	∞	∞	0
S365MOD	IPOPT	7	5	unknown	6.00e+00	1.80e+01	6
	Percival	7	5	exception	∞	∞	0
SAWPATH	**IPOPT**	**583**	**774**	**small_step**	**7.50e+02**	**2.66e−15**	**84**
	Percival	583	774	stalled	1.15e+02	2.68e−01	298

Continua na próxima página

name	solver	nvar	ncon	status	objective	primal_feas	evals
SIMPLLPA	**IPOPT**	**2**	**2**	**first_order**	**1.00e+00**	**0.00e+00**	**27**
	Percival	**2**	**2**	**first_order**	**1.00e+00**	**0.00e+00**	**57**
SIMPLLPB	**IPOPT**	**2**	**3**	**first_order**	**1.10e+00**	**0.00e+00**	**33**
	Percival	**2**	**3**	**first_order**	**1.10e+00**	**2.22e−16**	**43**
SINROSNB	**IPOPT**	**1000**	**999**	**first_order**	**6.88e−08**	**0.00e+00**	**754**
	Percival	1000	999	first_order	3.18e+01	6.53e−10	688
SMBANK	**IPOPT**	**117**	**64**	**first_order**	**−7.13e+06**	**1.31e−10**	**54**
	Percival	**117**	**64**	**first_order**	**−7.13e+06**	**1.05e−08**	**2157**
SMMPSF	**IPOPT**	**720**	**263**	**first_order**	**1.03e+06**	**2.97e−13**	**1165**
	Percival	720	263	max_time	1.28e+06	1.50e−01	24971
SNAKE	**IPOPT**	**2**	**2**	**first_order**	**−1.96e−04**	**0.00e+00**	**37**
	Percival	2	2	max_time	−1.34e+05	9.47e+00	2292421
SPANHYD	**IPOPT**	**97**	**33**	**first_order**	**2.40e+02**	**1.02e−12**	**1223**
	Percival	**97**	**33**	**first_order**	**2.40e+02**	**5.67e−13**	**1999**
SPIN2OP	**IPOPT**	**102**	**100**	**first_order**	**2.83e−19**	**3.18e−13**	**357**
	Percival	**102**	**100**	**first_order**	**3.08e−28**	**3.18e−09**	**1413**
SPIRAL	**IPOPT**	**3**	**2**	**first_order**	**1.72e−07**	**0.00e+00**	**192**
	Percival	**3**	**2**	**first_order**	**9.95e−09**	**1.41e−08**	**509**
SSEBLIN	**IPOPT**	**194**	**72**	**first_order**	**1.62e+07**	**2.61e−11**	**441**
	Percival	**194**	**72**	**first_order**	**1.62e+07**	**1.75e−08**	**20048**
SSEBNLN	**IPOPT**	**194**	**96**	**max_iter**	**1.62e+07**	**2.73e−11**	**9015**
	Percival	**194**	**96**	**first_order**	**1.62e+07**	**6.34e−09**	**20819**
STANCMIN	**IPOPT**	**3**	**2**	**first_order**	**4.25e+00**	**0.00e+00**	**33**
	Percival	**3**	**2**	**first_order**	**4.25e+00**	**1.57e−09**	**92**
STATIC3	**IPOPT**	**434**	**96**	**unknown**	**−3.28e+44**	**4.44e−16**	**2946**
	Percival	434	96	max_time	−5.37e+29	7.67e+03	223394
STEENBRA	**IPOPT**	**432**	**108**	**first_order**	**1.70e+04**	**3.02e−13**	**72**
	Percival	432	108	max_time	1.90e+04	8.41e−03	447295
STEENBRB	**IPOPT**	**468**	**108**	**first_order**	**9.08e+03**	**3.04e−13**	**186**
	Percival	468	108	max_time	1.84e+04	2.71e−02	318432
STEENBRC	**IPOPT**	**540**	**126**	**first_order**	**2.75e+04**	**1.74e−12**	**386**
	Percival	540	126	max_time	6.58e+04	1.28e−01	213995
STEENBRD	**IPOPT**	**468**	**108**	**first_order**	**9.04e+03**	**2.27e−13**	**575**
	Percival	468	108	max_time	1.82e+04	4.51e−02	251654
STEENBRE	**IPOPT**	**540**	**126**	**first_order**	**2.89e+04**	**7.69e−13**	**256**
	Percival	540	126	max_time	7.04e+04	9.43e−02	264844
STEENBRF	**IPOPT**	**468**	**108**	**first_order**	**8.99e+03**	**1.92e−13**	**183**
	Percival	468	108	max_time	1.77e+04	2.09e−02	246400
STEENBRG	**IPOPT**	**540**	**126**	**first_order**	**2.74e+04**	**1.90e−12**	**640**
	Percival	540	126	max_time	6.83e+04	9.65e−02	243101
SUPERSIM	**IPOPT**	**2**	**2**	**first_order**	**6.67e−01**	**2.22e−16**	**6**
	Percival	**2**	**2**	**first_order**	**6.67e−01**	**2.48e−16**	**18**
SWOPF	**IPOPT**	**83**	**92**	**first_order**	**6.79e−02**	**8.51e−08**	**48**
	Percival	83	92	max_time	6.82e−02	2.59e−04	61538
SYNTHES1	**IPOPT**	**6**	**6**	**first_order**	**7.59e−01**	**0.00e+00**	**30**
	Percival	**6**	**6**	**first_order**	**7.59e−01**	**3.04e−07**	**77**
SYNTHES2	**IPOPT**	**11**	**14**	**first_order**	**−5.54e−01**	**0.00e+00**	**42**
	Percival	**11**	**14**	**first_order**	**−5.54e−01**	**3.34e−08**	**199**
SYNTHES3	**IPOPT**	**17**	**23**	**first_order**	**1.51e+01**	**5.55e−17**	**39**
	Percival	**17**	**23**	**first_order**	**1.51e+01**	**3.39e−07**	**208**
TABLE7	**IPOPT**	**624**	**230**	**first_order**	**5.96e+04**	**9.28e−11**	**60**
	Percival	**624**	**230**	**first_order**	**5.96e+04**	**3.96e−09**	**5236**
TAME	**IPOPT**	**2**	**1**	**first_order**	**0.00e+00**	**0.00e+00**	**18**
	Percival	**2**	**1**	**first_order**	**0.00e+00**	**0.00e+00**	**18**
TARGUS	**IPOPT**	**162**	**63**	**first_order**	**1.08e+03**	**8.83e−13**	**54**
	Percival	**162**	**63**	**first_order**	**1.08e+03**	**1.69e−09**	**2076**
TENBARS1	**IPOPT**	**18**	**9**	**first_order**	**2.30e+03**	**3.41e−13**	**95**
	Percival	**18**	**9**	**first_order**	**2.30e+03**	**4.85e−11**	**1502**
TENBARS2	IPOPT	18	8	first_order	2.30e+03	1.14e−13	96
	Percival	**18**	**8**	**first_order**	**2.28e+03**	**1.48e−10**	**1908**
TENBARS3	**IPOPT**	**18**	**8**	**first_order**	**2.25e+03**	**4.15e−13**	**51**
	Percival	**18**	**8**	**first_order**	**2.25e+03**	**8.94e−10**	**2096**
TENBARS4	**IPOPT**	**18**	**9**	**first_order**	**3.68e+02**	**5.53e−11**	**54**
	Percival	**18**	**9**	**first_order**	**3.68e+02**	**1.62e−10**	**29214**
TFI1	**IPOPT**	**3**	**101**	**first_order**	**5.33e+00**	**0.00e+00**	**170**
	Percival	**3**	**101**	**first_order**	**5.33e+00**	**9.48e−10**	**516**
TFI2	**IPOPT**	**3**	**101**	**first_order**	**6.49e−01**	**0.00e+00**	**43**

Continua na próxima página

name	solver	nvar	ncon	status	objective	primal_feas	evals
	Percival	**3**	**101**	**first_order**	**6.49e−01**	**3.76e−10**	**120**
TFI3	**IPOPT**	**3**	**101**	**first_order**	**4.30e+00**	**0.00e+00**	**48**
	Percival	**3**	**101**	**first_order**	**4.30e+00**	**9.44e−07**	**164**
TRIMLOSS	**IPOPT**	**142**	**75**	**first_order**	**9.06e+00**	**1.42e−14**	**96**
	Percival	142	75	max_time	9.20e+00	1.55e−02	24640
TRO11X3	IPOPT	150	61	max_iter	−2.41e+09	3.78e−01	17484
	Percival	150	61	max_time	−5.94e+03	1.51e+00	37996
TRO21X5	**IPOPT**	**540**	**201**	**unknown**	**−9.41e+14**	**7.51e−10**	**2731**
	Percival	540	201	max_time	−2.65e+02	1.60e+00	10019
TRO3X3	**IPOPT**	**30**	**13**	**first_order**	**9.00e+00**	**8.88e−16**	**49**
	Percival	30	13	max_time	−3.50e+04	5.58e−01	459382
TRO4X4	**IPOPT**	**63**	**25**	**first_order**	**9.00e+00**	**6.03e−15**	**68**
	Percival	63	25	max_time	−3.45e+12	1.00e+00	215086
TRO5X5	**IPOPT**	**108**	**41**	**first_order**	**9.00e+00**	**2.50e−15**	**80**
	Percival	108	41	max_time	−1.10e+11	1.42e+00	60098
TRO6X2	IPOPT	45	21	max_iter	−1.29e+18	1.93e+03	9051
	Percival	45	21	max_time	−4.86e+12	1.41e+00	331882
TRUSPYR1	**IPOPT**	**11**	**4**	**first_order**	**1.12e+01**	**1.01e−08**	**43**
	Percival	**11**	**4**	**first_order**	**1.12e+01**	**4.51e−08**	**1288**
TRUSPYR2	**IPOPT**	**11**	**11**	**first_order**	**1.12e+01**	**4.63e−09**	**39**
	Percival	**11**	**11**	**first_order**	**1.12e+01**	**1.31e−09**	**1507**
TRY-B	**IPOPT**	**2**	**1**	**first_order**	**2.07e−15**	**7.80e−13**	**60**
	Percival	2	1	first_order	1.00e+00	7.72e−08	63
TWIRISM1	**IPOPT**	**343**	**313**	**first_order**	**−1.01e+00**	**8.56e−09**	**177**
	Percival	343	313	max_time	−1.00e+00	3.90e−03	5506
TWOBARS	**IPOPT**	**2**	**2**	**first_order**	**1.51e+00**	**0.00e+00**	**30**
	Percival	**2**	**2**	**first_order**	**1.51e+00**	**2.07e−07**	**73**
WACHBIEG	IPOPT	3	2	infeasible	−1.00e+00	1.50e+00	44
	Percival	**3**	**2**	**first_order**	**1.00e+00**	**4.64e−07**	**71**
WATER	**IPOPT**	**31**	**10**	**first_order**	**1.05e+04**	**5.68e−14**	**78**
	Percival	**31**	**10**	**first_order**	**1.05e+04**	**1.91e−10**	**526**
WOMFLET	**IPOPT**	**3**	**3**	**first_order**	**6.05e+00**	**0.00e+00**	**36**
	Percival	**3**	**3**	**first_order**	**6.05e+00**	**1.48e−09**	**125**
YORKNET	IPOPT	312	256	unknown	1.44e+04	7.41e−13	634
	Percival	**312**	**256**	**max_time**	**1.39e+04**	**4.51e−07**	**9325**
ZAMB2-10	**IPOPT**	**270**	**96**	**first_order**	**−1.58e+00**	**1.02e−08**	**81**
	Percival	270	96	max_time	−1.44e+00	1.86e−02	125507
ZAMB2-11	**IPOPT**	**270**	**96**	**first_order**	**−1.12e+00**	**9.41e−08**	**63**
	Percival	270	96	max_time	−1.04e+00	1.40e−02	109055
ZAMB2-8	**IPOPT**	**138**	**48**	**first_order**	**−1.53e−01**	**2.47e−07**	**60**
	Percival	**138**	**48**	**first_order**	**−1.53e−01**	**2.41e−10**	**12280**
ZAMB2-9	**IPOPT**	**138**	**48**	**first_order**	**−3.55e−01**	**3.90e−09**	**63**
	Percival	138	48	max_time	−3.54e−01	1.56e−02	216701
ZECEVIC2	**IPOPT**	**2**	**2**	**first_order**	**−4.12e+00**	**0.00e+00**	**27**
	Percival	**2**	**2**	**first_order**	**−4.12e+00**	**9.49e−15**	**53**
ZECEVIC3	**IPOPT**	**2**	**2**	**first_order**	**9.73e+01**	**0.00e+00**	**66**
	Percival	**2**	**2**	**first_order**	**9.73e+01**	**4.88e−08**	**105**
ZECEVIC4	**IPOPT**	**2**	**2**	**first_order**	**7.56e+00**	**0.00e+00**	**30**
	Percival	**2**	**2**	**first_order**	**7.56e+00**	**8.39e−07**	**82**
ZY2	**IPOPT**	**3**	**2**	**first_order**	**2.00e+00**	**0.00e+00**	**30**
	Percival	**3**	**2**	**first_order**	**2.00e+00**	**7.86e−07**	**85**

Bibliografia

[1] D. Orban, S. Arreckx, A. Lambe e J. R. R. A. Martins, "A matrix-free augmented Lagrangian algorithm with application to large-scale structural design optimization", *Optimization and Engineering*, v. 17, pp. 359–384, 2015. DOI: 10.1007/s11081-015-9287-9.

[2] N. Gould, D. Orban e Ph. L. Toint, "Numerical methods for large-scale nonlinear optimization", *Acta Numerica*, v. 14, pp. 299 –361, mai. de 2005. DOI: 10.1017/S0962492904000248.

[3] J. Nocedal e S. J. Wright, *Numerical Optimization*, 2ª ed. New York: Springer, 2006.

[4] D. P. Bertsekas, *Constrained optimization and Lagrange multiplier methods*. New York: Academic Press, 1982.

[5] A. A. Ribeiro e E. W. Karas, *Otimização Contínua – Aspectos Teóricos e Computacionais*. Cengage Learning, 2013.

[6] A. Izmailov e M. Solodov, *Otimização: condições de otimalidade, elementos de análise convexa e de dualidade*, 3ª ed. IMPA, 2014, vol. 1.

[7] ——, *Otimização: métodos computacionais*, 2ª ed. IMPA, 2012, vol. 2.

[8] J. M. Martínez e S. A. Santos, *Métodos Computacionais de Otimização*. IMPA - Colóquio Brasileiro de Matemática, 1995.

[9] M. Nasri, L. C. Matioli, E. Silva Ferreira e A. Silveira, "Implementation of Augmented Lagrangian Methods for Equilibrium Problems", *Journal of Optimization Theory and Applications*, v. 168, dez. de 2015. DOI: 10.1007/s10957-015-0846-1.

[10] S. A. Santos, "Trust-region-based methods for nonlinear programming: recent advances and perspectives", *Pesquisa Operacional*, v. 34, pp. 447 –462, 2014. DOI: 10.1590/0101-7438.2014.034.03.0447.

[11] A. R. Conn, N. I. M. Gould e Ph. L. Toint, *Trust-region methods*. SIAM, 2000.

[12] J. J. Moré, *Trust regions and projected gradients*. Argonne National Laboratory, Mathematics and Computer Science Division, 1999. DOI: 10.1007/BFb0042769.

[13] C. J. Lin e J. J. Moré, "Newton's method for large bound-constrained optimization problems", *SIAM J. Optimization*, v. 9, pp. 1100–1127, 1999. DOI: 10.1137/S1052623498345075.

[14] J. V. Burke e J. J. Moré, "Exposing constraints", *SIAM J. Optimization*, v. 4, pp. 573–595, 1994. DOI: 10.1137/0804032.

[15] J. V. Burke, J. J. Moré e G. Toraldo, "Convergence properties of trust region methods for linear and convex constraints", *Mathematical Programming*, v. 47, pp. 305–336, 1990. DOI: 10.1007/BF01580867.

[16] P. H. Calamai e J. J. Moré, "Projected gradient methods for linearly constrained problems", *Mathematical Programming*, v. 39, pp. 93–116, 1987. DOI: 10.1007/BF02592073.

[17] D. Orban e A. S. Siqueira, "JuliaSmoothOptimizers: Infrastructure and Solvers for Continuous Optimization in Julia", 2019. DOI: 10.5281/zenodo.2655082. endereço: https://juliasmoothoptimizers.github.io.

[18] E. A. dos Santos e A. S. Siqueira, *Percival.jl: an augmented Lagrangian method*, https://github.com/JuliaSmoothOptimizers/Percival.jl, jul. de 2020. DOI: 10.5281/zenodo.3969045.

[19] N. Gould, D. Orban e Ph. L. Toint, "CUTEst: a Constrained and Unconstrained Testing Environment with safe threads for Mathematical Optimization", *Computational Optimization and Applications*, v. 60, jul. de 2014. DOI: 10.1007/s10589-014-9687-3.

[20] A. Wächter e L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming", *Mathematical Programming*, v. 106, pp. 25–57, 2006. DOI: 10.1007/s10107-004-0559-y.

[21] E. L. Lima, *Análise Real: Funções de uma Variável*, 12ª ed. IMPA, 2013, vol. 1.

[22] J. Stoer e C. Witzgall, *Convexity and Optimization in Finite Dimensions I*. Springer, 1970.

[23] J. A. Ramos, "Tópicos em condições de otimalidade para otimização não linear", tese de dout., Universidade de São Paulo, 2015.

[24] R. Andreani, E. G. Birgin, J. M. Martínez e M. L. Schuverdt, "Augmented Lagrangian methods under the constant positive linear dependence constraint qualification", *Mathematical Programming*, v. 111, pp. 5–32, 2008. DOI: 10.1007/s10107-006-0077-1.

[25] E. R. Torrealba, "Uma classe de métodos de Lagrangiano aumentado: convergência e testes numéricos", diss. de mestr., Universidade Federal do Paraná, 2014.

[26] J. M. Martínez, *Otimização prática usando o Lagrangiano Aumentado*. IMECC-UNICAMP, 2009.

[27] E. G. Birgin e J. M. Martínez, *Practical Augmented Lagrangian Methods for Constrained Optimization*. Philadelphia, PA: Society for Industrial e Applied Mathematics, 2014.

[28] A. R. Conn, N. I. M. Gould e Ph. L. Toint, "A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bounds", *SIAM J. Numerical Analysis*, v. 28, pp. 545–572, 1991. DOI: 10.1137/0728030.

[29] D. Orban, A. S. Siqueira e contributors, *LinearOperators.jl: A Julia Linear Operator Package*, https://github.com/JuliaSmoothOptimizers/LinearOperators.jl, 2019. DOI: 10.5281/zenodo.3747979.

[30] T. Steihaug, "The conjugate gradient method and trust regions in large scale optimization", *SIAM J. Numerical Analysis*, v. 20, pp. 626–637, 1983. DOI: 10.1137/0720042.

[31] A. Montoison, D. Orban e contributors, *Krylov.jl: A Julia Basket of Hand-Picked Krylov Methods*, https://github.com/JuliaSmoothOptimizers/Krylov.jl, 2020. DOI: 10.5281/zenodo.3548984.

[32] D. Orban e A. S. Siqueira, *CUTEst.jl: Julia's CUTEst Interface*, https://github.com/JuliaSmoothOptimizers/CUTEst.jl, 2020. DOI: 10.5281/zenodo.3925838.

[33] D. Orban, A. S. Siqueira e contributors, *NLPModelsIpopt.jl: A thin IPOPT wrapper for NLPModels*, https://github.com/JuliaSmoothOptimizers/NLPModelsIpopt.jl, 2020. DOI: 10.5281/zenodo.2629034.

[34] E. W. Karas, A. S. Santos e F. B. Svaiter, "Algebraic rules for quadratic regularization of Newton's method", *Computational Optimization and Applications*, v. 60, pp. 343–376, 2015. DOI: 10.1007/s10589-014-9671-y.

[35] E. D. Dolan e J. J. Moré, "Benchmarking optimization software with performance profiles", *Mathematical Programming*, v. 91, pp. 201–213, 2002. DOI: 10.1007/s101070100263.

[36] C. Gonzaga, E. W. Karas e D. Rossetto, "An Optimal Algorithm for Constrained Differentiable Convex Optimization", *Optimization-online preprint 3053*, v. 23, pp. 1–16, 2011. DOI: 10.1137/110836602.

[37] L. F. Prudente, "Inviabilidade em métodos de Lagrangiano aumentado", tese de dout., Universidade Estadual de Campinas, 2002.

[38] E. Birgin, J. M. Martínez e L. F. Prudente, "Optimality properties of an Augmented Lagrangian method on infeasible problems", *Computational Optimization and Applications*, v. 60, pp. 609–631, abr. de 2014. DOI: 10.1007/s10589-014-9685-5.

[39] D. Orban e A. S. Siqueira, *SolverBenchmark.jl: Benchmarking tools for solvers in Julia*, https://github.com/JuliaSmoothOptimizers/SolverBenchmark.jl, 2019. DOI: 10.5281/zenodo.3948381.

[40] E. A. dos Santos e A. S. Siqueira, *Percival-tests: implementation of numerical tests for the augmented Lagrangian solver Percival.jl*, https://github.com/CiDAMO/Percival-tests, set. de 2020. DOI: 10.5281/zenodo.4044696.

[41] Y. Bello-Cruz, O. Ferreira e L. Prudente, "On the global convergence of the inexact semi-smooth Newton method for absolute value equation", *Computational Optimization and Applications*, v. 65, ago. de 2015. DOI: 10.1007/s10589-016-9837-x.