

ELVIS RODRIGUEZ TORREALBA

UMA CLASSE DE MÉTODOS DE LAGRANGIANO
AUMENTADO: CONVERGÊNCIA E
TESTES NUMÉRICOS

ELVIS RODRIGUEZ TORREALBA

UMA CLASSE DE MÉTODOS DE LAGRANGIANO
AUMENTADO: CONVERGÊNCIA E
TESTES NUMÉRICOS

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Paraná, como requisito parcial à obtenção do grau de Mestre em Matemática.
Orientadora: Dr.^a Elizabeth Wegner Karas.
Coorientador: Dr. Romulo Castillo.

Curitiba
2014

T689c

Torrealba, Elvis Rodriguez

Uma classe de métodos de lagrangiano aumentado : convergência e testes numéricos / Elvis Rodriguez Torrealba. – Curitiba, 2014.
67f. : il. color. ; 30 cm.

Dissertação (mestrado) - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-graduação em Matemática, 2014.

Orientador: Elizabeth Wegner Karas -- Coorientador: Romulo Castillo.
Bibliografia: p. 65-67.

1. Lagrange, Funções de. 2. Otimização matemática. 3. Algoritmos computacionais. I. Universidade Federal do Paraná. II. Karas, Elizabeth Wegner III. Castillo, Romulo. IV. Título.

CDD: 519.6

TERMO DE APROVAÇÃO

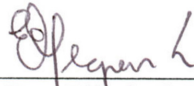
**“UMA CLASSE DE MÉTODOS DE LAGRANGIANO
AUMENTADO: CONVERGÊNCIA E TESTES
NUMÉRICOS”**

por

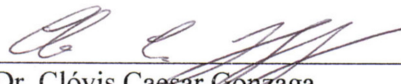
ELVIS MANUEL RODRIGUEZ TORREALBA

Dissertação aprovada como requisito parcial para obtenção do grau de
Mestre no Programa de Pós-Graduação em Matemática Aplicada,
pela Comissão Examinadora composta por:

Orientador:



Prof. Dra. Elizabeth Wegner Karas
Dep. de Matemática – UFPR



Prof. Dr. Clóvis César Gonzaga
Dep. de Matemática – UFSC



Prof. Dr. Roberto Andreani
IMECC - UNICAMP

Curitiba, 21 de fevereiro de 2014.



Ministério da Educação
Universidade Federal do Paraná
Setor de Ciências Exatas/Departamento de Matemática
Programa de Pós-Graduação em Matemática Aplicada - PPGMA

ATA DA 59ª DEFESA DE DISSERTAÇÃO DE MESTRADO

Aos vinte e um dias do mês de fevereiro de 2014, no Salão Antonina, do Hotel Nacional Inn, foi instalada pela Professora Elizabeth Wegner Karas, a Banca Examinadora para a quinquagésima nona Defesa de Dissertação de Mestrado em Matemática Aplicada. Estiveram presentes ao Ato, professores, alunos e visitantes.

A banca examinadora, homologada pelo Colegiado do Programa de Pós-Graduação em Matemática Aplicada, ficou constituída pelos professores: Prof. Dr. Clóvis Caesar Gonzaga, da Universidade Federal de Santa Catarina, Prof. Dr. Roberto Andreani, da Universidade de Campinas, e a Profa. Dra. Elizabeth Wegner Karas, orientadora da dissertação, a quem coube a presidência dos trabalhos.

Às quatorze horas, a banca iniciou seus trabalhos, convidando o candidato **ELVIS MANUEL RODRIGUEZ TORREALBA** a fazer a apresentação do tema da dissertação intitulada "UMA CLASSE DE MÉTODOS DE LAGRANGIANO AUMENTADO: CONVERGÊNCIA E TESTES NUMÉRICOS". Encerrada a apresentação, iniciou-se a fase de arguição pelos membros participantes. Após a arguição, a banca com pelo menos 03 (três) membros, reuniu-se para apreciação do desempenho do pós-graduando.

A banca considerou que o pós-graduando fez uma apresentação com a necessária concisão. A Dissertação apresenta contribuição à área de estudos e não foram registrados problemas fundamentais de estrutura e redação, resultando em plena e satisfatória compreensão dos objetivos pretendidos.

Tendo em vista a dissertação e a arguição, os membros presentes da banca decidiram pela sua aprovação.

Curitiba, 21 de fevereiro de 2014.

Profa. Dra. Elizabeth Wegner Karas
Presidente

Prof. Dr. Clóvis Caesar Gonzaga
Titular

Prof. Dr. Roberto Andreani
Titular



Ministério da Educação
Universidade Federal do Paraná
Setor de Ciências Exatas/Departamento de Matemática
Programa de Pós-Graduação em Matemática Aplicada - PPGMA

PARECER DA BANCA EXAMINADORA

Após a apresentação, a banca deliberou pela aprovação da dissertação do candidato **ELVIS MANUEL RODRIGUEZ TORREALBA** devendo, para tanto, incorporar as sugestões feitas pelos membros da banca, no prazo estabelecido pelo regimento correspondente.

Curitiba, 21 de fevereiro de 2014.

Profa. Dra. Elizabeth Wegner Karas
Presidente

Prof. Dr. Clóvis Caesar Gonzaga
Titular

Prof. Dr. Roberto Andreani
Titular

Aos meus avós e minha tia Petra.

Agradecimentos

A Deus pela vida, saúde, capacitação e força para continuar.

Aos meus pais Irma e Rafael, aos meus irmãos Toño, Ney e Nohe pela compreensão e apoio incondicional.

A Aura, meu amor, pelo carinho, pelos mimos e pelos abraços nos momentos difíceis.

Ao resto da minha família que tanto me incentivaram e apoiaram ao longo deste caminho, em especial, àqueles que perdi fisicamente porém que estarão sempre no meu coração.

Aos meus orientadores Elizabeth e Rómulo pela dedicação, tempo e esforço neste trabalho.

Aos professores Rómulo Castillo pelo incentivo e pela ajuda para vir ao Brasil, Carlos Durán pela ajuda quando cheguei no Brasil e ao professor Ademir Ribeiro pela tempo e pela ajuda sempre que precisei.

Aos membros da banca, professores Roberto Andreani e Clóvis Gonzaga, por aceitarem o convite e pelas correções para melhorar meu trabalho.

Aos meus amigos da Venezuela, do Brasil e aos não brasileiros que conheci aqui também. Em especial, Stela, Leonardo, Marcos, Felipe e Dion por me ajudarem sempre que precisei, e também a minha turma de mestrado por ficarmos juntos até o final.

Ao programa de Pós-Graduação em Matemática pela oportunidade e formação de excelência.

E finalmente, à CAPES pelo apoio financeiro.

“A arte de vencer se aprende nas derrotas.”

Simón Bolívar.

Resumo

Estudamos uma classe de métodos de Lagrangiano aumentado para problemas de minimização, num conjunto convexo e compacto, sujeito a restrições de desigualdade. Esta classe de métodos envolve uma função de penalização. Mostramos, sob certas hipóteses, a convergência global dos métodos desde que as funções de penalização envolvidas satisfaçam certas propriedades. Para os testes computacionais escolhemos três funções de penalização que satisfazem as propriedades exigidas, sendo uma delas a clássica função de penalização quadrática devida a Powell-Hestenes-Rockafellar. São apresentados resultados numéricos de comparação do desempenho computacional desta classe de métodos de Lagrangiano aumentado com as três funções de penalização, na resolução de problemas da coleção CUTEr.

Palavras-chave: Lagrangiano aumentado, função de penalização, convergência global, testes numéricos.

Abstract

We study a class of Augmented Lagrangian methods for solving inequality constrained problems. The class of methods involves penalty functions. Under reasonable hypotheses, we prove the global convergence of the methods since the penalty function satisfies some conditions. We discuss three penalty functions satisfying these conditions. One of them is the classical quadratic penalty function proposed by Powell-Hestenes-Rockafellar. Finally, we present some numerical experiments to compare the computational performance of the algorithm with these three penalty functions for solving problems of the **CUTEr** collection.

Palavras-chave: Augmented Lagrangian, penalty function, global convergence, numerical experiments.

Sumário

Introdução	1
1 Métodos de Penalização	3
1.1 Penalização Externa	3
1.1.1 Penalização externa exata	8
1.1.2 Penalização quadrática	18
1.2 Penalização interna	21
2 Métodos de Lagrangiano aumentado	26
2.1 Método clássico	26
2.1.1 Análise de convergência	28
2.1.2 Extensão para restrições de desigualdade	31
2.2 Método geral	32
2.2.1 Análise de convergência	34
2.2.2 Diferentes funções de penalização	38
3 Experimentos numéricos	42
3.1 Metodologia	42
3.1.1 Conjunto de problemas-teste	42
3.1.2 Resolução dos subproblemas	44
3.1.3 Critérios numéricos de convergência	44
3.1.4 Atualização do parâmetro de penalidade	45
3.1.5 Critério de seleção	46
3.2 Testes numéricos	47
3.2.1 Escolha dos Parâmetros	47
3.2.2 Resultados numéricos	48
Conclusão	50
Apêndices - Tabelas referentes aos testes numéricos	51
A Combinações dos parâmetros	51
B Resultados numéricos	55
Referências Bibliográficas	65

Introdução

Neste trabalho concentramos nossa atenção no seguinte problema

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & g(x) \leq 0 \\ & x \in \Omega, \end{array}$$

onde Ω é um subconjunto de \mathbb{R}^n geralmente simples. Muitas vezes Ω é o próprio \mathbb{R}^n e frequentemente uma caixa $\Omega = \{x \in \mathbb{R}^n \mid \ell \leq x \leq u\}$. As funções $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são continuamente diferenciáveis num conjunto aberto que contém Ω .

Este tipo de problema tem muitas aplicações práticas e existem muitos métodos para resolvê-lo. O método de Lagrangiano aumentado ou também conhecido como método dos multiplicadores é um deles. O método de Lagrangiano aumentado surge de maneira independente por Hestenes [22] e Powell [36] em 1969 e devido a sua eficiência computacional em problemas de grande porte, é muito utilizado para minimização de funções sujeitas a restrições mais gerais [1, 2]. Trabalhos recentes [9, 13, 14, 19, 21, 26] na área comprovam o grande interesse pelos métodos de Lagrangiano aumentado, tanto do ponto de vista teórico como computacional.

O método de Lagrangiano aumentado consiste em eliminar as restrições $g(x) \leq 0$ e incluí-las na função objetivo e assim gerar outro problema mais fácil de ser resolvido e que tenha soluções exatas ou aproximadas. Com esta estratégia é definido um algoritmo que gera uma sequência iterativa na qual, em cada iteração é resolvido de forma aproximada um problema de minimização com restrições simples e para esta solução encontrada o multiplicador de Lagrange e o parâmetro de penalidade são atualizados segundo a regra definida no algoritmo. Quando o parâmetro de penalidade é muito grande, encontrar a solução do subproblema pode ser uma tarefa difícil. É importante que os métodos de Lagrangiano aumentado tenham a propriedade de que o parâmetro de penalidade não cresça indefinidamente. Estratégias para evitar tal crescimento são consideradas, por exemplo em [9, 31], onde o parâmetro de penalidade fica limitado com certa frequência sob condições suficientes em problemas práticos. A vantagem sobre outros métodos de utilizar o método de Lagrangiano aumentado é que para resolver os subproblemas em cada iteração podem ser utilizados métodos que podem lidar com problemas que possuem um número grande de variáveis. Por exemplo, se no problema original, Ω é uma caixa, no momento de resolver os subproblemas no algoritmo de Lagrangiano aumentado uma boa ferramenta para executar este trabalho é o método **GENCAN** [10], que para resolver problemas de minimização com restrições de caixa, combina um procedimento de conjunto ativo com o método de gradiente espectral projetado [11, 12].

Nosso objetivo é estudar do ponto de vista teórico e computacional uma classe de métodos de Lagrangiano aumentado. Nossa principal referência foi o artigo [8] onde são comparados computacionalmente 65 métodos de Lagrangiano aumentado que diferem

entre si na função de penalização escolhida. Entre as funções de penalização testadas naquele artigo estão as propostas em [4, 5, 6, 7, 15, 24, 27, 28, 29, 32, 33, 35, 41, 42], além da função quadrática de Powell-Hestenes-Rockafellar (PHR), apresentada em [22, 36, 39], que foi a vencedora.

Nesta dissertação, estudamos uma classe de métodos de Lagrangiano aumentado cujas funções de penalização satisfazem certas propriedades as quais são verificadas por uma grande classe de funções que engloba a clássica penalização quadrática PHR. Discutimos sua análise de convergência global, sob hipóteses razoáveis.

Cabe ressaltar que a análise de convergência dos métodos de penalização apresentados nesta dissertação foi feita considerando a condição de independência linear dos gradientes das restrições ativas ou a condição de qualificação de Mangasarian-Fromovitz. No entanto, condições de qualificação mais fracas podem ser consideradas, como por exemplo, a condição de Dependência Linear Positiva Constante (CPLD) introduzida em [37] e discutida em [3]. A prova de convergência global do método de Lagrangiano aumentado com a penalização PHR, supondo a condição de qualificação CPLD é feita, por exemplo, em [1, 31].

Nesta dissertação, apresentamos também testes numéricos comparando o desempenho computacional da classe de métodos de Lagrangiano aumentado estudada com diferentes funções de penalização. Para os testes computacionais escolhemos duas funções de penalização sugeridas em [16] e que não haviam sido testadas em [8]. Mostramos que estas funções satisfazem as propriedades exigidas e que portanto a convergência global do método, com o uso das mesmas, está garantida. Fazemos, então testes computacionais, seguindo a metodologia de [8], para comparar o desempenho do método de Lagrangiano aumentado com estas duas funções de penalização e com a penalização quadrática clássica. Os testes foram feitos em **Fortran**, com base no algoritmo cedido pelos autores de [8], usando o conjunto de problema da coleção **CUTEr**.

O trabalho está organizado da seguinte forma: no Capítulo 1 é feita uma revisão dos métodos de penalização externa e interna. No Capítulo 2 apresentamos o método de Lagrangiano aumentado clássico e sua generalização para uma classe de métodos de Lagrangiano aumentado que envolve funções de penalização. A prova de convergência global é discutida desde que as funções de penalização satisfaçam certas propriedades. No Capítulo 2, ainda, exibimos três funções de penalização quadráticas que satisfazem as propriedades exigidas, sendo uma delas a clássica função de penalização devida a Powell-Hestenes-Rockafellar. No capítulo 3 são realizados testes numéricos envolvendo cada uma das três penalidades quadráticas, utilizando 84 problemas da coleção **CUTEr** que envolvem restrições de desigualdade e caixa em todas as variáveis. Nestes testes, o método de Lagrangiano aumentado com penalização quadrática clássica (PHR) foi o mais robusto e eficiente, corroborando as conclusões de [8].

Capítulo 1

Métodos de Penalização

Os métodos de penalização têm como objetivo aproximar um problema com restrições por uma sequência de problemas irrestritos. Neste tipo de métodos, a violação (ou risco de violação) das restrições é sancionada com o acréscimo de um termo adicionado na função objetivo de maneira que as funções que definem as restrições são eliminadas como tal e incluídas como um termo na função objetivo. Neste capítulo estudaremos os métodos de penalização externa e penalização interna. As principais referências utilizadas foram [25, 34].

1.1 Penalização Externa

Iniciamos estudando a penalização externa que consiste em adicionar à função objetivo um termo cujo valor é zero no conjunto viável e positivo para os não viáveis. O nome de penalização externa é dado porque em geral a solução do problema penalizado estará fora do conjunto viável.

O problema em forma geral a ser tratado com este tipo de penalidade é dado por

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && x \in D, \end{aligned} \tag{1.1}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $D \subset \mathbb{R}^n$ é um conjunto viável qualquer. Como dissemos, à função objetivo adicionamos um termo que deve penalizar a violação das restrições, isto é, uma função penalidade $P : \mathbb{R}^n \rightarrow \mathbb{R}_+$ contínua que satisfaz

$$P(x) \begin{cases} = 0 & \forall x \in D \\ > 0 & \forall x \notin D. \end{cases} \tag{1.2}$$

A seguir vemos um exemplo de uma função de penalização.

Exemplo 1.1

Sejam $h : \mathbb{R}^n \rightarrow \mathbb{R}^\ell$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $D = \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\}$. Defina inicialmente a função $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}^\ell \times \mathbb{R}^m$ como sendo

$$\Psi(x) = (h(x), \max\{0, g_1(x)\}, \dots, \max\{0, g_m(x)\}). \tag{1.3}$$

Então, dado $q > 0$, a função $P : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por

$$P(x) = \|\Psi(x)\|^q, \quad (1.4)$$

é uma função de penalização. Com efeito, se $x \in D$ temos que $h(x) = 0$ e $g(x) \leq 0$, então $\max\{0, g(x)\} = 0$ e portanto, $P(x) = \|\Psi(x)\|^q = \|0\|^q = 0$. Agora, se $x \notin D$ temos que $h(x) \neq 0$ ou alguma componente da função $g(x)$ é estritamente positiva. Assim, por propriedade das normas concluímos que $P(x) = \|\Psi(x)\|^q > 0$. \square

Como salientado em [25], na expressão (1.4), além da norma Euclidiana também podem ser consideradas as normas $\|\cdot\|_q$ ou $\|\cdot\|_\infty$ (é claro que para $q \in (0, 1)$, $\|\cdot\|_q$ não define uma norma). Assim, com estas escolhas temos que

$$P(x) = \|\Psi(x)\|_q^q = \|h(x)\|_q^q + \sum_{i=1}^m (\max\{g_i(x), 0\})^q, \quad (1.5)$$

ou

$$P(x) = \|\Psi(x)\|_\infty^q = \max\{\|h(x)\|_\infty, 0, g_1(x), g_2(x), \dots, g_m(x)\}^q. \quad (1.6)$$

Notemos que para h e g diferenciáveis, a função P definida por (1.5) é diferenciável quando $q > 1$, e não é diferenciável no caso contrário. Analogamente, a função P definida em (1.6) não é diferenciável, mesmo que h e g sejam diferenciáveis. Mais adiante estudaremos alguns casos particulares de funções de penalização diferenciável e não-diferenciável.

Consideremos uma família de problemas irrestritos dados por

$$\underset{x \in \mathbb{R}^n}{\text{minimizar}} \quad \Phi(x, \rho), \quad (1.7)$$

sendo $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por

$$\Phi(x, \rho) = f(x) + \rho P(x), \quad (1.8)$$

onde $\rho > 0$ é o parâmetro de penalidade e P é uma função que satisfaz (1.2). À medida em que o valor do parâmetro ρ aumenta, a violação das restrições torna-se cada vez mais cara. Podemos esperar então que as soluções dos problemas irrestritos em (1.7) se aproximem do conjunto viável para ρ suficientemente grande e que seus pontos de acumulação sejam soluções do problema (1.1). Desta forma podemos considerar o seguinte algoritmo base.

Algoritmo 1.1 - Penalização Externa

Escolha uma função $P : \mathbb{R}^n \rightarrow \mathbb{R}_+$ (penalidade externa para o conjunto D), $\rho^0 > 0$.

$k = 0$.

REPITA

DETERMINE $x^k = x(\rho^k)$, solução de (1.7), com $\rho = \rho^k$.

ESCOLHA $\rho^{k+1} > \rho^k$,

$k = k + 1$.

Observe que em cada iteração do Algoritmo 1.1, é resolvido um problema de minimização irrestrita para a função $\Phi(\cdot, \rho^k)$ e é natural utilizar o ponto da iteração anterior x^{k-1} como ponto inicial na resolução de cada subproblema.

Exemplo 1.2

Consideremos o problema dado por

$$\begin{aligned} &\text{minimizar} && x_1^2 - x_2 \\ &\text{sujeito a} && x_1 + x_2 - 1 = 0, \\ &&& x_1 \geq 0. \end{aligned}$$

Observe que o ponto $\bar{x} = (0, 1)$ é a única solução deste problema. Consideremos a função de penalidade definida no Exemplo 1.1 com $q = 2$. Assim, dado $\rho > 0$ o subproblema gerado pelo Algoritmo 1.1 consiste em minimizar a função

$$\Phi(x, \rho) = x_1^2 - x_2 + \rho(x_1 + x_2 - 1)^2 + \rho(\max\{0, -x_1\})^2.$$

Note que esta função é convexa e diferenciável. Logo o minimizador é caracterizado pela solução do sistema de equações

$$0 = \nabla\Phi(x, \rho) = \begin{pmatrix} 2x_1 + 2\rho(x_1 + x_2 - 1) - 2\rho \max\{0, -x_1\} \\ -1 + 2\rho(x_1 + x_2 - 1) \end{pmatrix},$$

de onde obtemos que

$$2x_1 + 1 - 2\rho \max\{0, -x_1\} = 0.$$

Observe que esta última equação não possui solução se $x_1 \geq 0$. Assim, resolvendo o sistema para $x_1 < 0$ temos que

$$x_1(\rho) = \frac{-1}{2(1 + \rho)}, \quad x_2(\rho) = 1 + \frac{1}{2\rho} + \frac{1}{2(1 + \rho)}.$$

É claro que para $\rho > 0$ o ponto $x(\rho) = (x_1(\rho), x_2(\rho))$ não é viável. Em particular, se $\rho \rightarrow \infty$ obtemos que $x(\rho) = (x_1(\rho), x_2(\rho)) \rightarrow (0, 1)$. Ou seja, para ρ suficientemente grande a solução do problema (1.5) se aproxima da solução do problema original. A Figura 1.1 mostra a trajetória $x(\rho)$ das soluções dos problemas penalizados.

Estudemos algumas das propriedades das funções de penalização externa e convergência do Algoritmo 1.1.

Proposição 1.3 [25, Prop. 4.3.1] *Seja $(x^k)_{k \in \mathbb{N}}$ uma sequência gerada pelo Algoritmo 1.1, onde cada x^k é solução global do problema (1.7) com parâmetro de penalidade $\rho = \rho^k$. Então para todo k , tem-se que*

$$\Phi(x^{k+1}, \rho^{k+1}) \geq \Phi(x^k, \rho^k) \tag{1.9}$$

$$P(x^{k+1}) \leq P(x^k) \tag{1.10}$$

$$f(x^{k+1}) \geq f(x^k). \tag{1.11}$$

Demonstração. Como x^k é solução global de (1.7) então para $\rho = \rho^k$ temos que

$$\begin{aligned} \Phi(x^k, \rho^k) &\leq \Phi(x^{k+1}, \rho^k) && (1.12) \\ &= f(x^{k+1}) + \rho^k P(x^{k+1}) \\ &\leq f(x^{k+1}) + \rho^{k+1} P(x^{k+1}) \\ &= \Phi(x^{k+1}, \rho^{k+1}) \end{aligned}$$

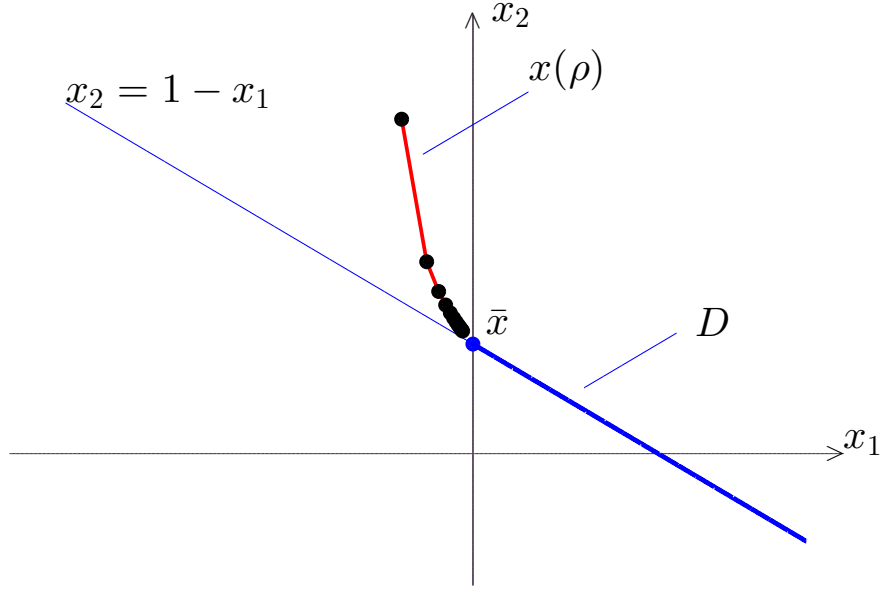


Figura 1.1: Trajetória $x(\rho)$ das soluções dos subproblemas

onde a desigualdade na terceira linha é válida pois $P \geq 0$ e $\rho^k < \rho^{k+1}$, mostrando assim (1.9). Agora para mostrar (1.10) observe que de (1.12) temos que

$$f(x^k) + \rho^k P(x^k) \leq f(x^{k+1}) + \rho^k P(x^{k+1}). \quad (1.13)$$

Por outro lado, x^{k+1} é solução global de (1.7) com $\rho = \rho^{k+1}$. Logo temos que $\Phi(x^{k+1}, \rho^{k+1}) \leq \Phi(x^k, \rho^{k+1})$, ou seja,

$$f(x^{k+1}) + \rho^{k+1} P(x^{k+1}) \leq f(x^k) + \rho^{k+1} P(x^k). \quad (1.14)$$

Assim, somando (1.13) e (1.14) tem-se que

$$(\rho^k - \rho^{k+1})P(x^k) \leq (\rho^k - \rho^{k+1})P(x^{k+1})$$

e usando o fato que $(\rho^k - \rho^{k+1}) < 0$, concluímos $P(x^k) \geq P(x^{k+1})$, o que prova (1.10).

Para finalizar, vejamos que de (1.10) e (1.12) temos que

$$f(x^k) + \rho^k P(x^k) \leq f(x^{k+1}) + \rho^k P(x^{k+1}) \leq f(x^{k+1}) + \rho^k P(x^k),$$

implicando que $f(x^k) + \rho^k P(x^k) \leq f(x^{k+1}) + \rho^k P(x^k)$ de onde obtemos (1.11). \square

O seguinte teorema prova a convergência assintótica do Algoritmo 1.1. Também mostra que em geral os pontos x^k gerados pelo Algoritmo 1.1 não são viáveis. Com efeito, se o método gera um ponto viável, ele já é solução do problema (1.1). É claro que tal situação não pode ser esperada num caso mais geral.

Teorema 1.4 (Convergência global do método de penalização externa) [25, Teo. 4.3.1]

Seja $(x^k)_{k \in \mathbb{N}}$ uma sequência gerada pelo Algoritmo 1.1 onde x^k é solução global do problema (1.7) com $\rho = \rho^k$. Então $x^k \in D$ se e somente se x^k é solução global do problema (1.1).

Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $P : \mathbb{R}^n \rightarrow \mathbb{R}_+$ funções contínuas no \mathbb{R}^n e $\bar{v} > -\infty$ o valor ótimo do problema (1.1). Se $\rho^k \rightarrow +\infty$, então todo ponto de acumulação da sequência $(x^k)_{k \in \mathbb{N}}$ é solução global do problema (1.1).

Demonstração. Observe que, se x^k é solução de (1.7) então para $x \in D$ vale que

$$\begin{aligned} f(x^k) &\leq f(x^k) + \rho^k P(x^k) \\ &= \Phi(x^k, \rho^k) \\ &\leq \Phi(x, \rho^k) \\ &= f(x), \end{aligned}$$

pois $\rho^k > 0$, $P(x^k) \geq 0$ e $P(x) = 0$ para todo $x \in D$. Assim, para todo k temos

$$f(x^k) \leq \Phi(x^k, \rho^k) \leq \inf_{x \in D} f(x) = \bar{v}. \quad (1.15)$$

Logo, se $x^k \in D$ temos que x^k é solução de (1.1) o que mostra a primeira parte do teorema.

Note que pela Proposição 1.3, a sequência de números reais $(\Phi(x^k, \rho^k))_{k \in \mathbb{N}}$ é não decrescente e por (1.15) temos que ela é limitada superiormente por \bar{v} e por conseguinte é convergente. Logo,

$$\lim_{k \in \mathbb{N}} \Phi(x^k, \rho^k) = \lim_{k \in \mathbb{N}} (f(x^k) + \rho^k P(x^k)) \leq \bar{v}. \quad (1.16)$$

Por outro lado, se \bar{x} é um ponto de acumulação da sequência $(x^k)_{k \in \mathbb{N}}$ então existe um conjunto $\mathbb{K} \subset \mathbb{N}$ tal que $(x^k)_{k \in \mathbb{K}}$ converge para \bar{x} . Observe que por (1.15) para cada k ocorre que

$$0 \leq P(x^k) \leq \frac{\bar{v} - f(x^k)}{\rho^k}.$$

Portanto tomando limite na desigualdade acima e usando a continuidade de f segue que

$$\lim_{k \in \mathbb{K}} P(x^k) = 0.$$

Finalmente, se utilizamos a continuidade de P e que $x^k \rightarrow \bar{x}$ obtemos que $P(\bar{x}) = 0$ e concluímos que $\bar{x} \in D$. Então pela primeira parte da prova \bar{x} é solução do problema (1.1). \square

Pelo teorema anterior, a hipótese de que o parâmetro de penalidade ρ creça de maneira ilimitada é fundamental, porém na hora de aplicar o método de penalização externa dada pelo algoritmo 1.1 este pode ter uma limitação prática devido à necessidade de resolver problemas cada vez pior condicionados. O seguinte exemplo mostra como isto ocorre.

Exemplo 1.5

Considere o problema

$$\begin{aligned} &\text{minimizar } \frac{1}{2}(x_1^2 + x_2^2) \\ &\text{sujeito a } x_1 - 1 = 0. \end{aligned} \quad (1.17)$$

Observe que a solução deste problema é dada pelo ponto $\bar{x} = (1, 0)$. Considere a função penalidade dada como no Exemplo 1.1, com $q = 2$. Logo o problema irrestrito a serem resolvido é dado por

$$\Phi(x, \rho) = \frac{1}{2}(x_1^2 + x_2^2) + \rho(x_1 - 1)^2. \quad (1.18)$$

Os pontos críticos de (1.18) são dados pela solução do sistema

$$0 = \nabla\Phi(x, \rho) = \begin{pmatrix} x_1 + 2\rho(x_1 - 1) \\ x_2 \end{pmatrix}$$

De onde obtemos que a solução deste sistema é dada por

$$x_1(\rho) = \frac{2\rho}{1+2\rho} = 1 - \frac{1}{1+2\rho} \quad \text{e} \quad x_2 = 0.$$

Observe que se $\rho \rightarrow \infty$ temos que o ponto $x(\rho) = (x_1(\rho), x_2) \rightarrow \bar{x}$. Embora isto tem uma complicação prática, pois se calculamos a matriz Hessiana do problema (1.18)

$$\nabla^2\Phi(x, \rho) = \begin{pmatrix} 1+2\rho & 0 \\ 0 & 1 \end{pmatrix},$$

temos que evidentemente para valores grandes do parâmetro de penalidade ρ , esta matriz ficará cada vez pior condicionada, o que significa que computacionalmente o termo penalizador absorve ao termo relativo à função objetivo do problema (1.17). \square

1.1.1 Penalização externa exata

Na seção anterior mostramos convergência para o Algoritmo 1.1. Observamos que conforme o parâmetro de penalidade aumenta a solução gerada pelo subproblema estarão-se aproximando mais da solução original do problema. Porém no Exemplo 1.5 pudimos observar que o fato de que o parâmetro de penalidade aumente pode gerar problemas de condicionamento na matriz hessiana nos subproblemas. Uma exceção importante para os métodos de penalidade externa é dada pelos métodos de penalização externa exata que embora o esquema algorítmico é o mesmo, estes métodos para um parâmetro penalizador suficientemente grande (porém fixo) permite a recuperação exata de uma solução do problema original. Tipicamente, isso é possível, porém ao custo da utilização de funções penalidade que geralmente não são diferenciáveis conforme veremos a seguir.

Nesta seção centraremos nosso interesse no seguinte problema

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & x \in D, \end{array} \quad (1.19)$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^\ell$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são funções dadas e

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

A função penalidade a serem considerada é dada por

$$P(x) = \sum_{i=1}^{\ell} |h_i(x)| + \sum_{i=1}^m \max\{g_i(x), 0\}. \quad (1.20)$$

Observemos que a função P assim definida não é diferenciável em D , mesmo quando h e g sejam diferenciáveis. Para efeito de alguns resultados no decorrer desta seção, definiremos para uma função $\xi : \mathbb{R}^n \rightarrow \mathbb{R}$ a derivada direcional no sentido de Gâteaux [17], em x na direção d , como a função $\xi'_+ : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por

$$\xi'_+(x, d) = \lim_{t \rightarrow 0^+} \frac{\xi(x + td) - \xi(x)}{t}$$

para $d \in \mathbb{R}^n$ e x no domínio da função ξ . Consideremos a família de problemas penalizados

$$\begin{aligned} & \text{minimizar} && \Phi(x, \rho) \\ & \text{sujeito a} && x \in \mathbb{R}^n, \end{aligned} \quad (1.21)$$

onde

$$\Phi(\cdot, \rho) : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \Phi(x, \rho) = f(x) + \rho P(x) \quad (1.22)$$

e $\rho > 0$ é o parâmetro penalizador.

O seguinte exemplo mostra que é possível obter solução do problema (1.19) a partir dos subproblemas (1.21) para um valor de ρ fixo.

Exemplo 1.6

Considere o problema

$$\begin{aligned} & \text{minimizar} && \frac{1}{2}(x_1 - x_2)^2 + \frac{1}{2}x_2^2 \\ & \text{sujeito a} && -x_1 \leq -1. \end{aligned}$$

Este problema tem solução única $\bar{x} = (1, \frac{1}{2})$. Logo aplicando a definição (1.21) temos que a função a serem minimizada sem restrições é dada por

$$\Phi(x, \rho) = \frac{1}{2}(x_1 - x_2)^2 + \frac{1}{2}x_2^2 + \rho \max\{0, 1 - x_1\}.$$

Como a função não é diferenciável então para achar a solução, temos que aproveitar alguma propriedade da função, neste caso temos que $\Phi(x, \rho)$ pode ser reescrita como

$$\Phi(x, \rho) = \left(\frac{x_1}{2} - x_2\right)^2 + \frac{1}{4}x_1^2 + \rho \max\{0, 1 - x_1\},$$

assim, como cada termo da soma é não negativo então temos que o ponto minimizador é dado pelo ponto $x_2 = \frac{x_1}{2}$ onde x_1 minimiza a função

$$\theta(x) = \begin{cases} \frac{1}{4}x^2 + \rho(1 - x) & \text{se } x < 1 \\ \frac{1}{4}x^2 & \text{se } x \geq 1 \end{cases}$$

Observe que a função $\theta(x)$ pode ser reescrita como

$$\theta(x) = \begin{cases} \frac{1}{4}(x - 2\rho)^2 + \rho - \rho^2 & \text{se } x < 1 \\ \frac{1}{4}x^2 & \text{se } x \geq 1 \end{cases}$$

Logo, se $\rho < \frac{1}{2}$ temos que o minimizador da função $\theta(x)$ é determinado pelo valor de ρ pois para valores de $x < 1$ o mínimo da parábola $\frac{1}{4}(x - 2\rho)^2 + \rho - \rho^2$ é dado pelo ponto $x = 2\rho < 1$ e porque a função $\frac{1}{4}x^2$ é sempre crescente para $x \geq 1$. Agora, se $\rho \geq \frac{1}{2}$ a função $\frac{1}{4}(x - 2\rho)^2 + \rho - \rho^2$ é sempre decrescente para $x < 1$ e como a função $\frac{1}{4}x^2$ é sempre crescente para $x \geq 1$ então o minimizador da função $\theta(x)$ será o ponto $x = 1$. Assim, se

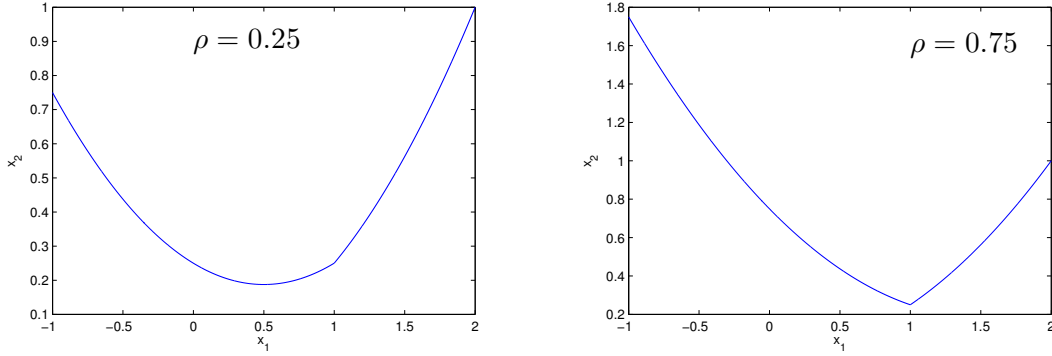


Figura 1.2: Variação do minimizador da função $\theta(x)$

consideramos $\rho \geq \frac{1}{2}$ qualquer garantimos que o ponto $x_1 = x = 1$ é minimizador de $\theta(x)$ e em consequência $x_2 = \frac{1}{2}$. O que implica que o ponto \bar{x} é minimizador da função de penalização Φ . A Figura 1.2 mostra como varia o minimizador da função $\theta(x)$ de acordo com o valor do parâmetro de penalidade ρ .

Devido que a função penalidade (1.21) não é diferenciável, o lema a seguir mostra uma caracterização para esta função penalidade. É importante mencionar que estes resultados são válidos para várias outras formas de penalização não-diferenciável. Por exemplo

$$P(x) = \max \{|h_1(x)|, \dots, |h_\ell(x)|, 0, g_1(x), \dots, g_m(x)\}. \quad (1.23)$$

Lema 1.7 *Sejam $h : \mathbb{R}^n \rightarrow \mathbb{R}^\ell$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ funções diferenciáveis no ponto $x \in \mathbb{R}^n$ e P definida por (1.20). Então para todo $x \in \mathbb{R}^n$ e $d \in \mathbb{R}^n$, $P'_+(x, d)$ existe e*

$$\begin{aligned} P'_+(x, d) = & \sum_{i \in J^+(x)} \nabla h_i(x)^T d + \sum_{i \in J^0(x)} |\nabla h_i(x)^T d| - \sum_{i \in J^-(x)} \nabla h_i(x)^T d \\ & + \sum_{i \in I^+(x)} \nabla g_i(x)^T d + \sum_{i \in I^0(x)} \max \{0, \nabla g_i(x)^T d\} \end{aligned} \quad (1.24)$$

onde

$$\begin{aligned} J^+(x) &= \{i = 1, \dots, \ell \mid h_i(x) > 0\}, \\ J^0(x) &= \{i = 1, \dots, \ell \mid h_i(x) = 0\}, \\ J^-(x) &= \{i = 1, \dots, \ell \mid h_i(x) < 0\}, \\ I^+(x) &= \{i = 1, \dots, m \mid g_i(x) > 0\}, \\ I^0(x) &= \{i = 1, \dots, m \mid g_i(x) = 0\}. \end{aligned}$$

Em particular, se $x \in D$, tem-se que

$$P'_+(x, d) = \sum_{i=1}^{\ell} |\nabla h_i(x)^T d| + \sum_{i \in I^0(x)} \max \{0, \nabla g_i(x)^T d\}. \quad (1.25)$$

Demonstração. Considere $\bar{x} \in \mathbb{R}^n$, $d \in \mathbb{R}^n$ e $i \in \{1, \dots, \ell\}$. Caso em que $h_i(\bar{x}) > 0$, por continuidade $h_i(x) > 0$ numa vizinhança V de \bar{x} . Portanto, $|h_i(x)| = h_i(x)$ para todo

$x \in V$. Assim, para t suficientemente pequeno

$$\frac{|h_i(\bar{x} + td)| - |h_i(\bar{x})|}{t} = \frac{h_i(\bar{x} + td) - h_i(\bar{x})}{t}. \quad (1.26)$$

Caso em que $h_i(\bar{x}) < 0$, de modo análogo, concluímos que para t suficientemente pequeno

$$\frac{|h_i(\bar{x} + td)| - |h_i(\bar{x})|}{t} = - \left(\frac{h_i(\bar{x} + td) - h_i(\bar{x})}{t} \right). \quad (1.27)$$

Caso em que $h_i(\bar{x}) = 0$, para $t > 0$

$$\frac{|h_i(\bar{x} + td)| - |h_i(\bar{x})|}{t} = \left| \frac{h_i(\bar{x} + td) - h_i(\bar{x})}{t} \right|. \quad (1.28)$$

Agora seja $i \in \{1, \dots, m\}$. Caso $g_i(\bar{x}) > 0$, por continuidade $g_i(x) > 0$ numa vizinhança V de \bar{x} . Portanto $\max\{0, g_i(x)\} = g_i(x)$ para todo $x \in V$. Assim, para t suficientemente pequeno

$$\frac{\max\{0, g_i(\bar{x} + td)\} - \max\{0, g_i(\bar{x})\}}{t} = \frac{g_i(\bar{x} + td) - g_i(\bar{x})}{t}. \quad (1.29)$$

Caso $g_i(\bar{x}) < 0$, por continuidade $g_i(x) < 0$, numa vizinhança de \bar{x} . Portanto $\max\{0, g_i(x)\} = 0$ para todo $x \in V$. Assim, para t suficientemente pequeno

$$\frac{\max\{0, g_i(\bar{x} + td)\} - \max\{0, g_i(\bar{x})\}}{t} = 0. \quad (1.30)$$

Caso $g_i(\bar{x}) = 0$, para $t > 0$

$$\begin{aligned} \frac{\max\{0, g_i(\bar{x} + td)\} - \max\{0, g_i(\bar{x})\}}{t} &= \frac{\max\{0, g_i(\bar{x} + td)\}}{t} \\ &= \max \left\{ 0, \frac{g_i(\bar{x} + td)}{t} \right\} \\ &= \max \left\{ 0, \frac{g_i(\bar{x} + td) - g_i(\bar{x})}{t} \right\}. \end{aligned} \quad (1.31)$$

Logo, passando limite quando $t \rightarrow 0^+$ em (1.26)-(1.31) e levando em conta a diferenciabilidade de h, g na definição de $P'_+(x, d)$ obtemos que a igualdade (1.24) é satisfeita. Para a segunda parte, basta observar que se $x \in D$ então $h_i(x) = 0$ para todo $i = 1, \dots, \ell$ e $g_i(x) \leq 0$ para todo $i = 1, \dots, m$. Assim, tomado limite quando $t \rightarrow 0^+$ em (1.28), (1.30) e (1.31) temos que a igualdade (1.25) também é satisfeita. \square

Daqui em diante o conjunto $I^0(x) = \{i \mid g_i(x) = 0\}$, será dito conjunto dos índices das restrições ativas em x . O seguinte teorema mostra as condições necessárias para que a penalização seja exata.

Teorema 1.8 [25, Teo. 4.3.4] *Seja $\bar{x} \in \mathbb{R}^n$ uma solução local (estrita) do problema penalizado (1.21).*

Se $\bar{x} \in D$ então \bar{x} é solução local (estrita) do problema original (1.19).

Se além da hipótese que $\bar{x} \in D$, as funções f, h e g são diferenciáveis em \bar{x} , então existe um par de multiplicadores de Lagrange $(\nu, \mu) \in \mathbb{R}^\ell \times \mathbb{R}^m$ tal que as condições de otimalidade de Karush-Kuhn-Tucker são satisfeitas no ponto \bar{x} .

Se além das hipóteses acima, \bar{x} satisfaz que o conjunto

$$\{\nabla h_i(\bar{x}) \mid i = 1, \dots, \ell\} \cup \{\nabla g_i(\bar{x}) \mid i \in I^0(\bar{x})\}$$

é linearmente independente, então $\rho \geq \|(\nu, \mu)\|_\infty$.

Demonstração. Como \bar{x} é um minimizador local do problema (1.21) então existe uma vizinhança V de \bar{x} tal que

$$\Phi(\bar{x}, \rho) \leq \Phi(x, \rho),$$

para todo $x \in V$. Agora, como $\bar{x} \in D$ então $\Phi(x, \rho) = f(x)$ para todo $x \in D$, assim obtemos

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

para todo $x \in V \cap D$, isto é, \bar{x} é minimizador do problema (1.19). Observe que se \bar{x} for uma solução local estrita de (1.21), todas as desigualdades acima são estritas e portanto \bar{x} é solução local estrita de (1.19).

Suponhamos agora que f, g e h são diferenciáveis em \bar{x} . Logo pelo Lema (1.7), $\Phi'_+(\cdot, \rho)$ existe em \bar{x} para cada $d \in \mathbb{R}^n$. Assim, para todo $t > 0$ suficientemente pequeno

$$0 \leq \Phi(\bar{x} + td, \rho) - \Phi(\bar{x}, \rho) = t\Phi'_+(\bar{x}, d, \rho) + o(t),$$

onde $\Phi'_+(\bar{x}, d, \rho) = \nabla f(x)^T d + \rho P'_+(x, d)$. Logo dividindo os dois lados da relação acima por $t > 0$ e passando limite quando $t \rightarrow 0^+$, obtemos que

$$\Phi'_+(\bar{x}, d, \rho) \geq 0$$

para todo $d \in \mathbb{R}^n$. Como $\bar{x} \in D$, de (1.25) obtemos que para todo $d \in \mathbb{R}^n$

$$0 \leq \nabla f(\bar{x})^T d + \rho \sum_{i=1}^{\ell} |\nabla h_i(\bar{x})^T d| + \rho \sum_{i \in I^0(\bar{x})} \max\{0, \nabla g_i(\bar{x})^T d\}. \quad (1.32)$$

Assim, em particular se consideramos $d \in \mathbb{R}^n$ tal que $\nabla h_i(\bar{x})^T d = 0$ para cada $i = 1, \dots, \ell$ e $\nabla g_i(\bar{x})^T d \leq 0$ para cada $i \in I^0(\bar{x})$ tem-se que

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

Logo, $d = 0$ é solução do problema de programação linear

$$\begin{aligned} \text{minimizar} \quad & \nabla f(x)^T d \\ \text{sujeito a} \quad & \nabla h_i(x)^T d = 0, \quad i = 1, \dots, \ell \\ & \nabla g_i(x)^T d \leq 0, \quad i \in I^0(x). \end{aligned} \quad (1.33)$$

Como este problema tem restrições lineares então pela condição de otimalidade (KKT) para o problema (1.33), existem $\nu \in \mathbb{R}^{\ell}$ e $\mu \in \mathbb{R}_+^{|I^0(\bar{x})|}$ (onde $|I^0(\bar{x})|$ é a cardinalidade do conjunto das restrições ativas) tais que

$$\nabla f(\bar{x}) + \sum_{i=1}^{\ell} \nu_i \nabla h_i(\bar{x}) + \sum_{i \in I^0(\bar{x})} \mu_i \nabla g_i(\bar{x}) = 0. \quad (1.34)$$

Agora, como $\bar{x} \in D$ então definindo $\mu_i = 0$ para $i \notin I^0(\bar{x})$, de (1.34) obtemos que \bar{x} satisfaz *KKT* para o problema (1.19), com multiplicadores $(\nu, \mu) \in \mathbb{R}^{\ell} \times \mathbb{R}^m$.

Finalmente, suponhamos que no ponto \bar{x} o conjunto

$$\{\nabla h_i(\bar{x}) \mid i = 1, \dots, \ell\} \cup \{\nabla g_i(\bar{x}) \mid i \in I^0(\bar{x})\}$$

é linearmente independente. De (1.32) e (1.34) obtemos que

$$\begin{aligned}
& \sum_{i=1}^{\ell} \nu_i (\nabla h_i(\bar{x}))^T d + \sum_{i \in I^0(\bar{x})} \mu_i \nabla(g_i(\bar{x}))^T d \\
& \leq \rho \sum_{i=1}^{\ell} |(\nabla h_i(\bar{x}))^T d| + \rho \sum_{i \in I^0(\bar{x})} \max\{0, (\nabla g_i(\bar{x}))^T d\}. \tag{1.35}
\end{aligned}$$

para todo $d \in \mathbb{R}^n$. Observe que dado $(u, v_{I^0}) \in \mathbb{R}^{\ell} \times \mathbb{R}^{|I^0|}$ o sistema

$$\begin{pmatrix} \nabla h(\bar{x}) \\ \nabla g_{I^0}(\bar{x}) \end{pmatrix} d = \begin{pmatrix} u \\ v_{I^0} \end{pmatrix},$$

tem sempre solução pois o conjunto $\{\nabla h_i(x) \mid i = 1, \dots, \ell\} \cup \{\nabla g_i(x) \mid i \in I^0(x)\}$ é linearmente independente. Assim, dado $(u, v_{I^0}) \in \mathbb{R}^{\ell} \times \mathbb{R}^{|I^0|}$ existe $d \in \mathbb{R}^n$ tal que

$$\nabla h_i(\bar{x})^T d = u_i, i = 1, \dots, \ell, \quad \nabla g_i(\bar{x})^T d = v_i, i \in I^0.$$

Agora, tomando $(u, v_{I^0}) \in \mathbb{R}^{\ell} \times \mathbb{R}^{|I^0|}$ arbitrário, de (1.35) obtemos que

$$\begin{aligned}
(\nu, \mu_{I^0})^T (u, v_{I^0}) &= \sum_{i=1}^{\ell} \nu_i u_i + \sum_{i \in I} \mu_i v_i = \sum_{i=1}^{\ell} \nu_i \nabla h_i(\bar{x})^T d + \sum_{i \in I^0} \mu_i \nabla g_i(\bar{x})^T d \\
&\leq \rho \sum_{i=1}^{\ell} |\nabla h_i(\bar{x})^T d| + \rho \sum_{i \in I} \max\{0, \nabla g_i(\bar{x})^T d\} \\
&= \rho \left(\sum_{i=1}^{\ell} |u_i| + \sum_{i \in I} \max\{0, v_i\} \right) \\
&\leq \rho \left(\|u\|_1 + \sum_{i \in I} |v_i| \right) \\
&= \rho \|(u, v_{I^0})\|_1,
\end{aligned}$$

onde também utilizamos o fato de que $\max\{0, t\} \leq |t|$ para todo $t \in \mathbb{R}$.

Pela definição da norma dual, temos que

$$\begin{aligned}
\|(\nu, \mu_{I^0})\|_{\infty} &= \sup \{ (\nu, \mu_{I^0})^T (u, v_{I^0}) \mid \|(u, v_{I^0})\|_1 = 1 \} \\
&= \sup \left\{ ((\nu, \mu_{I^0}))^T \frac{(u, v_{I^0})}{\|(u, v_{I^0})\|_1} \mid (u, v_{I^0}) \in \mathbb{R}^{\ell} \times \mathbb{R}^{|I^0|} \right\} \\
&\leq \rho.
\end{aligned}$$

Agora devido a que $\mu_i = 0$ para todo $i \notin I^0(\bar{x})$ então temos que $\|(\nu, \mu_{I^0})\|_{\infty} = \|(\nu, \mu)\|_{\infty}$, de onde concluímos a última afirmação do teorema. \square

Na segunda parte do teorema acima, a existência dos multiplicadores não é única. No entanto, ao acrescentarmos a hipótese de independência linear, temos a garantia da unicidade dos multiplicadores [25, Teo. 1.2.3].

Passaremos agora a estudar as condições que são suficientes para que uma penalização seja exata, isto é, condições que garantem que uma solução do problema penalizado (1.21) também seja uma solução do problema original (1.19).

Primeiro lembremos que a Lagrangiana do problema (1.19) é dada por

$$\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^\ell \times \mathbb{R}^m \rightarrow \mathbb{R},$$

$$\mathcal{L}(x, \nu, \mu) = f(x) + \nu^T h(x) + \mu^T g(x).$$

Mostraremos que, sob certas condições, o valor da Lagrangiana fornece uma cota inferior para o valor da função penalizada.

Proposição 1.9 [25, prop. 4.3.2] *Para todo $x \in \mathbb{R}^n$ e todo $(\nu, \mu) \in \mathbb{R}^\ell \times \mathbb{R}_+^m$, tem-se que*

$$\mathcal{L}(x, \nu, \mu) \leq f(x) + \|(\nu, \mu)\|_\infty P(x), \quad (1.36)$$

onde P é dada por (1.20). Em particular, se $\rho \geq \|(\nu, \mu)\|_\infty$, tem-se que

$$\mathcal{L}(x, \nu, \mu) \leq \Phi(x, \rho), \quad (1.37)$$

onde Φ é dada por (1.22).

Demonstração. Para todo $x \in \mathbb{R}^n$, tem-se que

$$\begin{aligned} \nu^T h(x) + \mu^T g(x) &= \sum_{i=1}^{\ell} \nu_i h_i(x) + \sum_{i=1}^m \mu_i g_i(x) \\ &\leq \|\nu\|_\infty \|h(x)\|_1 + \sum_{i=1}^m \mu_i \max\{0, g_i(x)\} \\ &\leq \|\nu\|_\infty \|h(x)\|_1 + \|\mu\|_\infty \sum_{i=1}^m \max\{0, g_i(x)\} \\ &\leq \|(\nu, \mu)\|_\infty P(x), \end{aligned} \quad (1.38)$$

onde utilizamos a desigualdade de Cauchy-Schwarz generalizada, o fato de que $\mu \geq 0$ e que $\|\nu\|_\infty \leq \|(\nu, \mu)\|_\infty$ e $\|\mu\|_\infty \leq \|(\nu, \mu)\|_\infty$. Agora somando $f(x)$ a ambos lados de (1.38) obtemos $\mathcal{L}(x, \nu, \mu) \leq f(x) + \|(\nu, \mu)\|_\infty P(x)$ como queríamos. A relação (1.37) é uma consequência imediata de (1.36), se utilizamos $\|(\nu, \mu)\|_\infty \leq \rho$ em (1.38). \square

O seguinte teorema mostra condições suficientes para que uma penalização seja exata no caso convexo. Notemos que neste caso, a função $\Phi(\cdot, \rho)$ também é convexa.

Teorema 1.10 (Penalização exata no caso convexo) [25, Teo. 4.3.5] *Sejam f e g funções convexas, diferenciáveis no ponto $\bar{x} \in \mathbb{R}^n$, e seja h uma função afim. Suponhamos que existam os multiplicadores de Lagrange $(\nu, \mu) \in \mathbb{R}^\ell \times \mathbb{R}^m$ tais que as condições de otimalidade de Karush-Kuhn-Tucker são satisfeitas (logo, \bar{x} é uma solução global do problema (1.19)).*

Então para todo $\rho \geq \|(\nu, \mu)\|_\infty$, o ponto \bar{x} é uma solução (global) do problema (1.21).

Demonstração. Como h é uma função afim, $\mathcal{L}(\cdot, \nu, \mu)$ é uma função convexa e $\nabla \mathcal{L}(\bar{x}, \nu, \mu) = 0$, (onde $\nabla \mathcal{L}$ denota o gradiente respeito à primeira variável) assim o ponto \bar{x} é um minimizador global de $\mathcal{L}(\cdot, \nu, \mu)$ no \mathbb{R}^n , ou seja,

$$\mathcal{L}(\bar{x}, \nu, \mu) \leq \mathcal{L}(x, \nu, \mu)$$

para todo $x \in \mathbb{R}^n$. Como \bar{x} satisfaz KKT então pela viabilidade e complementaridade em \bar{x} vale que $P(\bar{x}) = 0$, $h(\bar{x}) = 0$ e $\bar{\mu}^T g(\bar{x}) = 0$ logo, para $\rho \geq 0$ qualquer tem-se que

$$\begin{aligned} \Phi(\bar{x}, \rho) &= f(\bar{x}) + \rho P(\bar{x}) \\ &= f(\bar{x}) \\ &= f(\bar{x}) + \nu^T h(\bar{x}) + \mu^T g(\bar{x}) \\ &= \mathcal{L}(\bar{x}, \nu, \mu) \\ &\leq \mathcal{L}(x, \nu, \mu) \end{aligned} \tag{1.39}$$

para todo $x \in \mathbb{R}^n$. Por outro lado, como $\mu \geq 0$ e $\rho \geq \|(\nu, \mu)\|_\infty$, de (1.37) temos que

$$\mathcal{L}(x, \nu, \mu) \leq \Phi(x, \rho)$$

para todo $x \in \mathbb{R}^n$. Combinando esta desigualdade com (1.39), temos

$$\Phi(\bar{x}, \rho) \leq \Phi(x, \rho)$$

para todo $x \in \mathbb{R}^n$, o implica que \bar{x} é minimizador global de $\Phi(\cdot, \rho)$ no \mathbb{R}^n . \square

Consideremos, finalmente, o caso não convexo, mas em que uma solução satisfaz a condição suficiente de segunda ordem (portanto ela é estrita). Precisaremos inicialmente do seguinte lema.

Lema 1.11 *Sejam $h : \mathbb{R}^n \rightarrow \mathbb{R}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}$ funções duas vezes diferenciáveis em $\bar{x} \in \mathbb{R}^n$. Então a função definida como $\Gamma(x) = |h(x)| + \max\{0, g(x)\}$ é Lipschitz-contínua numa vizinhança V de \bar{x} .*

Demonstração. Como soma de funções localmente Lipschitz é localmente Lipschitz então basta mostrar que as funções $|h(x)|$ e $\max\{0, g(x)\}$ são localmente Lipschitz. Como $h(x)$ é Lipschitz em V (Pois h é duas vezes diferenciável em \bar{x}), digamos com constante L_1 então para $x, y \in V$,

$$\left| |h(x)| - |h(y)| \right| \leq |h(x) - h(y)| \leq L_1 \|x - y\| \tag{1.40}$$

assim, segue que $|h(x)|$ é localmente Lipschitz em V . Por outro lado, considerando que g é Lipschitz em V com constante digamos L_2 então para a função $\max\{0, g(x)\}$ consideremos os seguintes casos, com $x, y \in V$.

Caso que $g(x) \geq 0$ e $g(y) \geq 0$ temos que

$$\max\{0, g(x)\} = g(x) \quad \text{e} \quad \max\{0, g(y)\} = g(y)$$

em consequência

$$\left| \max\{0, g(x)\} - \max\{0, g(y)\} \right| = |g(x) - g(y)| \leq L_2 \|x - y\|. \tag{1.41}$$

Caso que $g(x) < 0$ e $g(y) < 0$ obtemos que

$$\left| \max\{0, g(x)\} - \max\{0, g(y)\} \right| = 0 \leq L_2 \|x - y\|. \tag{1.42}$$

Caso que $g(x) > 0$ e $g(y) \leq 0$ temos que

$$\left| \max\{0, g(x)\} - \max\{0, g(y)\} \right| = |g(x)| \leq |g(x) - g(y)| \leq L_2 \|x - y\|, \tag{1.43}$$

Considerando na primeira desigualdade que $-g(x) \geq 0$. Portanto de (1.41)-(1.43) temos que a função $\max\{0, g(x)\}$ é localmente Lipschitz, o que finaliza a prova. \square

Teorema 1.12 (Penalização exata no caso da condição suficiente de segunda ordem) [25, Teo. 4.3.6] *Sejam f , h e g funções duas vezes diferenciáveis em $\bar{x} \in \mathbb{R}^n$. Suponhamos também que existam os multiplicadores de Lagrange $(\nu, \mu) \in \mathbb{R}^\ell \times \mathbb{R}^m$ tais que a condição suficiente de segunda ordem seja satisfeita (logo, \bar{x} é uma solução local estrita do problema (1.19)).*

Então para todo $\rho > \|(\nu, \mu)\|_\infty$, o ponto \bar{x} é uma solução local estrita do problema (1.21).

Demonstração. Suponhamos que \bar{x} não seja um minimizador local estrito do problema (1.21), isto é, que exista uma sequência $(x^k)_{k \in \mathbb{N}} \subset \mathbb{R}^n \setminus \{\bar{x}\}$ tal que $x^k \rightarrow \bar{x}$ e para todo k ,

$$\Phi(x^k, \rho) \leq \Phi(\bar{x}, \rho). \quad (1.44)$$

Consideremos a sequência $d^k = \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|}$, como d^k é limitada então existe um conjunto $\mathbb{K} \subset \mathbb{N}$ tal que d^k converge a $d \in \mathbb{R}^n$, com $\|d\| = 1$. Definindo, para todo $k \in \mathbb{K}$, $t^k = \|x^k - \bar{x}\|$ temos que

$$x^k = \bar{x} + t^k d^k = \bar{x} + t^k d + t^k (d^k - d) = \bar{x} + t^k d + o(t^k).$$

Pelo Lema 1.11, é fácil ver que a função $\Phi(\cdot, \rho)$ é Lipschitz-contínua (digamos, com módulo $L > 0$) numa vizinhança de \bar{x} . Portanto,

$$|\Phi(x^k, \rho) - \Phi(\bar{x} + t^k d, \rho)| \leq L \|x^k - \bar{x} - t^k d\| = o(t^k).$$

Assim, obtemos que

$$\Phi'_+(\bar{x}, d, \rho) = \lim_{k \rightarrow \infty} \frac{\Phi(\bar{x} + t^k d, \rho) - \Phi(\bar{x}, \rho)}{t^k} = \lim_{k \rightarrow \infty} \frac{\Phi(x^k, \rho) - \Phi(\bar{x}, \rho) + o(t^k)}{t^k} \leq 0,$$

onde a desigualdade segue de (1.44). Como pela hipótese $\bar{x} \in D$, combinando (1.25) com a desigualdade acima e obtemos que

$$0 \geq \nabla f(\bar{x})^T d + \rho \sum_{i=1}^{\ell} |\nabla h_i(\bar{x})^T d| + \rho \sum_{i \in I^0(\bar{x})} \max\{0, \nabla g_i(\bar{x})^T d\}. \quad (1.45)$$

Em particular, como os dois últimos termos são não negativos, isto implica que

$$0 \geq \nabla f(\bar{x})^T d. \quad (1.46)$$

Por outro lado, pela condição suficiente de segunda ordem (que inclui as condições de otimalidade KKT),

$$\nabla f(\bar{x}) + \sum_{i=1}^{\ell} \nu_i \nabla h_i(\bar{x}) + \sum_{i \in I^0(\bar{x})} \mu_i \nabla g_i(\bar{x}) = 0.$$

Agora, multiplicando a relação acima por d^T e usando (1.45) obtemos que

$$\begin{aligned}
0 &\geq -\sum_{i=1}^{\ell} \nu_i \nabla h_i(\bar{x})^T d - \sum_{i \in I^0(\bar{x})} \mu_i \nabla g_i(\bar{x})^T d \\
&\quad + \rho \sum_{i=1}^{\ell} |\nabla h_i(\bar{x})^T d| + \rho \sum_{i \in I^0(\bar{x})} \max\{0, \nabla g_i(\bar{x})^T d\} \\
&\geq (\rho - \|(\nu, \mu)\|_{\infty}) \left(\sum_{i \in I^0(\bar{x})} \max\{0, \nabla g_i(\bar{x})^T d\} + \sum_{i=1}^{\ell} |\nabla h_i(\bar{x})^T d| \right),
\end{aligned}$$

onde $\bar{\mu} \geq 0$ foi levado em conta. Como $\rho > \|(\nu, \mu)\|_{\infty}$ e todos os termos no lado direito da relação acima são não negativos, concluímos que todos eles têm que ser nulos. Portanto,

$$\nabla h_i(\bar{x})^T d = 0, \quad i = 1, \dots, \ell, \quad (1.47)$$

$$\nabla g_i(\bar{x})^T d \leq 0, \quad i \in I^0(\bar{x}). \quad (1.48)$$

Assim, com estas condições e levando em conta que $\|d\| = 1$ então pela condição suficiente de segunda ordem temos que

$$d^T \nabla^2 \mathcal{L}(\bar{x}, \nu, \mu) d \geq \gamma \|d\|^2 = \gamma,$$

onde $\gamma > 0$.

Levando em conta que $\nabla \mathcal{L}(\bar{x}, \nu, \mu) = 0$, obtemos

$$\begin{aligned}
\mathcal{L}(x^k, \nu, \mu) &= \mathcal{L}(\bar{x}, \nu, \mu) + (\nabla \mathcal{L}(\bar{x}, \nu, \mu))^T (x^k - \bar{x}) \\
&\quad + \frac{1}{2} (x^k - \bar{x})^T \nabla^2 \mathcal{L}(\bar{x}, \nu, \mu) (x^k - \bar{x}) + o\left(\|x^k - \bar{x}\|^2\right) \\
&= \mathcal{L}(\bar{x}, \nu, \mu) + \frac{(t^k)^2}{2} d^T \nabla^2 \mathcal{L}(\bar{x}, \nu, \mu) d + o((t^k)^2) \\
&\geq \mathcal{L}(\bar{x}, \nu, \mu) + \frac{(t^k)^2}{2} \gamma + o((t^k)^2) \\
&\geq \mathcal{L}(\bar{x}, \nu, \mu) + \frac{(t^k)^2}{4} \gamma \\
&> \mathcal{L}(\bar{x}, \nu, \mu),
\end{aligned} \quad (1.49)$$

onde a segunda desigualdade vale para todo k suficientemente grande (tal que $o((t^k)^2) \geq -\gamma(t^k)^2/4$).

De (1.44) e (1.49), para todo k suficientemente grande obtemos

$$\Phi(x^k, \rho) \leq \Phi(\bar{x}, \rho) = f(\bar{x}) = \mathcal{L}(\bar{x}, \nu, \mu) < \mathcal{L}(x^k, \nu, \mu) \leq \Phi(x^k, \rho),$$

onde a última igualdade segue de (1.37), o qual é absurdo e assim, \bar{x} é solução de (1.21).

□

Para finalizar a análise do método de penalização exata, consideremos a família de problemas parametrizada por $\alpha \in \mathbb{R}^m$ e $\beta \in \mathbb{R}^{\ell}$

$$\begin{aligned}
& \text{minimizar} && f(x) \\
& \text{sujeito a} && g(x) + \alpha \leq 0, \\
& && h(x) + \beta = 0, \\
& && x \in C,
\end{aligned} \tag{1.50}$$

onde f, g, h são Lipschitz contínuas definidas como em (1.19) e C um subconjunto fechado de \mathbb{R}^n . Observe que se $\alpha = 0, \beta = 0, C = \mathbb{R}^n$ o problema (1.50) coincide com o problema (1.19). A seguinte definição é conhecida na literatura como condição de calma.

Definição 1.13 [17, Def. 6.4.1] *Seja x solução do problema (1.19). O problema (1.19) é calmo em x quando existem números positivos ε e M tais que para todo $(\alpha, \beta) \in B(0, \varepsilon) \subset \mathbb{R}^{m+\ell}$ e para todo ponto $x' \in B(x, \varepsilon)$ viável para o problema (1.50) temos que*

$$f(x') - f(x) + M \|(\alpha, \beta)\| \geq 0. \tag{1.51}$$

Perturbação das restrições são estudadas na teoria de análise de sensibilidade [30].

O conceito de calma de um problema está relacionado com a penalização exata [17, Prop. 6.4.3]. Além disso, a condição de calma de um problema no ponto x representa uma condição de qualificação [17, Prop. 6.4.4], no sentido da garantia da existência dos multiplicadores de Lagrange. Mas cabe ressaltar, que não há garantia da unicidade dos multiplicadores.

1.1.2 Penalização quadrática

Embora os métodos de penalidade externa no contexto geral já foram considerados, apresentamos um método específico para restrições de igualdade.

O problema de minimização a resolver por este método é dado por

$$\begin{aligned}
& \text{minimizar} && f(x) \\
& \text{sujeito a} && h(x) = 0,
\end{aligned} \tag{1.52}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}^\ell$ são funções suaves.

Ou seja, consideramos agora o caso particular em que

$$D = \{x \in \mathbb{R}^n | h(x) = 0\}.$$

Considere a função penalidade quadrática $P : \mathbb{R}^n \rightarrow \mathbb{R}_+$ definida por

$$P(x) = \frac{1}{2} \|h(x)\|^2 = \frac{1}{2} \sum_{i=1}^{\ell} (h_i(x))^2.$$

Note que esta função é continuamente diferenciável e satisfaz (1.2), sendo portanto uma penalização externa.

Consideramos então a família de problemas irrestritos (1.7) como

$$\text{minimizar}_{x \in \mathbb{R}^n} \Phi(x, \rho)$$

com

$$\Phi(x, \rho) = f(x) + \frac{\rho}{2} \sum_{i=1}^{\ell} (h_i(x))^2. \quad (1.53)$$

Como no Algoritmo 1.1, em uma iteração k , dado um parâmetro de penalidade $\rho^k > 0$, x^k será um minimizador de $\Phi(x^k, \rho^k)$. Se x^k é um minimizador global exato então o Teorema 1.4 garante a convergência global do algoritmo. No entanto, este minimizador não precisa ser exato. Basta que seja um ponto estacionário de Φ com precisão $\varepsilon^k > 0$, ou seja, x^k deve satisfazer

$$\|\nabla\Phi(x^k, \rho^k)\| \leq \varepsilon^k$$

onde $\nabla\Phi(x^k, \rho^k)$ denota o gradiente de Φ em relação à primeira variável x .

Descrevemos a seguir o algoritmo com penalização quadrática, com os subproblemas resolvidos da forma inexata. Considere $(\varepsilon^k)_{k \in \mathbb{N}}$ uma sequência de valores positivos que converge para zero.

Algoritmo 1.2 -*Penalização quadrática*

Dados $\rho^0 > 0$, $\varepsilon^0 > 0$, $x_s^0 \in \mathbb{R}^n$.

$k = 0$.

REPITA

Obtenha x^k como ponto estacionário de $\Phi(\cdot, \rho^k)$ com precisão ε^k ,
a partir de x_s^k .

ESCOLHA $\rho^{k+1} > \rho^k$,

ESCOLHA $x_s^{k+1} \in \mathbb{R}^n$

$k = k + 1$

Note que a cada iteração é considerado um ponto x_s^k como ponto inicial na resolução do subproblema a maioria dos casos considera-se $x_s^{k+1} = x^k$ para a iteração seguinte.

O próximo teorema garante que todo ponto de acumulação da sequência gerada pelo Algoritmo 1.2, com independência linear dos gradientes das restrições é um ponto estacionário para o problema original, ou seja, satisfaz condições KKT [38].

Teorema 1.14 [34, Thm. 17.2] *Suponha que $\varepsilon^k \rightarrow 0$ e $\rho \rightarrow \infty$. Considere x^* ponto de acumulação da sequência $(x^k)_{k \in \mathbb{N}}$ no qual o conjunto dos gradientes das restrições $\{\nabla h_i(x^*) | i = 1, \dots, \ell\}$ é linearmente independente. Então x^* é um ponto KKT para o problema original (1.52). Além disso, para tal ponto, existe um subconjunto de índices \mathbb{K} tal que $\lim_{k \in \mathbb{K}} x^k = x^*$ e*

$$\lim_{k \in \mathbb{K}} \rho^k h(x^k) = \mu^*$$

onde μ^* é o vetor multiplicador de Lagrange associado a x^* .

Demonstração. Diferenciando (1.53) com respeito à primeira variável temos

$$\nabla\Phi(x^k, \rho^k) = \nabla f(x^k) + \rho^k \sum_{i=1}^{\ell} h_i(x^k) \nabla h_i(x^k). \quad (1.54)$$

Do Algoritmo 1.2 segue que

$$\left\| \nabla f(x^k) + \rho^k \sum_{i=1}^{\ell} h_i(x^k) \nabla h_i(x^k) \right\| \leq \varepsilon^k,$$

para todo k . Por propriedades de norma ($\|a + b\| \geq \|a\| - \|b\|$), obtém-se que

$$\left\| \sum_{i=1}^{\ell} h_i(x^k) \nabla h_i(x^k) \right\| \leq \frac{1}{\rho^k} (\varepsilon^k + \|\nabla f(x^k)\|). \quad (1.55)$$

Como x^* é ponto de acumulação então existe um conjunto de índices $\mathbb{K} \subset \mathbb{N}$ tal que $x^k \rightarrow x^*$ em \mathbb{K} . Assim tomando limite sobre os índices do \mathbb{K} em (1.55) e usando a continuidade do ∇f , a convergência $\varepsilon^k \rightarrow 0$ e o fato que $\rho^k \rightarrow \infty$ obtemos

$$\sum_{i=1}^{\ell} h_i(x^*) \nabla h_i(x^*) = 0.$$

Como por hipótese o conjunto $\{\nabla h_i(x^*), i = 1, \dots, \ell\}$ é linearmente independente tem-se que $h(x^*) = 0$, ou seja, x^* é um ponto viável.

Por outro lado, para que x^* satisfaça KKT é suficiente provar que

$$\lim_{k \in \mathbb{K}} \rho^k h(x^k) = \mu^*.$$

Vamos considerar $A(x)$ a matriz dos gradientes das restrições, isto é,

$$A(x)^T = [\nabla h_1(x), \nabla h_2(x), \dots, \nabla h_\ell(x)],$$

além disso definamos o vetor $\mu^k = \rho^k h(x^k)$, logo podemos escrever (1.54) como

$$\nabla \Phi(x^k, \rho^k) = \nabla f(x^k) + A(x^k)^T \mu^k, \quad (1.56)$$

de onde

$$A(x^k)^T \mu^k = -[\nabla f(x^k) - \nabla \Phi(x^k, \rho^k)]. \quad (1.57)$$

Como a matriz $A(x^k)$ tem posto máximo para $k \in \mathbb{K}$ suficientemente grande, a matriz $[A(x^k)A(x^k)^T]$ é não singular. Logo multiplicando (1.57) por $A(x^k)$ temos

$$[A(x^k)A(x^k)^T] \mu^k = -A(x^k)[\nabla f(x^k) - \nabla \Phi(x^k, \rho^k)]$$

e assim para k suficientemente grande temos

$$\mu^k = -[A(x^k)A(x^k)^T]^{-1} A(x^k)[\nabla f(x^k) - \nabla \Phi(x^k, \rho^k)].$$

Logo aplicando limite e utilizando a continuidade das funções f, h_i e usando que $\nabla \Phi(x^k, \rho^k) \rightarrow 0$ (pois $\varepsilon^k \rightarrow 0$) temos

$$\mu^* = \lim_{k \in \mathbb{K}} \mu^k = -[A(x^*)A(x^*)^T]^{-1} A(x^*) \nabla f(x^*).$$

Portanto, tomando limite em (1.56) temos que

$$\nabla f(x^*) + A(x^*)^T \mu^* = 0,$$

ou seja, x^* é um ponto KKT com multiplicador μ^* .

□

1.2 Penalização interna

Os métodos de penalidade interna, ou também conhecidos como métodos barreira, são feitos para problemas do tipo

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && g(x) \leq 0, \end{aligned} \quad (1.58)$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são funções suaves. Neste caso, o nosso conjunto viável é dada por $D = \{x \in \mathbb{R}^n \mid g(x) \leq 0\}$. Seu interior é o conjunto não vazio

$$D^0 = \{x \in \mathbb{R}^n \mid g(x) < 0\}. \quad (1.59)$$

Suponha ainda que

$$\inf_{x \in D^0} f(x) = \inf_{x \in D} f(x) = \bar{v} > -\infty, \quad (1.60)$$

em outras palavras, o ínfimo de f no conjunto D^0 é igual ao valor ótimo do problema (1.58). O objetivo desta hipótese é evitar alguns casos como o da Figura (1.3) onde a solução \bar{x} não pode ser aproximada por pontos do interior do conjunto viável.

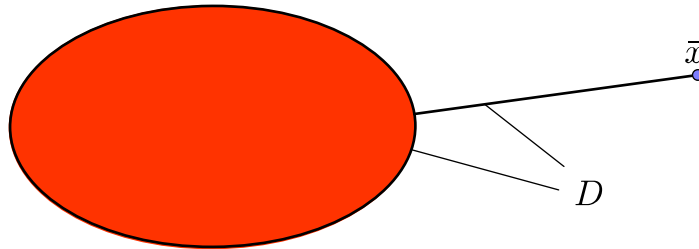


Figura 1.3: A solução \bar{x} não pode ser aproximada por pontos do interior de D .

A ideia dos métodos de penalização interna é a mesma que da penalização externa, no sentido que queremos aproximar o problema (1.58) por uma sequência de problemas que podem ser considerados irrestritos pois a solução de cada problema dependerá da função penalidade a ser considerada. A diferença com os métodos de penalização externa é que o termo penalizador agregado à função objetivo, tende para infinito para qualquer sequência que se aproxime à fronteira do conjunto viável e os pontos das sequências estarão sempre no interior do conjunto viável.

Definição 1.15 Uma função $P : D^0 \rightarrow \mathbb{R}_+$ chama-se barreira para o conjunto D , se ela é contínua em D^0 e para toda sequência $(x^k)_{k \in \mathbb{N}} \subset D^0$ tal que $g_i(x^k) \rightarrow 0$ para algum $i \in \{1, \dots, m\}$, tem-se $\lim_{k \rightarrow \infty} P(x^k) = +\infty$.

As funções barreira mais comuns são

$$P(x) = - \sum_{i=1}^m \log(-g_i(x)), \quad (1.61)$$

e

$$P(x) = -\sum_{i=1}^m \frac{1}{g_i(x)}, \quad (1.62)$$

as quais são conhecidas como barreira logarítmica e barreira inversa, respectivamente.

Consideremos a seguinte família de funções $\varphi : D^0 \times \mathbb{R}_{++} \rightarrow \mathbb{R}$ dada por

$$\varphi(x, \rho) = f(x) + \rho P(x), \quad (1.63)$$

onde $\rho > 0$ é o parâmetro penalizador. Com esta função, dado $\rho > 0$ considere o seguinte problema

$$\begin{array}{ll} \text{minimizar} & \varphi(x, \rho) \\ \text{sujeito a} & x \in D^0. \end{array} \quad (1.64)$$

O problema acima é irrestrito no sentido de que ao aplicar um método iterativo razoável a partir de um ponto inicial em D^0 , a sequência gerada permanecerá dentro de D^0 , já que pela definição da função barreira, à medida que um ponto se aproxime da fronteira de D os valores de $\varphi(\cdot, \rho)$ explodem.

Algoritmo 1.3 - Método de Barreiras

Escolha uma função barreira $P : D^0 \rightarrow \mathbb{R}$ para $D = \{x \in \mathbb{R}^n \mid g(x) \leq 0\}$, $\rho^0 > 0$.
 $k = 0$.

REPITA

DETERMINE $x^k \in D^0$, como solução de (1.64) para a função objetivo dada em (1.63) com $\rho = \rho^k$.

ESCOLHA $\rho^{k+1} < \rho^k$,

$k = k + 1$

Como as soluções do problema original geralmente estão na fronteira de D , então para a convergência destes métodos, é necessário que os valores de ρ tendam para zero e desta maneira amenizar a influência explosiva da função P .

Exemplo 1.16 Consideremos o problema

$$\begin{array}{ll} \text{minimizar} & x \\ \text{sujeito a} & -x \leq 0. \end{array} \quad (1.65)$$

Vejamos que utilizando o método de barreira a solução é o ponto $\bar{x} = 0$.

De fato, consideremos a função barreira inversa definida como em (1.62), assim obtemos que

$$\varphi(x, \rho) = x + \frac{\rho}{x}.$$

Logo os pontos estacionários para esta função são dados por

$$0 = \varphi'(x, \rho) = 1 - \frac{\rho}{x^2}$$

de onde obtemos que

$$x(\rho) = \sqrt{\rho}.$$

Observe que $x(\rho) \in D^0$ pois neste caso $D^0 = \{x \in \mathbb{R} \mid x > 0\}$. Em particular, quando $\rho \rightarrow 0$ então $x(\rho) \rightarrow 0$.

□

Estudemos agora algumas propriedades e as convergência das funções de penalização interna.

Proposição 1.17 [25, Prop. 4.3.3] *Suponhamos que o conjunto D satisfaça (1.59) e (1.60). Seja $(x^k)_{k \in \mathbb{N}}$ uma sequência gerada pelo Algoritmo 1.3, onde x^k é solução global do problema (1.64) com $\rho = \rho^k$. Então para todo k , tem-se que*

$$\varphi(x^{k+1}, \rho^{k+1}) \leq \varphi(x^k, \rho^k) \quad (1.66)$$

$$P(x^{k+1}) \geq P(x^k) \quad (1.67)$$

$$f(x^{k+1}) \leq f(x^k). \quad (1.68)$$

Demonstração. Como x^{k+1} é solução global de (1.64) com $\rho = \rho^{k+1}$ então temos que

$$\begin{aligned} f(x^{k+1}) + \rho^{k+1}P(x^{k+1}) &= \varphi(x^{k+1}, \rho^{k+1}) \\ &\leq \varphi(x^k, \rho^{k+1}) \\ &= f(x^k) + \rho^{k+1}P(x^k) \\ &\leq f(x^k) + \rho^k P(x^k) \\ &= \varphi(x^k, \rho^k), \end{aligned} \quad (1.69)$$

onde a segunda desigualdade é válida pois $\rho^{k+1} < \rho^k$ e $P(x^k) \geq 0$. O que mostra (1.66). Para mostrar (1.67) vejamos que pela otimalidade de x^k , com $\rho = \rho^k$ temos que

$$f(x^k) + \rho^k P(x^k) = \varphi(x^k, \rho^k) \leq \varphi(x^{k+1}, \rho^k) = f(x^{k+1}) + \rho^k P(x^{k+1}).$$

Logo somando esta desigualdade com (1.69) obtemos

$$(\rho^k - \rho^{k+1})P(x^k) \leq (\rho^k - \rho^{k+1})P(x^{k+1}). \quad (1.70)$$

Agora como $(\rho^k - \rho^{k+1}) > 0$ então de (1.70) concluímos que $P(x^{k+1}) \geq P(x^k)$. Finalmente combinando (1.67) e (1.69) obtemos

$$f(x^{k+1}) - f(x^k) \leq \rho^{k+1}(P(x^k) - P(x^{k+1})) \leq 0,$$

o que mostra (1.68). □

Temos a seguir o resultado de convergência dos métodos de barreira.

Teorema 1.18 (Convergência dos métodos de penalização interna) [25, Teo. 4.3.7] *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ funções contínuas no \mathbb{R}^n . Suponha que D satisfaça (1.59) e (1.60). considere $(x^k)_{k \in \mathbb{N}}$ uma sequência gerada pelo Algoritmo 1.3, onde x^k é solução do problema (1.64) com $\rho = \rho^k$. Se $\rho^k \rightarrow 0$, então todo ponto de acumulação de $(x^k)_{k \in \mathbb{N}}$ é solução global do problema (1.58).*

Demonstração. Seja x^k um ponto de acumulação da sequência $(x^k)_{k \in \mathbb{N}}$. Logo existe um conjunto $\mathbb{K} \subset \mathbb{N}$ tal que $\lim_{k \in \mathbb{K}} x^k = \bar{x}$. Como $(x^k)_{k \in \mathbb{K}} \subset D^0$ então temos que $g(x^k) < 0$ para todo k e pela continuidade de g temos que $g(\bar{x}) \leq 0$, ou seja, $\bar{x} \in D$.

Se $f(\bar{x}) = \bar{v}$ então não há nada para mostrar, pois neste caso \bar{x} já é solução de (1.58). Agora suponha que $f(\bar{x}) > \bar{v}$. Consideremos

$$\delta = \frac{(f(\bar{x}) - \bar{v})}{2} > 0. \quad (1.71)$$

Como (1.60) é satisfeito então pela definição de ínfimo temos que existe $x' \in D^0$ tal que

$$f(x') < \bar{v} + \delta. \quad (1.72)$$

Por outro lado pela desigualdade (1.68) temos que a sequência de imagens $(f(x^k))_{k \in \mathbb{K}}$ é não crescente e pela continuidade de f segue que

$$f(x^k) \geq \lim_{k \in \mathbb{K}} f(x^k) = f(\bar{x}) = \bar{v} + 2\delta.$$

Agora subtraindo $f(x')$ da desigualdade acima e utilizando (1.72) obtemos que

$$f(x^k) - f(x') \geq \bar{v} + 2\delta - f(x') > \delta. \quad (1.73)$$

Logo para $k \in \mathbb{K}$ suficientemente grande

$$\begin{aligned} 0 &\geq \varphi(x^k, \rho^k) - \varphi(x', \rho^k) \\ &= f(x^k) + \rho^k P(x^k) - f(x') - \rho^k P(x') \\ &> \delta + \rho^k (P(x^k) - P(x')) \\ &\geq \delta + \rho^k (P(x^0) - P(x')) \\ &> 0, \end{aligned}$$

onde a primeira desigualdade é válida pois x^k é solução global do problema (1.64) com $\rho = \rho^k$, a segunda desigualdade segue por (1.73), a terceira por (1.67) e a última pelo fato de que $\rho^k \rightarrow 0$. O absurdo mostrado acima implica que $f(\bar{x}) = \bar{v}$, isto é, \bar{x} é solução do problema (1.58). \square

Como no caso da penalização externa, os métodos de penalização interna também tem o problema de mal-condicionamento, neste caso para valores do parâmetro penalizador tendendo para zero. O seguinte exemplo mostra esta situação.

Exemplo 1.19

Considere o problema

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

A solução deste problema é $\bar{x} = (0, 1)$ e o conjunto $D^0 = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 > 0\}$. Consideremos a função barreira logarítmica definida definida em (1.61). Logo o problema penalizado é dado por

$$\varphi(x, \rho) = (x_1 + 1)^2 + (x_2 - 1)^2 - \rho \log(x_1). \quad (1.74)$$

Para obter os pontos críticos da função acima, tomamos gradiente respeito à primeira variável e obtemos

$$0 = \nabla\varphi(x, \rho) = \begin{pmatrix} 2(x_1 + 1) - \frac{\rho}{x_1} \\ 2(x_2 - 1) \end{pmatrix}.$$

Logo, a solução deste sistema é dada por

$$x(\rho) = \left(\frac{-1 + \sqrt{1 + 2\rho}}{2}, 1 \right).$$

Assim, se $\rho \rightarrow 0$ temos que $x(\rho) \rightarrow (0, 1) = \bar{x}$. Por outro lado, a matriz hessiana da função (1.74) é dada por

$$\nabla^2\varphi(x, \rho) = \begin{pmatrix} 2 + \frac{\rho}{x_1^2} & 0 \\ 0 & 2 \end{pmatrix}$$

e avaliando no ponto $x(\rho)$ tem-se que

$$\nabla^2\varphi(x(\rho), \rho) = \begin{pmatrix} 2 + \frac{2\rho}{1 + \rho - \sqrt{1 + 2\rho}} & 0 \\ 0 & 2 \end{pmatrix}$$

como

$$\lim_{\rho \rightarrow 0} \frac{2\rho}{1 + \rho - \sqrt{1 + 2\rho}} = +\infty,$$

vemos que esta matriz fica cada vez pior condicionada quando $\rho \rightarrow 0$.

□

Capítulo 2

Métodos de Lagrangiano aumentado

No capítulo anterior discutimos alguns métodos de penalização. Vimos que nos métodos de penalização externa, o parâmetro de penalidade pode crescer ilimitadamente, o que pode causar problemas de instabilidade do método. Uma maneira de evitar crescimento ilimitado do parâmetro de penalização é acrescentar um termo de penalização diferenciável à Lagrangiana, e não à função objetivo do problema, dando origem aos métodos de Lagrangiano aumentado.

Neste capítulo estudaremos o método de Lagrangiano aumentado e apresentaremos duas propostas para a função de penalização que dão origem a diferentes métodos de Lagrangiano aumentado. Discutiremos seus resultados de convergência global. O Capítulo 3 será dedicado a testes computacionais para comparação do desempenho dos métodos obtidos com as diferentes funções de penalização estudadas.

2.1 Método clássico

Considere inicialmente o problema com restrições de igualdade

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && h(x) = 0, \end{aligned} \tag{2.1}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}^\ell$ são funções continuamente diferenciáveis.

A função Lagrangiana $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^\ell \rightarrow \mathbb{R}$ associada ao problema é definida por

$$\mathcal{L}(x, \mu) = f(x) + \sum_{i=1}^{\ell} \mu_i h_i(x).$$

No método de Lagrangiano aumentado adicionamos à função Lagrangiana uma função de penalização. No método clássico, adicionamos uma função de penalização quadrática multiplicada pelo parâmetro de penalidade. Temos assim a função Lagrangiano aumentado

$$L(x, \mu, \rho) = f(x) + \sum_{i=1}^{\ell} \mu_i h_i(x) + \frac{\rho}{2} \sum_{i=1}^{\ell} (h_i(x))^2 \tag{2.2}$$

onde ρ é o parâmetro de penalidade. Note que esta função difere da função Lagrangiana padrão pelo termo quadrático adicionado à direita.

Cabe ressaltar que o método clássico de Lagrangiano aumentado pode ser visto como um caso particular do método de penalização quadrática discutida na Seção 1.1.2. Ao aplicarmos o método de penalização quadrática ao problema

$$\begin{aligned} \text{minimizar} \quad & f(x) + \sum_{i=1}^{\ell} \mu_i h_i(x) \\ \text{sujeito a} \quad & h(x) = 0, \end{aligned}$$

vemos que os problemas a serem resolvidos no Algoritmo 1.2 são justamente de minimização de (2.2).

Como no Algoritmo 1.2, o ponto x^k será um minimizador inexato da função penalizada $L(x, \mu, \rho)$, ou seja, dado $\mu^k \in \mathbb{R}^{\ell}$, $\rho^k > 0$, e uma tolerância $\varepsilon^k > 0$, o ponto x^k deve satisfazer

$$\|\nabla L(x^k, \mu^k, \rho^k)\| \leq \varepsilon^k. \quad (2.3)$$

Devemos discutir agora como atualizar o vetor $\mu \in \mathbb{R}^{\ell}$ dos multiplicadores de Lagrange.

Se derivamos a expressão (2.2) com respeito a x temos que

$$\begin{aligned} \nabla L(x, \mu, \rho) &= \nabla f(x) + \sum_{i=1}^{\ell} \mu_i \nabla h_i(x) + \rho \sum_{i=1}^{\ell} h_i(x) \nabla h_i(x) \\ &= \nabla f(x) + \sum_{i=1}^{\ell} (\mu_i + \rho h_i(x)) \nabla h_i(x). \end{aligned} \quad (2.4)$$

Fixando ρ e μ nesta última igualdade e se consideramos as hipóteses do Teorema 1.14 temos que se x^k é o minimizador de $L(x, \mu, \rho)$, para k suficientemente grande

$$\mu_i^* \approx \mu_i^k + \rho^k h_i(x^k) \quad (2.5)$$

para todo $i = 1, \dots, \ell$. Assim,

$$h_i(x^k) \approx \frac{1}{\rho^k} (\mu_i^* - \mu_i^k).$$

Note que quando μ^k se aproxima de μ^* , $h_i(x^k)$ está próximo de zero, e portanto, x^k torna-se viável. Agora para aproveitar este cálculo de x^k em cada iteração e calcular uma estimativa para o multiplicador associado, de (2.5) temos uma ideia natural para atualizar o vetor de multiplicadores, ou seja,

$$\mu_i^{k+1} = \mu_i^k + \rho^k h_i(x^k)$$

para todo $i = 1, \dots, \ell$. Considerando uma sequência de $(\varepsilon^k)_{k \in \mathbb{K}}$ de valores positivos convergindo para zero, definimos o seguinte algoritmo.

Algoritmo 2.1 - *Lagrangiano aumentado*

Dados $\rho^0 > 0$, $\varepsilon^0 > 0$, $x_s^0 \in \mathbb{R}^n$ e $\mu^0 \in \mathbb{R}^{\ell}$, $c > 1$.

$k = 0$

REPITA

Obtenha x^k , a partir de x_s^k , satisfazendo (2.3)

Atualizar para cada $i = 1, \dots, \ell$, $\mu_i^{k+1} = \mu_i^k + \rho^k h_i(x^k)$

$$\rho^{k+1} = c\rho^k$$

Escolha um novo ponto inicial $x_s^{k+1} \in \mathbb{R}^n$

$k = k + 1$

Da mesma forma como no Algoritmo 1.2 em cada iteração este algoritmo considera um ponto inicial para resolver o subproblema. Na maioria das vezes este ponto é tomado como $x_s^{k+1} = x^k$.

2.1.1 Análise de convergência

Nesta seção apresentamos a análise de convergência global do Algoritmo 2.1, supondo a condição de qualificação de independência linear dos gradientes das restrições ativas. No entanto, em [1] esta análise é feita considerando uma condição de qualificação mais fraca, ou seja, a condição de Dependência Linear Positiva Constante (CPLD), introduzida em [37] e discutida em [3].

Teorema 2.1 [34, Thm. 17.5] *Seja x^* solução local do problema (1.52), tal que o conjunto dos gradientes das restrições $\{\nabla h_1(x^*), \dots, \nabla h_\ell(x^*)\}$ é linearmente independente e a condição suficiente de segunda ordem para o problema (2.1) é satisfeita para $\mu = \mu^*$. Então existe um número $\bar{\rho}$ tal que para todo $\rho > \bar{\rho}$, x^* é um minimizador local estrito de $L(x, \mu^*, \rho)$.*

Demonstração. Mostremos que x^* satisfaz as condições necessárias e suficientes para o problema de otimização sem restrições, ou seja,

$$\nabla L(x^*, \mu^*, \rho) = 0 \quad e \quad \nabla^2 L(x^*, \mu^*, \rho) \quad \text{é positiva definida.} \quad (2.6)$$

Como x^* é um ponto viável então

$$\begin{aligned} \nabla L(x^*, \mu^*, \rho) &= \nabla f(x^*) + \sum_{i=1}^{\ell} (\mu_i^* + \rho h_i(x^*)) \nabla h_i(x^*) \\ &= \nabla f(x^*) + \sum_{i=1}^{\ell} \mu_i^* \nabla h_i(x^*) \\ &= \mathcal{L}(x^*, \mu^*). \end{aligned} \quad (2.7)$$

Logo, como para x^* e μ^* as condições de segunda ordem são válidas, temos que $\mathcal{L}(x^*, \mu^*) = 0$ e conseqüentemente $\nabla L(x^*, \mu^*, \rho) = 0$. Assim, é satisfeita a primeira igualdade de (2.6).

Vejam agora que existe $\rho > 0$ tal que $u^T \nabla^2 L(x^*, \mu^*, \rho) u > 0$ para todo $u \in \mathbb{R}^n$ não nulo. Considere a matriz $A^T = [\nabla c_1(x^*), \nabla c_2(x^*), \dots, \nabla c_\ell(x^*)]$. Como o conjunto dos gradientes $\{\nabla h_i(x^*), i = 1, \dots, \ell\}$ é linearmente independente então A tem posto máximo, assim

$$\begin{aligned} \nabla^2 L(x^*, \mu^*, \rho) &= \nabla^2 f(x^*) + \sum_{i=1}^{\ell} \mu_i^* \nabla^2 h_i(x^*) + \rho \left(A^T A + \sum_{i=1}^{\ell} h_i(x^*) \nabla^2 h_i(x^*) \right) \\ &= \nabla^2 f(x^*) + \sum_{i=1}^{\ell} \mu_i^* \nabla^2 h_i(x^*) + \rho A^T A \\ &= \nabla^2 \mathcal{L}(x^*, \mu^*) + \rho A^T A \end{aligned} \quad (2.8)$$

Agora pelo teorema fundamental da álgebra $\mathbb{R}^n = \mathcal{N}(A) \oplus \text{Im}(A^T)$, onde $\mathcal{N}(A)$ e $\text{Im}(A^T)$ são o núcleo e imagem da matriz A . Logo dado $u \in \mathbb{R}^n$, $u = w + A^T v$ onde

$w \in \mathcal{N}(A)$ e $v \in \mathbb{R}^\ell$, por tanto

$$\begin{aligned} u^T \nabla^2 L(x^*, \mu^*, \rho) u &= u^T \nabla^2 \mathcal{L}(x^*, \mu^*) u + \rho u^T A^T A u \\ &= w^T \nabla^2 \mathcal{L}(x^*, \mu^*) w + 2w^T \nabla^2 \mathcal{L}(x^*, \mu^*) A^T v \\ &\quad + v^T A \nabla^2 \mathcal{L}(x^*, \mu^*) A^T v + \rho v^T A (A^T A) A^T v. \end{aligned} \quad (2.9)$$

Estudemos os termos à direita da última igualdade. Como a Hessiana da função Lagrangiana é definida positiva no $\mathcal{N}(A)$ (pela condição suficiente de segunda ordem), para (x^*, μ^*) vale que $w^T \nabla^2 \mathcal{L}(x^*, \mu^*) w > 0$ pois $w \in \mathcal{N}(A)$. Por outro lado, como $\mathcal{N}(A) \cap S^{n-1}$ é compacto, pela continuidade da Hessiana da função Lagrangiana, a imagem é um conjunto compacto assim, existe $a > 0$ tal que

$$\frac{w^T}{\|w\|} \nabla^2 \mathcal{L}(x^*, \mu^*) \frac{w^T}{\|w\|} \geq a$$

o que implica

$$w^T \nabla^2 \mathcal{L}(x^*, \mu^*) w \geq a \|w\|^2, \text{ para todo } w \in \mathcal{N}(A). \quad (2.10)$$

Para analisar o segundo termo de (2.9) definamos $b = \|\nabla^2 \mathcal{L}(x^*, \mu^*) A^T\|$. Assim

$$|w^T \nabla^2 \mathcal{L}(x^*, \mu^*) A^T v| \leq \|w\| \|\nabla^2 \mathcal{L}(x^*, \mu^*) A^T v\| \leq \|w\| \|\nabla^2 \mathcal{L}(x^*, \mu^*) A^T\| \|v\| = \|w\| \|v\| b$$

e conseqüentemente,

$$2w^T \nabla^2 \mathcal{L}(x^*, \mu^*) A^T v \geq -2b \|v\| \|w\|. \quad (2.11)$$

De maneira semelhante para $c = \|A \nabla^2 \mathcal{L}(x^*, \mu^*) A^T\|$ temos que

$$v^T A \nabla^2 \mathcal{L}(x^*, \mu^*) A^T v \geq -c \|v\|^2, \quad (2.12)$$

limitando inferiormente o terceiro termo da soma em (2.9). Finalmente, como AA^T é simétrica positiva definida então pelo teorema espectral temos que existe uma decomposição em matriz de AA^T como sendo $AA^T = UDU^T$ com U ortogonal e $D = \text{diag}(\alpha_1, \dots, \alpha_\ell)$ diagonal onde o conjunto $\{\alpha_i, i = 1, \dots, \ell\}$ são os autovalores de AA^T . Tomando $y = U^T v$ e α_{\min} o menor autovalor de AA^T temos

$$\begin{aligned} v^T A (A^T A) A^T v &= \|AA^T v\|^2 \\ &= \|UDU^T v\|^2 \\ &= (UDU^T v)^T (UDU^T v) \\ &= v^T U D^T D U^T v \\ &= y^T D^2 y \\ &= \sum_{i=1}^m \alpha_i^2 y_i^2 \\ &\geq \alpha_{\min}^2 \|y\|^2 \\ &= \alpha_{\min}^2 \|v\|^2, \end{aligned} \quad (2.13)$$

sendo que a última igualdade é válida pois a matriz U é ortogonal. Em conclusão usando (2.10), (2.11), (2.12) e (2.13) em (2.9) temos que

$$\begin{aligned}
u^T \nabla^2 \mathcal{L}(x^*, \mu^*, \rho) u &\geq a \|w\|^2 - 2b \|v\| \|w\| + (\rho \alpha_{\min}^2 - c) \|v\|^2 \\
&= a \|w\|^2 - 2b \|v\| \|w\| + \frac{b^2}{a} \|v\| - \frac{b^2}{a} \|v\| + (\rho \alpha_{\min}^2 - c) \|v\|^2 \\
&= a \left(\|w\| - \frac{b}{a} \|v\| \right)^2 + \left(\rho \alpha_{\min}^2 - c - \frac{b^2}{a} \right) \|v\|. \tag{2.14}
\end{aligned}$$

Logo existe $\bar{\rho} = \frac{ac + b^2}{a\alpha_{\min}^2}$ tal que

$$\nabla^2 L(x^*, \mu^*, \rho) > 0,$$

para todo $\rho > \bar{\rho}$. Pois para que $\nabla^2 L(x^*, \mu^*, \rho) = 0$ pela escolha de $\bar{\rho}$, os termos da soma em (2.14) serão nulos se, e somente se, $w = v = 0$ o que é absurdo pois $u \neq 0$. Logo é satisfeita a segunda parte de (2.6) e assim x^* é minimizador local estrito de $\mathcal{L}(x, \mu^*, \rho)$.

□

O resultado a seguir, dado por Bertsekas em [7, Prop. 4.23] e apresentado em [34, Thm. 17.6], estabelece condições sob as quais existe um minimizador de $L(x, \mu, \rho)$ que está próximo de x^* e fornece limitantes para os erros em x^k e μ^{k+1} obtidos resolvendo o subproblema na iteração k .

Teorema 2.2 *Considere $x^* \in \mathbb{R}^n$ e $\mu^* \in \mathbb{R}^\ell$ satisfazendo as hipóteses do teorema anterior e seja $\bar{\rho} > 0$ garantido pelo teorema. Então existem escalares δ, ϵ e M tais que*

i) Para todo μ^k e ρ^k satisfazendo

$$\|\mu^k - \mu^*\| \leq \delta \rho^k, \quad \rho^k \geq \bar{\rho}, \tag{2.15}$$

o problema

$$\begin{array}{ll}
\text{minimizar} & L(x, \mu^k, \rho^k) \\
\text{sujeito a} & \|x - x^*\| \leq \epsilon,
\end{array}$$

tem uma única solução x^k . Além disso, temos que

$$\|x^k - x^*\| \leq \frac{M}{\rho^k} \|\mu^k - \mu^*\|.$$

ii) Para todo μ^k e ρ^k satisfazendo (2.15) tem-se que

$$\|\mu^{k+1} - \mu^*\| \leq \frac{M}{\rho^k} \|\mu^k - \mu^*\|,$$

onde as coordenadas de μ^{k+1} são dadas como no Algoritmo 2.1, isto é, $\mu_i^{k+1} = \mu_i^k + \rho^k h_i(x^k)$, para todo $i \in \{1, \dots, \ell\}$.

iii) Para todo μ^k e ρ^k satisfazendo (2.15), a matriz $\nabla^2 L(x^k, \mu^k, \rho^k)$ é positiva definida e os gradientes $\nabla h_i(x^k), i = 1, \dots, \ell$, são linearmente independentes.

2.1.2 Extensão para restrições de desigualdade

Vamos introduzir o método de Lagrangiano aumentado para problemas com restrições de desigualdade. O problema a considerar é

$$\begin{aligned} &\text{minimizar } f(x) \\ &\text{sujeito a } g(x) \leq 0, \end{aligned} \quad (2.16)$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são funções continuamente diferenciáveis. A ideia é reformular o problema por um problema com restrições de igualdade com a introdução de variáveis de folga e usar o Algoritmo (2.1). Estas variáveis serão adicionadas às restrições de desigualdade como segue

$$g_i(x) + \theta_i = 0, \quad \text{com } \theta_i \geq 0,$$

para cada $i = 1, \dots, m$. Neste sentido, o Lagrangiano aumentado do problema (2.16) é dado por

$$\begin{aligned} &\text{minimizar}_{x, \theta} \quad L(x, \mu, \rho, \theta) \\ &\text{sujeito a } \quad \theta_i \geq 0, \quad i \in \{1, \dots, m\} \end{aligned} \quad (2.17)$$

onde

$$L(x, \mu, \rho, \theta) = f(x) + \left(\sum_{i=1}^m \mu_i (g_i(x) + \theta_i) + \frac{\rho}{2} (g_i(x) + \theta_i)^2 \right) \quad (2.18)$$

Observe que cada termo da soma, $c_i(\theta_i) = \mu_i (g_i(x) + \theta_i) + \frac{\rho}{2} (g_i(x) + \theta_i)^2$, pode ser visto como uma função convexa da variável de folga θ_i , com x, μ, ρ fixos. Assim, o minimizador para cada variável θ_i em (2.17) ocorre quando a derivada parcial de $L(x, \mu, \rho, \theta)$ com respeito a θ_i se anula, isto é, para cada $i = 1, \dots, m$

$$\mu_i + \rho(g_i(x) + \theta_i) = 0$$

o qual implica que

$$\theta_i = - \left(g_i(x) + \frac{\mu_i}{\rho} \right).$$

Logo, como $\theta_i \geq 0$, o valor ótimo para cada $i = 1, \dots, m$ é dado por

$$\theta_i = \max \left\{ - \left(g_i(x) + \frac{\mu_i}{\rho} \right), 0 \right\}. \quad (2.19)$$

Substituindo a expressão acima em (2.18) temos que para cada $i = 1, \dots, m$,

$$\mu_i (g_i(x) + \theta_i) + \frac{\rho}{2} (g_i(x) + \theta_i)^2 = \begin{cases} \mu_i g_i(x) + \frac{\rho}{2} (g_i(x))^2 & \text{se } g_i(x) + \frac{\mu_i}{\rho} \geq 0 \\ -\frac{(\mu_i)^2}{2\rho} & \text{caso contrário.} \end{cases}$$

Com este resultado, a função Lagrangiano aumentado dada em (2.18) pode ser reescrita independente da variável de folga, como

$$L(x, \mu, \rho) = f(x) + \sum_{i=1}^m P(g_i(x), \mu_i, \rho) \quad (2.20)$$

onde a função $P : \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_{++} \rightarrow \mathbb{R}$ é dada por

$$P(y, t, s) = \begin{cases} ty + \frac{s}{2}y^2 & \text{se } y + \frac{t}{s} \geq 0 \\ -\frac{t^2}{2s} & \text{caso contrário.} \end{cases}$$

Note que a função P pode ser rescrita na forma,

$$P(y, t, s) = \frac{1}{2s} (\max\{0, sy + t\}^2 - t^2). \quad (2.21)$$

Esta última igualdade é conhecida na literatura como a função de penalidade de Powell-Hestenes-Rockafellar (PHR) [22, 36, 39].

A expressão (2.21) representa uma extensão natural do Lagrangiano aumentado para problemas com restrições de desigualdade. Com esta extensão, podemos aplicar o Algoritmo 2.1 para o caso com restrições de desigualdade, modificando a atualização dos multiplicadores de Lagrange para

$$\mu_i^{k+1} = \max\{0, \mu_i^k + \rho^k g_i(x^k)\}$$

para todo $i = 1, \dots, m$.

2.2 Método geral

Do ponto de vista teórico, esta é a seção mais importante desta dissertação. Nela pretende-se generalizar o conceito do método de Lagrangiano aumentado estudado na seção anterior usando diferentes funções de penalização na definição do Lagrangiano aumentado.

Considere o problema

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && g(x) \leq 0 \\ &&& x \in \Omega \end{aligned} \quad (2.22)$$

onde Ω é um conjunto convexo e compacto $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são funções continuamente diferenciáveis. As restrições definidas em Ω são geralmente simples. De modo que este conjunto Ω é, por exemplo, uma caixa ou um polítopo limitado.

Nossa proposta é discutir o método de Lagrangiano aumentado, baseado no artigo [8] para resolver o problema (2.22). As restrições dadas por $g(x) \leq 0$ serão incluídas na definição do Lagrangiano aumentado.

Para definir a função Lagrangiano aumentado, considere uma função de penalidade $P : \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_{++} \rightarrow \mathbb{R}$, tal que $P'(y, t, s) = \frac{\partial}{\partial y} P(y, t, s)$ existe e é contínua para todo $y \in \mathbb{R}$, $t \in \mathbb{R}_+$ e $s \in \mathbb{R}_{++}$. Considere $0 < t_{\min} < t_{\max} < \infty$. Assumimos que P satisfaz as seguintes propriedades:

A1. $P'(y, t, s) \geq 0$, para todo $y \in \mathbb{R}, t \in \mathbb{R}_+, s \in \mathbb{R}_{++}$.

A2. Se para todo $k \in \mathbb{N}$, $t^k \in [t_{\min}, t_{\max}]$ então

$$\lim_{k \rightarrow \infty} s^k = \infty \text{ e } \lim_{k \rightarrow \infty} y^k = y > 0 \text{ implica que } \lim_{k \rightarrow \infty} P'(y^k, t^k, s^k) = \infty.$$

A3. Se para todo $k \in \mathbb{N}$, $t^k \in [t_{\min}, t_{\max}]$ então

$$\lim_{k \rightarrow \infty} s^k = \infty \text{ e } \lim_{k \rightarrow \infty} y^k = y < 0 \text{ implica que } \lim_{k \rightarrow \infty} P'(y^k, t^k, s^k) = 0.$$

Com estas propriedades a função Lagrangiano aumentado $L : \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}_{++}^m \rightarrow \mathbb{R}$ é definida por

$$L(x, \mu, \rho) = f(x) + \sum_{i=1}^m P(g_i(x), \mu_i, \rho_i), \quad (2.23)$$

para todo $x \in \Omega$, $\mu \in \mathbb{R}_+^m$ e $\rho \in \mathbb{R}_{++}^m$.

Para cada $k \in \mathbb{N}$ dados $\rho^k \in \mathbb{R}_{++}^m$ e $\bar{\mu}^k \in \mathbb{R}_{++}^m$, obtemos $x^k \in \Omega$ como uma solução aproximada do problema

$$\begin{aligned} &\text{minimizar} && L(x, \bar{\mu}^k, \rho^k) \\ &\text{sujeito a} && x \in \Omega. \end{aligned} \quad (2.24)$$

Seja $(\varepsilon^k)_{k \in \mathbb{N}}$ uma sequência de números positivos que converge para zero. Um ponto x^k é dito solução aproximada do problema (2.24) quando

$$\|\mathcal{P}(x^k - \nabla L(x^k, \bar{\mu}^k, \rho^k)) - x^k\| \leq \varepsilon^k, \quad (2.25)$$

onde $\mathcal{P}(z)$ denota a função projeção euclidiana do ponto $z \in \mathbb{R}^n$ sobre o conjunto Ω . Se um ponto $x^{k-1} \in \Omega$ for dado, este é normalmente usado como ponto inicial na resolução do problema (2.24). Apresentamos a seguir o algoritmo para resolver o problema (2.22).

Algoritmo 2.2 Algoritmo de Lagrangiano aumentado

Dados: $x^0 \in \Omega$, $\mu_{\max} > \mu_{\min} > 0$, $\mu_0 \in \mathbb{R}_{++}^m$, $\rho^1 \in \mathbb{R}_{++}^m$, $r \in (0, 1)$, $\gamma > 1$.
 $k = 1$.

REPITA

DETERMINE $\bar{\mu}_i^k \in [\mu_{\min}, \mu_{\max}]$, $i = 1, \dots, m$.

DETERMINE $x^k \in \Omega$ solução aproximada do subproblema (2.24).

$\mu_i^k = P'(g_i(x^k), \bar{\mu}_i^k, \rho_i^k)$, para $i = 1, \dots, m$.

PARA CADA $i = 1, \dots, m$,

SE $\max\{0, g_i(x^k)\} \leq r \max\{0, g_i(x^{k-1})\}$ e $|g_i(x^k) \mu_i^k| \leq r |g_i(x^{k-1}) \mu_i^{k-1}|$

$\rho_i^{k+1} = \rho_i^k$

SENÃO

$\rho_i^{k+1} = \gamma \rho_i^k$

$k = k + 1$.

O Algoritmo 2.2 é bastante geral. Para cada função de penalidade P escolhida tem-se um método de Lagrangiano aumentado diferente. Na próxima seção discutimos a convergência global do algoritmo, independente da penalidade escolhida desde que satisfaça as propriedades **A1-A3**. Na Seção 2.2.2 discutiremos algumas funções de penalidade que serão usadas nos testes numéricos.

2.2.1 Análise de convergência

A análise de convergência que faremos a seguir é dada a partir da classe de Lagrangianos aumentados que satisfaçam as propriedades **A1-A3** apresentadas na seção anterior. Iniciaremos estudando algumas definições e conceitos básicos a serem utilizados no teorema principal.

Definição 2.3 Dizemos que um ponto $\bar{x} \in \Omega$ é degenerado, quando existe $\bar{\lambda} \in \mathbb{R}_+^m$ tal que

$$\sum_{i \in I(\bar{x})} \bar{\lambda}_i > 0 \quad e \quad \mathcal{P} \left(\bar{x} - \sum_{i \in I(\bar{x})} \bar{\lambda}_i \nabla g_i(\bar{x}) \right) - \bar{x} = 0, \quad (2.26)$$

onde $I(\bar{x}) = \{i \mid g_i(\bar{x}) \geq 0\}$.

Em particular, se \bar{x} é viável então $I(\bar{x})$ coincide com o conjunto dos índices das restrições ativas.

Exemplo 2.4

Considere $g_1, g_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por,

$$g_1(x_1, x_2) = -x_2 - x_1^2,$$

$$g_2(x_1, x_2) = x_2 - x_1^2.$$

Por simplicidade tomemos $\Omega = \mathbb{R}^2$. Vejamos que $\bar{x} = (0, 0)$ é degenerado.

Com efeito, $\nabla g_1(0, 0) = (0, -1)$, $\nabla g_2(0, 0) = (0, 1)$ e como $\Omega = \mathbb{R}^2$ temos que

$$\mathcal{P} \left(\bar{x} - \sum_{i=1}^2 \bar{\lambda}_i \nabla g_i(\bar{x}) \right) - \bar{x} = - \sum_{i=1}^2 \bar{\lambda}_i \nabla g_i(\bar{x}).$$

Tomando $\bar{\lambda}_i = 1$, com $i = 1, 2$, temos que $\sum_{i=1}^2 \bar{\lambda}_i \nabla g_i(\bar{x}) = 0$, o que implica que \bar{x} é degenerado. A Figura 2.1 mostra geometricamente o ponto $\bar{x} = (0, 0)^T$.

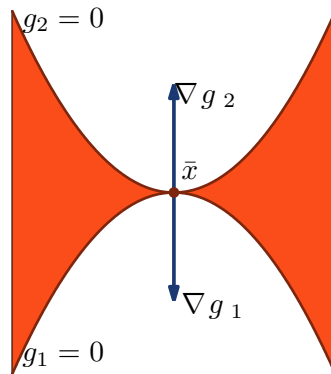


Figura 2.1: Ponto degenerado

Dado $\bar{x} \in \Omega$ degenerado com $\bar{\lambda} \in \mathbb{R}_+^m$ satisfazendo (2.26), considere a função $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por

$$\psi(z) = \sum_{i \in I(\bar{x})} \bar{\lambda}_i \nabla g_i(\bar{x})^T (z - \bar{x}). \quad (2.27)$$

A função ψ é linear, portanto convexa. No caso particular em que $\Omega = \mathbb{R}^n$, de (2.26) temos que $\sum_{i \in I(\bar{x})} \bar{\lambda}_i \nabla g_i(\bar{x}) = 0$ e assim, $\psi \equiv 0$. Neste caso, \bar{x} é um minimizador de ψ em Ω . Veremos a seguir que este resultado é válido quando Ω é considerado como convexo e compacto.

Lema 2.5 *Seja $\bar{x} \in \Omega$ degenerado, então \bar{x} é minimizador global em Ω da função ψ definida em (2.27).*

Demonstração. Seja $z = \bar{x} - \sum_{i \in I(\bar{x})} \bar{\lambda}_i \nabla g_i(\bar{x})$. Como \bar{x} é um ponto degenerado então vale que

$$\mathcal{P}(z) = \mathcal{P}\left(\bar{x} - \sum_{i \in I(\bar{x})} \bar{\lambda}_i \nabla g_i(\bar{x})\right) = \bar{x}.$$

Considere $x \in \Omega$. Pela convexidade do conjunto Ω , temos que para cada $t \in (0, 1)$ o ponto $((1-t)\bar{x} + tx) \in \Omega$. Como $\bar{x} = \mathcal{P}(z)$ segue que para qualquer $t \in (0, 1)$,

$$\begin{aligned} \|z - \bar{x}\| &\leq \|z - [(1-t)\bar{x} + tx]\| \\ &= \|z - \bar{x} + t(\bar{x} - x)\|. \end{aligned}$$

Assim,

$$\|z - \bar{x}\|^2 \leq \|z - \bar{x}\|^2 + 2t(z - \bar{x})^T(\bar{x} - x) + t^2 \|\bar{x} - x\|^2.$$

Logo

$$-(z - \bar{x})^T(\bar{x} - x) \leq \frac{t}{2} \|\bar{x} - x\|^2$$

e assim, para t suficientemente pequeno temos

$$(z - \bar{x})^T(x - \bar{x}) \leq 0,$$

para todo $x \in \Omega$. Agora, substituindo z na desigualdade acima obtemos,

$$\sum_{i \in I(\bar{x})} \bar{\lambda}_i \nabla g_i(\bar{x})^T (x - \bar{x}) \geq 0, \quad (2.28)$$

para todo $x \in \Omega$. Utilizando o fato de que $\sum_{i \in I(\bar{x})} \bar{\lambda}_i \nabla g_i(\bar{x}) = \nabla \psi(\bar{x})$ e a convexidade de ψ na desigualdade (2.28) concluímos que

$$0 \leq \nabla \psi(\bar{x})^T (x - \bar{x}) \leq \psi(x) - \psi(\bar{x}),$$

para todo $x \in \Omega$, isto é, \bar{x} é minimizador de ψ em Ω . □

Observações

- Se \bar{x} é degenerado, da desigualdade (2.28), temos que o conjunto

$$\{z \in \Omega \mid (z - \bar{x})^T \nabla g_i(\bar{x}) < 0, \text{ para todo } i \in I(\bar{x})\} = \emptyset.$$

Em outras palavras, não existem direções viáveis de descida para todas as funções g_i tais que $g_i(\bar{x}) \geq 0$.

- Os pontos que não satisfaçam (2.26) serão ditos pontos não degenerados. Note que pela contrapositiva do item anterior, temos que se um ponto \bar{x} satisfaz Mangasarian Fromovitz então \bar{x} é não degenerado.
- Reciprocamente, se um ponto é viável não degenerado então satisfaz a condição de qualificação de Mangasarian-Fromovitz [20, Proposição 4.10].

Vejam agora que o Algoritmo 2.2 está bem definido. Isto significa que enquanto o algoritmo não para, sempre é possível obter um ponto x^k solução do subproblema (2.24).

Lema 2.6 *O Algoritmo 2.2 está bem definido.*

Demonstração. O subproblema (2.24) que define a sequência $(x^k)_{k \in \mathbb{N}}$ é um problema de minimização de uma função contínua definida num conjunto compacto Ω . Logo suas soluções satisfazem (2.25), ou seja, o algoritmo está bem definido. \square

Discutimos a seguir a convergência global da classe de métodos de Lagrangiano aumentado considerada. Em [31] é discutida a prova de convergência global do método de Lagrangiano aumentado com a penalização PHR considerando a condição de qualificação CPLD. O próximo teorema garante a convergência global do Algoritmo 2.2 independente da função de penalização utilizada desde que esta função satisfaça as propriedades A1-A3 usando o conceito de ponto degenerado como vimos há pouco relacionado com a condição de qualificação de Mangasarian-Fromovitz.

Teorema 2.7 [8, Thm. 2] *Considere $(x^k)_{k \in \mathbb{N}}$ uma sequência gerada pelo Algoritmo 2.2 e x^* um ponto de acumulação da sequência. Então uma das seguintes possibilidades ocorre*

- (i) x^* é degenerado;
- (ii) x^* é um ponto estacionário do problema (2.22).

Demonstração. Considerando que a sequência $(\varepsilon^k)_{k \in \mathbb{N}}$ converge para zero, de (2.25) temos que

$$\lim_{k \rightarrow \infty} \left\| \mathcal{P} \left(x^k - \nabla f(x^k) - \sum_{i=1}^m P'(g_i(x^k), \bar{\mu}_i^k, \rho_i^k) \nabla g_i(x^k) \right) - x^k \right\| = 0, \quad (2.29)$$

pela atualização do multiplicador μ^k dada pelo Algoritmo 2.2 segue que

$$\lim_{k \rightarrow \infty} \left\| \mathcal{P} \left(x^k - \nabla f(x^k) - \sum_{i=1}^m \mu_i^k \nabla g_i(x^k) \right) - x^k \right\| = 0. \quad (2.30)$$

Por outro lado, como x^* é um ponto de acumulação da sequência $(x^k)_{k \in \mathbb{N}}$, então existe um conjunto de índices $\mathbb{K} \subset \mathbb{N}$ tal que $\lim_{k \in \mathbb{K}} x^k = x^*$. Deste modo, para o conjunto \mathbb{K} , existirão duas possibilidades:

- A sequência $(\mu^k)_{k \in \mathbb{K}}$ é ilimitada.

Neste caso, existe um conjunto infinito de índices $\mathbb{K}_1 \subset \mathbb{K}$ e $j \in \{1, \dots, m\}$ tal que $\|\mu^k\|_\infty = \mu_j^k \geq 1$ para todo $k \in \mathbb{K}_1$ e $\lim_{k \in \mathbb{K}_1} \mu_j^k = \infty$. Além disso, observe que para todo $x \in \Omega$, $v \in \mathbb{R}^n$, $t \in [0, 1]$

$$\|\mathcal{P}(x + tv) - x\| \leq \|\mathcal{P}(x + v) - x\|.$$

Assim, combinando o resultado acima e (2.30), considerando $x = x^k$, $v = -\nabla f(x^k) - \sum_{i=1}^m \mu_i^k \nabla g_i(x^k)$ e $t = 1/\mu_j^k$ obtemos que

$$\lim_{k \in \mathbb{K}_1} \left\| \mathcal{P} \left(x^k - \frac{1}{\mu_j^k} \nabla f(x^k) - \sum_{i=1}^m \frac{\mu_i^k}{\mu_j^k} \nabla g_i(x^k) \right) - x^k \right\| = 0. \quad (2.31)$$

A sequência $(\lambda^k)_{k \in \mathbb{K}_1}$ definida por $\lambda^k = \frac{\mu^k}{\mu_j^k}$ é limitada, logo admite uma sub-sequência convergente. Em outras palavras, existe um conjunto $\mathbb{K}_2 \subset \mathbb{K}_1$ tal que $\lim_{k \in \mathbb{K}_2} \lambda^k = \lambda$, isto é, $\lambda_i^k \rightarrow \lambda_i$ para cada $i \in \{1, \dots, m\}$, e além disso, observe que $\lambda_j = 1$. Logo, pela propriedade **A3**, pela continuidade de P' e pela atualização de μ^k , no algoritmo 2.2, para todos os índices $i \in \{1, \dots, m\}$ tais que $g_i(x^*) < 0$, temos que $(\mu_i^k)_{k \in \mathbb{K}_2}$ é limitada e além disso

$$\lim_{k \in \mathbb{K}_2} \frac{\mu_i^k}{\mu_j^k} = 0.$$

Portanto, tomando o limite em (2.31), sobre o conjunto \mathbb{K}_2 temos que

$$\mathcal{P} \left(x^* - \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) \right) - x^* = 0,$$

para $I(x^*) = \{i \mid g_i(x^*) \geq 0\}$. Como $\sum_{i \in I(x^*)} \lambda_i > 0$, pois $\lambda_j = 1$ concluímos que, x^* é degenerado.

- A sequência $(\mu^k)_{k \in \mathbb{K}}$ é limitada.

Pelo teorema de Bolzano-Weierstrass temos que existe um conjunto de índices $\mathbb{K}_3 \subset \mathbb{K}$ tal que $\lim_{k \in \mathbb{K}_3} \mu^k = \mu^* \geq 0$. Tomando limite em (2.30) sobre \mathbb{K}_3 temos que

$$\mathcal{P} \left(x^* - \nabla f(x^*) - \sum_{i=1}^m \mu_i^* \nabla g_i(x^*) \right) - x^* = 0.$$

Para cada $i \in 1, \dots, m$, considere as seguintes possibilidades:

- A sequência $(\rho_i^k)_{k \in \mathbb{K}_3}$ é limitada.

Pela atualização de ρ_i^k no algoritmo 2.2, temos que as condições $|g_i(x^k) \mu_i^k| \leq r |g_i(x^{k-1}) \mu_i^{k-1}|$ e $\max\{0, g_i(x^k)\} \leq r \max\{0, g_i(x^{k-1})\}$ são satisfeitas para $k \in \mathbb{K}_3$ suficientemente grande. Logo, $\lim_{k \in \mathbb{K}_3} |g_i(x^k) \mu_i^k| = \lim_{k \in \mathbb{K}_3} \max\{0, g_i(x^k)\} = 0$ e pela continuidade de g temos que $g_i(x^*) \leq 0$ e $\mu_i^* g_i(x^*) = 0$. Logo x^* é ponto estacionário do problema (2.22).

- A sequência $(\rho_i^k)_{k \in \mathbb{K}_3}$ é ilimitada.

Com esta hipótese temos que, $\rho_i^k \rightarrow \infty$, então se $g_i(x^*) > 0$ pela propriedade **A2** obtemos que

$$\lim_{k \in \mathbb{K}_3} P'(g_i(x^k), \bar{\mu}_i^k, \rho_i^k) = \infty,$$

assim, pela atualização no algoritmo 2.2, segue que a sequência μ^k é ilimitada, o que seria absurdo. Logo, obtemos que $g_i(x^*) \leq 0$, para todo $i \in \{1, \dots, m\}$. Se $g_i(x^*) < 0$ então pela propriedade **A3**,

$$\lim_{k \in \mathbb{K}_3} P'(g_i(x^k), \bar{\mu}_i^k, \rho_i^k) = 0$$

e por conseguinte, $\mu_i^* = 0$. Em outras palavras, a complementaridade é satisfeita. Assim, x^* é ponto estacionário. □

2.2.2 Diferentes funções de penalização

Vimos na seção anterior que o Algoritmo (2.2) é globalmente convergente independente da função de penalização usada na definição (2.23) do Lagrangiano aumentado. No artigo [8], os autores fazem um grande trabalho computacional testando o desempenho do método Lagrangiano aumentado com 65 funções de penalização diferentes. Neste conjunto de testes, o método Lagrangiano aumentado clássico, ou seja, com a função de penalidade (2.21), é o que obteve o melhor desempenho. Em [16] é proposta uma família de funções penalidade do tipo quadrática que inclui a penalidade do Rockafellar (PHR). Nesta família se consideram além da penalidade de Rockafellar (PHR) outras duas penalidades particulares que não foram consideradas nos testes realizados em [8] mas que em [16] são sugeridas como possíveis alternativas para serem implementadas.

Nesta seção discutimos estas funções de penalização e mostramos que satisfazem as propriedades **A1-A3**, o que garante a convergência global do Algoritmo 2.2. No próximo capítulo discutimos os testes computacionais com estas funções de penalização.

Considere as penalidades $P_0, P_1, P_2 : \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_{++} \rightarrow \mathbb{R}$ definidas por

$$P_0(y, t, s) = \frac{s}{2} \left(\max \left\{ 0, yt + \frac{1}{s} \right\}^2 - \frac{1}{s^2} \right) \quad (2.32)$$

$$P_1(y, t, s) = \frac{s}{2} t \left(\max \left\{ 0, y + \frac{1}{s} \right\}^2 - \frac{1}{s^2} \right) \quad (2.33)$$

$$P_2(y, t, s) = \frac{1}{2s} (\max \{0, t + sy\}^2 - t^2). \quad (2.34)$$

As duas primeiras penalidades são apresentadas em [16], enquanto P_2 é a clássica penalidade (PHR) [22, 36, 39] já discutida na Seção 2.1.2.

Discutiremos a seguir algumas relações e diferenças que satisfazem estas funções de penalização.

Considere $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ definida por $\varphi(w) = \frac{1}{2} (\max\{0, w + 1\}^2 - 1)$. As funções podem ser reescritas como segue:

$$P_0(y, t, s) = \frac{1}{s} \varphi(yts)$$

$$P_1(y, t, s) = \frac{t}{s} \varphi(y s)$$

$$P_2(y, t, s) = \frac{t^2}{s} \varphi\left(\frac{y s}{t}\right).$$

Note que a diferença entre cada uma destas funções está na posição do valor t . Note ainda que se consideramos $q \in \{0, 1, 2\}$, a função

$$P_q(y, t, s) = \frac{t^q}{s} \varphi(t^{1-q} y s)$$

gera cada uma das funções penalidade.

Outra característica importante está relacionada à forma como elas penalizam. Observe que se fixamos $t = 1$ em (2.32)-(2.34), as funções na variável y e s , coincidem, ou seja,

$$P_0(y, 1, s) = P_1(y, 1, s) = P_2(y, 1, s) = \frac{s}{2} \left(\max \left\{ 0, y + \frac{1}{s} \right\}^2 - \frac{1}{s^2} \right).$$

A Figura 2.2 exhibe o gráfico das três funções penalidade na variável y , quando fixado $t = s = 1$.

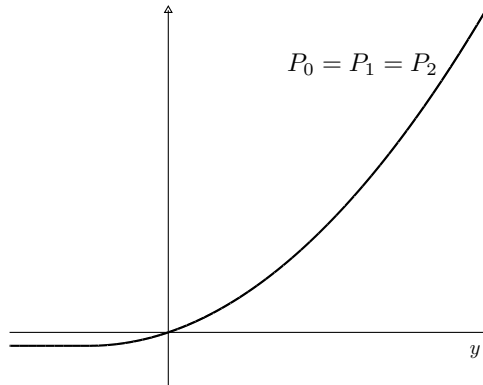


Figura 2.2: $P_0(y, 1, 1) = P_1(y, 1, 1) = P_2(y, 1, 1)$

No entanto, dependendo dos valores de s e t , ora uma função penaliza mais que as outras. Fixando-se $s = 1$ em (2.32)-(2.34), a Figura 2.3 mostra o comportamento das três funções na variável y . Na figura da esquerda, foi fixado $t = 1/2$, enquanto na da direita, $t = 2$. Note que na figura da esquerda, o gráfico da função penalidade $P_2(\cdot, t, 1)$ passa por cima de $P_1(\cdot, t, 1)$ que por sua vez passa por cima de $P_0(\cdot, t, 1)$. Quando $t > 1$, esta relação de ordem inverte-se como mostra a figura da direita.

Fixado $t = 1$, a Figura 2.4 mostra o comportamento das três funções na variável y , para diferentes valores de s . À medida que o valor de s cresce a penalização aumenta.

O próximo lema mostra que estas funções de penalização satisfazem as propriedades **A1-A3** para garantia da convergência do algoritmo.

Lema 2.8 *As funções penalidade definidas em (2.32), (2.33) e (2.34) satisfazem as propriedades **A1-A3**.*

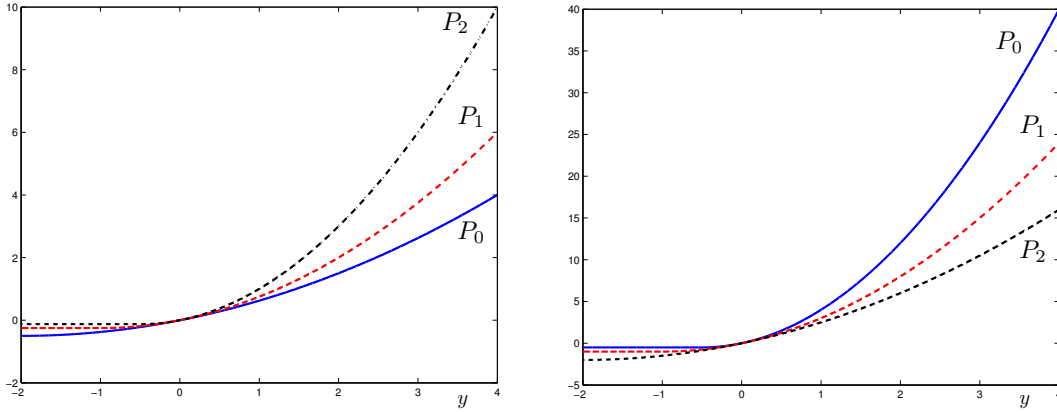


Figura 2.3: Fixado $s = 1$ com $t = 1/2$ (esquerda) e $t = 2$ (direita).

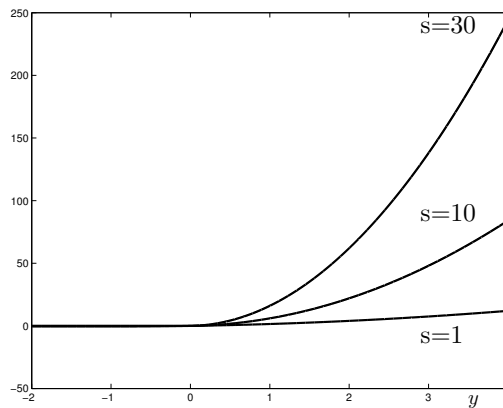


Figura 2.4: Fixado $t = 1$, gráfico das funções de penalização com diferentes valores de s .

Demonstração. Note inicialmente que as derivadas das funções penalidade com respeito à primeira variável são dadas por

$$P'_0(y, t, s) = t \max\{0, yst + 1\},$$

$$P'_1(y, t, s) = t \max\{0, ys + 1\},$$

$$P'_2(y, t, s) = \max\{0, ys + t\}.$$

Como $t \in \mathbb{R}_+$, temos que P'_0 , P'_1 e P'_2 são não negativas, ou seja vale a propriedade **A1**.

Agora suponha que $\lim_{k \rightarrow \infty} s^k = \infty$ e $t^k \in [t_{\min}, t_{\max}]$ com $0 < t_{\min} < t_{\max} < \infty$. Estudemos os seguintes casos:

(i) Se $\lim_{k \rightarrow \infty} y^k = y > 0$ temos que para k suficientemente grande,

$$\max\{0, y^k s^k t^k + 1\} = y^k s^k t^k + 1,$$

$$\max\{0, y^k s^k + 1\} = y^k s^k + 1,$$

$$\max\{0, y^k s^k + t^k\} = y^k s^k + t^k.$$

Logo

$$\lim_{k \rightarrow \infty} P'_0(y^k, t^k, s^k) = \lim_{k \rightarrow \infty} P'_1(y^k, t^k, s^k) = \lim_{k \rightarrow \infty} P'_2(y^k, t^k, s^k) = \infty,$$

ou seja, a propriedade **A2** é satisfeita.

(ii) Se $\lim_{k \rightarrow \infty} y^k = y < 0$ então para k suficientemente grande temos que

$$\max\{0, y^k s^k t^k + 1\} = \max\{0, y^k s^k + 1\} = \max\{0, t^k + s^k y^k\} = 0.$$

Assim $\lim_{k \rightarrow \infty} P'_i(y^k, t^k, s^k) = 0$, para cada $i \in \{1, 2, 3\}$, isto é, **A3** também é válido.

□

Como estas penalidades satisfazem as propriedades **A1-A3**, podemos utilizar o Algoritmo 2.2 para gerar uma sequência $(x^k)_{k \in \mathbb{N}}$ e com as hipóteses do Teorema 2.7 a convergência global do algoritmo está garantida.

Capítulo 3

Experimentos numéricos

No trabalho [8], foi comparado o desempenho computacional do Algoritmo 2.2 com 65 funções de penalidade diferentes. Naqueles testes o algoritmo com a função de penalidade (2.34), proposta por Powell-Hestenes-Rockafellar (PHR) em [22, 36, 39] venceu tanto em eficiência como em robustez. Neste capítulo, seguindo a metodologia adotada em [8], comparamos o desempenho do Algoritmo 2.2 com as funções de penalidade discutidas na Seção 2.2.2. Agradecemos aos autores do artigo [8] que gentilmente disponibilizaram o programa usado para os testes numéricos do referido artigo. A implementação do algoritmo, em Fortran, chama a rotina `ALGENCAN` [10, 31].

3.1 Metodologia

O algoritmo fornecido pelos autores de [8] foi modificado de modo a realizarmos a comparação do desempenho do Algoritmo 2.2 com as funções penalidade P_0 e P_1 discutidas na Seção 2.2.2 com a função penalidade P_2 que foi a vencedora na bateria de testes do artigo [8].

Neste capítulo consideramos três instâncias do Algoritmo 2.2, que diferem entre si pela função penalidade. Consideramos, então,

- P0 - Algoritmo 2.2 com função penalidade P_0 definida em (2.32).
- P1 - Algoritmo 2.2 com função penalidade P_1 definida em (2.33).
- P2 - Algoritmo 2.2 com função penalidade P_2 definida em (2.34).

Para os testes numéricos seguimos a mesma metodologia adotada em [8] como discutimos a seguir.

3.1.1 Conjunto de problemas-teste

Como em [8], para os testes numéricos, foi considerada a coleção de problemas `CUTEr` [23, 40]. Foram testados todos os problemas de coleção com restrições de desigualdade e de caixa, ou sejam, aqueles em que

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

Assim, consideramos todos os problemas da coleção `CUTEr` em que o conjunto viável é da forma

$$D = \{x \in \Omega \mid g(x) \leq 0\},$$

com $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ diferenciável.

Os problemas da coleção CUTEr vêm como uma classificação simples inspirada pelo esquema de Hock e Schittkowski [23].

Um problema é classificado por uma sequência de 8 caracteres:

XXXr-XX-N-M.

O primeiro caracter representa o tipo de função objetivo. As possíveis letras são:

N	Nenhuma função objetivo definida
C	Função objetivo constante
L	Função objetivo linear
Q	Função objetivo quadrática
S	Função objetivo dada por soma de quadrados
O	Função objetivo não está nas descrições acima.

O segundo caracter define o tipo de restrição mais complexa que envolve o problema. As possíveis letras são:

U	Problema irrestrito
X	Algumas componentes da variável são fixas
B	Restrições do tipo caixa
N	Restrições que representam a matriz de um grafo (Linear)
L	Restrições lineares
Q	Restrições quadráticas
O	Restrições mais gerais que qualquer uma acima

O terceiro caracter na sequência indica a suavidade do problema. Neste caso só tem dois caracteres possíveis:

R	Problema regular
I	Problema irregular.

Um problema é regular se as funções envolvidas no problema são de classe C^2 em todo ponto. Caso contrário o problema é dito irregular. O quarto caracter (r) só pode assumir os valores 0, 1 ou 2 e representa o maior grau da derivada fornecido na descrição do problema. O caracter após o primeiro hífen indica a origem principal e/ou a utilidade do problema. As possíveis letras são:

A	Problema acadêmico
M	Problema de modelagem
R	Problema que não foram desenvolvidos para testar algoritmos.

O próximo caracter indica se o problema contém ou não variáveis internas explícitas, assim, só existem dois valores possíveis:

Y	O problema contém variáveis explícitas
N	O problema não contém variáveis explícitas.

Finalmente, os dois últimos caracteres são representados por números inteiros os quais indicam o número de variáveis do problema (N) e o número de restrições (M) (exceto para variáveis fixadas) do problema. Observe que existe também um terceiro valor (V) o qual pode ser utilizado tanto no lugar de (N) e (M) e significa que o número de variáveis ou restrições é variável, podendo ser escolhidos pelo usuário.

Tendo em vista esta classificação, consideramos todos os problemas em que:

- O primeiro caracter do problema não é N nem C, ou seja, em que há uma função não constante a ser minimizada.
- O segundo caracter não é U, nem X, nem B, ou seja, em que há restrições de desigualdade além da restrição de caixa.

Assim todos os problemas de coleção com restrições de desigualdade e de caixa resultam num conjunto de 84 problemas com dimensão entre 1 e 800, e o número de restrições, além de caixa entre 1 e 2404.

3.1.2 Resolução dos subproblemas

Em cada iteração do Algoritmo 2.2 é preciso resolver um subproblema da forma (2.24), de minimização de uma função num conjunto Ω . Como nos problemas testados, Ω é uma caixa, foi utilizada a rotina **GENCAN** para a resolução dos subproblemas, como feito em [8].

A rotina **GENCAN** é uma implementação em **FORTTRAN** do algoritmo do conjunto ativo descrito em [10], para resolver problemas de otimização com restrições de caixa o qual para minimizar nas faces emprega um algoritmo tipo Newton truncado com busca linear, enquanto que para sair das faces emprega iterações do método gradiente espectral projetado definido em [11, 12].

Como uma solução x^k do subproblema deve satisfazer (2.25), o algoritmo **GENCAN** para quando tal critério é satisfeito com $\varepsilon^k = \varepsilon > 0$. No entanto às vezes esse critério pode ser muito exigente e impossível de ser atingido em ponto flutuante.

Portanto, tem que ser adotado um critério em **GENCAN** para diminuir este tipo de problema. Neste trabalho foram considerados os parâmetros descritos em [10], como o número máximo de iterações internas fixado em 100000.

3.1.3 Critérios numéricos de convergência

Pela maneira como o Algoritmo 2.2 foi descrito, ele não para. Ele não tem um critério de parada definido e gerará uma sequência infinita, ou seja, se x^k é solução do problema de programação não linear então o algoritmo calculará x^{k+1} e neste caso $x^{k+1} = x^k$ e assim de maneira indefinida. Por isso, os critérios que apresentaremos são definidos para estabelecer quando um ponto x^k é candidato a ser solução de um determinado problema e assim, computacionalmente, estabelecer um critério de parada prático.

Um ponto x^k associado à estimativa μ^k , do multiplicador será dito que satisfaz um critério ordinário de convergência (COC), se a viabilidade, otimalidade e complementaridade são válidos para uma tolerância pequena. Em outras palavras, dado $\varepsilon > 0$, dizemos que x^k com a estimativa μ^k do multiplicador de Lagrange, satisfaz COC se

$$\left\| \mathcal{P} \left(x^k - \nabla f(x^k) - \sum_{i=1}^m \mu_i^k \nabla g_i(x^k) \right) - x^k \right\| \leq \varepsilon, \quad (3.1)$$

$$g_i(x^k) \leq \varepsilon, \quad i = 1 \dots m, \quad (3.2)$$

$$\max_{i=1, \dots, m} \{ -g_i(x^k) \mu_i^k \mid g_i(x^k) < -\varepsilon \text{ e } \mu_i^k > \varepsilon \} \leq \varepsilon. \quad (3.3)$$

A desigualdade (3.1) corresponde ao critério de parada (2.25) para o subproblema (2.24), pois pela atualização do multiplicador μ^k no Algoritmo 2.2 temos que (3.1) é satisfeito se e somente se,

$$\| \mathcal{P} (x^k - \nabla L(x^k, \bar{\mu}^k, \rho^k)) - x^k \| \leq \varepsilon.$$

A desigualdade (3.2) diz que o ponto é quase viável, com tolerância ε . Finalmente (3.3) é uma versão aproximada da complementaridade com tolerância ε .

Estas condições (3.1)-(3.3) representam condições de otimalidade aproximadas para o problema (2.22), no sentido de que, se estas condições são satisfeitas com $\varepsilon = 0$, isto implica que as condições de KKT são satisfeitas em x^k . Estes critérios de convergência aproximada não são completamente seguros, pois eles são muito sensíveis ao escalamento das variáveis, da função objetivo e das restrições. Em problemas específicos, o usuário pode sentir a necessidade de usar tolerâncias diferentes no lado direito das desigualdades (3.1)-(3.3). Em nossos testes, utilizamos para todos os problemas, $\varepsilon = 10^{-4}$, ou seja, o algoritmo declarou ter obtido uma solução quando as condições (3.1)-(3.3) foram satisfeitas com tal tolerância.

3.1.4 Atualização do parâmetro de penalidade

O vetor dos parâmetros de penalidade ρ é atualizado de acordo com o critério definido no Algoritmo 2.2. A ideia é que o parâmetro de penalidade ρ_i não seja trocado entre iterações nas quais é produzido uma melhora entre a viabilidade e complementaridade. Assim, é considerado o seguinte critério prático para atualizar os parâmetros de penalidade o qual é compatível com o critério ordinário de convergência (COC) definido na seção anterior.

Dado $\varepsilon > 0$, definimos em cada iteração k feita do Algoritmo 2.2:

$$\begin{aligned} F_i^k &= \max\{0, g_i(x^k)\}, \\ U_i^k &= \begin{cases} \mu_i^k, & \text{se } g_i(x^k) < -\varepsilon \\ 0, & \text{caso contrário,} \end{cases} \\ W_i^k &= \begin{cases} -g_i(x^k), & \text{se } \mu_i^k > \varepsilon \\ 0, & \text{caso contrário,} \end{cases} \end{aligned}$$

para cada $i = 1, \dots, m$.

Observe que (3.2) e (3.3) da definição COC correspondem a

$$\|F_k\|_\infty \leq \varepsilon \quad \text{e} \quad \max_{i=1, \dots, m} \{|U_i^k W_i^k|\} \leq \varepsilon,$$

respectivamente. Assim, a atualização do parâmetro de penalidade no Algoritmo 2.2 em termos de F^k , U^k e W^k é dada por:

Para cada $i = 1 \dots, m$, se

$$F_i^k \leq r F_i^{k-1} \quad \text{e} \quad |U_i^k W_i^k| \leq r |U_i^{k-1} W_i^{k-1}|$$

então

$$\rho_i^{k+1} = \rho_i^k,$$

senão

$$\rho_i^{k+1} = \gamma \rho_i^k.$$

3.1.5 Critério de seleção

Ao resolver um determinado problema, precisamos de alguma maneira comparar as soluções obtidas por diferentes métodos para assim, escolher os métodos mais eficientes.

Considere um problema fixado e seja $x_{final}^{(M)}$ a solução encontrada por um método M aplicado a este problema. Definimos f_{melhor} como sendo o menor valor encontrado pelo conjunto dos métodos testados, ou seja,

$$f_{melhor} = \min_M \{f(x_{final}^{(M)}) \mid x_{final}^{(M)} \text{ é viável.}\}.$$

Dizemos que um método resolveu o problema se $x_{final}^{(M)}$ é viável e

$$f(x_{final}^{(M)}) \leq f_{melhor} + 10^{-3} |f_{melhor}| + 10^{-6}.$$

Além disso, considere $t^{(M)}$, o tempo de CPU que o método M gastou para chegar ao ponto $x_{final}^{(M)}$. Assim, definimos t_{melhor} como o menor tempo computacional gasto entre os métodos testados que resolveram o problema, ou seja,

$$t_{melhor} = \min_M \{t^{(M)} \mid \text{método } M \text{ encontrou uma solução}\}.$$

Dizemos que um determinado método M é mais rápido para um problema quando

$$t^{(M)} \leq t_{melhor} + 0.01 t_{melhor}.$$

Agora tendo como base todo o conjunto de problemas testados, para cada método definimos os seguintes índices:

- **Robustez** como sendo a porcentagem dos problemas para o qual o método encontrou uma solução.
- **Viabilidade** como sendo a porcentagem dos problemas para o qual o método encontrou um ponto viável.
- **Eficiência** como sendo a porcentagem dos problemas para o qual o método foi o mais rápido.

3.2 Testes numéricos

Para os experimentos numéricos foi utilizado um computador com 2 GB de RAM e um processador dual core de 2.1 GHz. Os cálculos em FORTRAN foram compilados usando GNU Fortran.

Nosso objetivo é comparar o desempenho algorítmico das três instâncias P0, P1 e P2 do Algoritmo 2.2.

3.2.1 Escolha dos Parâmetros

Uma primeira tarefa é determinar o melhor conjunto de parâmetros iniciais μ_0 , γ , ρ_1 , r , μ_{\min} e μ_{\max} para cada uma das instâncias. Seguindo a metodologia adotada em [8], foram consideradas as seguintes possibilidades para estes parâmetros:

$$\begin{aligned}\mu_0 &\in \{10^{-6}, 1\}, \\ \gamma &\in \{2, 10, 100\}, \\ \rho_1 &\in \{10^{-6}, 1, 10\}, \\ r &\in \{10^{-2}, 0.1, 0.5\}, \\ (\mu_{\min}, \mu_{\max}) &\in \{(10^{-6}, 10^{20}), (10^{-6}, 10^6), (10^{-3}, 10^3)\},\end{aligned}$$

o que gera $2 \times 3^4 = 162$ combinações possíveis. A relação das 162 combinações encontra-se na Tabela A.1 no Apêndice A.

A melhor combinação de parâmetros iniciais para a penalidade P_2 foi adotada a escolhida em [8], ou seja, a 69ª combinação em que

$$\mu_0 = 10^{-6}, \quad \gamma = 10, \quad \rho_1 = 10, \quad r = 0.1, \quad \mu_{\min} = 10^{-3}, \quad \mu_{\max} = 10^3.$$

Para escolher a melhor combinação possível para as outras duas penalidades P_0 e P_1 , seguimos o metodologia sugerida em [8]. Escolhemos um subconjunto de 46 problemas do conjunto de 84 problemas teste da coleção CUTEr. O Algoritmo 2.2 com cada uma destas penalidades e com cada uma das 162 combinações de parâmetros foi considerado como um método diferente. Ou seja, para cada instância P0 e P1 rodamos este subconjunto de problemas com cada uma das 162 combinações de parâmetros. Com isto foi gerada uma tabela com 14904 linhas ($162 \text{ combinações} \times 2 \text{ instâncias} \times 46 \text{ problemas} = 14904$), cujas colunas fornecem o valor ótimo e uma medida de viabilidade na solução encontrada e o respectivo tempo de CPU gasto para obter cada solução. Foram então calculado o índice de robustez e de eficiência de cada um destes métodos, segundo o que discutimos na seção anterior.

Com estes resultados, concluímos que para a instância P0 o melhor índice de robustez foi 76,09% obtido com a 111ª combinação de parâmetros dada por

$$\mu_0 = 1, \quad \gamma = 2, \quad \rho_1 = 1, \quad r = 10^{-2}, \quad \mu_{\min} = 10^{-3}, \quad \mu_{\max} = 10^3.$$

Enquanto o melhor índice de eficiência foi de 34,78% obtido pela 154ª combinação, dada por

$$\mu_0 = 1, \quad \gamma = 100, \quad \rho_1 = 10, \quad r = 10^{-2}, \quad \mu_{\min} = 10^{-6}, \quad \mu_{\max} = 10^{20}.$$

Para a instância P1, o melhor índice de robustez foi de 80,43% obtido com a 147ª combinação, dada por

$$\mu_0 = 1, \quad \gamma = 10, \quad \rho_1 = 10, \quad r = 10^{-2}, \quad \mu_{\min} = 10^{-3}, \quad \mu_{\max} = 10^3.$$

E o melhor índice de eficiência foi de 30,43% obtido com a 66ª combinação, dada por

$$\mu_0 = 10^{-6}, \quad \gamma = 10, \quad \rho_1 = 10, \quad r = 10^{-2}, \quad \mu_{\min} = 10^{-3}, \quad \mu_{\max} = 10^3.$$

A próxima etapa é resolver todo o conjunto de 84 problemas pelas três instâncias com a seleção de parâmetros discutida nesta seção.

3.2.2 Resultados numéricos

Tendo em vista os resultados da seção anterior, as três instâncias P0, P1, P2 consideradas na Seção 3.1 serão tomadas para resolver todo o conjunto de 84 problemas-teste da coleção CUTeR e cada método a ser comparado é tomado como segue

- (P0,111)– Algoritmo 2.2 com função penalidade (2.32) com a 111ª combinação de parâmetros iniciais.
- (P0,154)– Algoritmo 2.2 com função penalidade (2.32) com a 154ª combinação de parâmetros iniciais.
- (P1,147)– Algoritmo 2.2 com função penalidade (2.33) com a 147ª combinação de parâmetros iniciais.
- (P1,66)– Algoritmo 2.2 com função penalidade (2.33) com a 66ª combinação de parâmetros iniciais.
- (P2,69) Algoritmo 2.2 com função penalidade (2.34) com a 69ª combinação de parâmetros iniciais.

No Apêndice B, apresentamos a Tabela B.1 com os resultados numéricos. A primeira coluna indica o nome do problema como aparece na coleção CUTeR, a segunda coluna mostra o número de variáveis do problema, a terceira coluna mostra o número de restrições associadas ao problema além da restrição de caixa, a quarta e quinta coluna indicam as instâncias implementadas associada à combinação de parâmetro com melhor desempenho obtida na seção anterior. As últimas duas colunas mostram o tempo de gasto do CPU e o valor ótimo obtido por cada método no problema indicado.

Tendo então resolvido os 84 problemas com cada um dos cinco métodos, calculamos as porcentagens de robustez, viabilidade e eficiência como descritas na Seção 3.1.5. A Tabela 3.1 mostra estes resultados.

Instância	Robustez	Viabilidade	Eficiência
(P0,111)	73,81	83,33	20,24
(P0,154)	73,81	78,57	27,38
(P1,147)	79,76	84,52	28,57
(P1,66)	76,19	80,95	27,38
(P2,69)	80,85	88,10	47,62

Tabela 3.1: Porcentagem de problemas por instância.

A Figura 3.1 mostra o gráfico do perfil de desempenho, como proposto em [18], em relação ao tempo computacional. A diferença entre a figuras da esquerda e da direita está no limitante do eixo horizontal. Na figura da esquerda, vemos que a penalidade

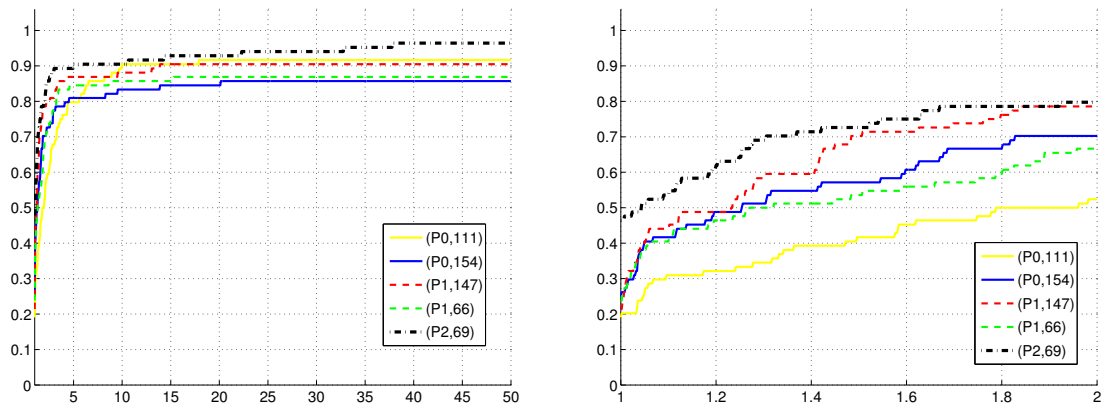


Figura 3.1: Gráfico de perfil de desempenho

quadrática clássica P_2 foi a mais robusta. Na figura da direita, percebemos que ela foi a mais eficiente ganhando em 47,62% dos problemas, condizente com a informação da Tabela 3.1.

As informações da Tabela 3.1 e da Figura 3.1 confirmam as conclusões extraídas em [8]. O Algoritmo 2.2 com a penalidade clássica (2.34) conhecida como PHR ganha tanto em eficiência como em robustez. As diferenças são mais expressivas sobretudo em relação à eficiência do método como salientado em [8]. Isto parece ser devido sobretudo pela simplicidade da expressão da função penalidade P_2 .

Conclusão

Neste trabalho estudamos uma classe de métodos de Lagrangiano aumentado para problemas de minimização de uma função diferenciável, num conjunto convexo e compacto, com restrições de desigualdade. Os métodos desta classe diferenciam entre si na função de penalização adotada. Sob hipóteses razoáveis como a condição de qualificação de Mangasarian-Fromovitz, a convergência global é provada desde que as penalidades satisfaçam algumas condições. Além da clássica penalização quadrática, exibimos outras duas funções penalidades, sugeridas em [16], que verificam as condições exigidas. Foram realizados então testes numéricos para comparar o desempenho computacional do método com cada uma destas três funções de penalização. Seguindo a metodologia de [8], foram escolhidas, para cada uma das duas funções penalidades exibidas, as duas melhores combinações de parâmetros iniciais entre 162 possibilidades. Foram então comparadas cinco instâncias do algoritmo de Lagrangiano aumentado, ou sejam, com a penalização quadrática e a combinação sugerida em [8], e com as outras duas penalidades, cada uma delas com os dois conjuntos de parâmetros iniciais selecionados. Os testes foram feitos em **Fortran**, com a implementação do algoritmo cedida pelos autores de [8], e com os subproblemas sendo resolvidos pelo **Algencan** [10, 31]. Foram considerados todos os 84 problemas da coleção **CUTEr** [23, 40], que envolvem restrições de desigualdade e de caixa em todas as variáveis. Ainda que as duas funções de penalidade exibidas neste trabalho não tivessem sido contempladas nos testes apresentados em [8], a conclusão de nosso trabalho coincide com a do artigo. O método de Lagrangiano aumentado com a penalização quadrática clássica (PHR) foi melhor em robustez e sobretudo em eficiência.

Apêndice A

Combinações dos parâmetros

Na Seção 3.2, referente aos testes numéricos, comentamos que os algoritmos foram testados com 162 combinações dos parâmetros iniciais μ_0 , γ , ρ_1 , r , μ_{\min} e μ_{\max} , com

$$\begin{aligned}\mu_0 &\in \{10^{-6}, 1\}, \\ \gamma &\in \{2, 10, 100\}, \\ \rho_1 &\in \{10^{-6}, 1, 10\}, \\ r &\in \{10^{-2}, 0.1, 0.5\}, \\ (\mu_{\min}, \mu_{\max}) &\in \{(10^{-6}, 10^{20}), (10^{-6}, 10^6), (10^{-3}, 10^3)\}.\end{aligned}$$

A Tabela A.1 mostra a relação destas combinações. A primeira coluna indica o número da combinação e as demais colunas os valores dos parâmetros μ_0 , γ , ρ_1 , r , μ_{\min} e μ_{\max} .

Combinação	μ^0	ρ^1	γ	r	μ_{\min}	μ_{\max}
1	10^{-6}	10^{-3}	2	10^{-2}	10^{-6}	10^{20}
2	10^{-6}	10^{-3}	2	10^{-2}	10^{-6}	10^6
3	10^{-6}	10^{-3}	2	10^{-2}	10^{-3}	10^3
4	10^{-6}	10^{-3}	2	0.1	10^{-6}	10^{20}
5	10^{-6}	10^{-3}	2	0.1	10^{-6}	10^6
6	10^{-6}	10^{-3}	2	0.1	10^{-3}	10^3
7	10^{-6}	10^{-3}	2	0.5	10^{-6}	10^{20}
8	10^{-6}	10^{-3}	2	0.5	10^{-6}	10^6
9	10^{-6}	10^{-3}	2	0.5	10^{-3}	10^3
10	10^{-6}	10^{-3}	10	10^{-2}	10^{-6}	10^{20}
11	10^{-6}	10^{-3}	10	10^{-2}	10^{-6}	10^6
12	10^{-6}	10^{-3}	10	10^{-2}	10^{-3}	10^3
13	10^{-6}	10^{-3}	10	0.1	10^{-6}	10^{20}
14	10^{-6}	10^{-3}	10	0.1	10^{-6}	10^6
15	10^{-6}	10^{-3}	10	0.1	10^{-3}	10^3
16	10^{-6}	10^{-3}	10	0.5	10^{-6}	10^{20}
17	10^{-6}	10^{-3}	10	0.5	10^{-6}	10^6
18	10^{-6}	10^{-3}	10	0.5	10^{-3}	10^3
19	10^{-6}	10^{-3}	100	10^{-2}	10^{-6}	10^{20}

Tabela A.1: Relação das combinações de parâmetro testadas.

(Continua na página seguinte.)

Combinação	μ^0	ρ^1	γ	r	μ_{\min}	μ_{\max}
20	10^{-6}	10^{-3}	100	10^{-2}	10^{-6}	10^6
21	10^{-6}	10^{-3}	100	10^{-2}	10^{-3}	10^3
22	10^{-6}	10^{-3}	100	0.1	10^{-6}	10^{20}
23	10^{-6}	10^{-3}	100	0.1	10^{-6}	10^6
24	10^{-6}	10^{-3}	100	0.1	10^{-3}	10^3
25	10^{-6}	10^{-3}	100	0.5	10^{-6}	10^{20}
26	10^{-6}	10^{-3}	100	0.5	10^{-6}	10^6
27	10^{-6}	10^{-3}	100	0.5	10^{-3}	10^3
28	10^{-6}	1	2	10^{-2}	10^{-6}	10^{20}
29	10^{-6}	1	2	10^{-2}	10^{-6}	10^6
30	10^{-6}	1	2	10^{-2}	10^{-3}	10^3
31	10^{-6}	1	2	0.1	10^{-6}	10^{20}
32	10^{-6}	1	2	0.1	10^{-6}	10^6
33	10^{-6}	1	2	0.1	10^{-3}	10^3
34	10^{-6}	1	2	0.5	10^{-6}	10^{20}
35	10^{-6}	1	2	0.5	10^{-6}	10^6
36	10^{-6}	1	2	0.5	10^{-3}	10^3
37	10^{-6}	1	10	10^{-2}	10^{-6}	10^{20}
38	10^{-6}	1	10	10^{-2}	10^{-6}	10^6
39	10^{-6}	1	10	10^{-2}	10^{-3}	10^3
40	10^{-6}	1	10	0.1	10^{-6}	10^{20}
41	10^{-6}	1	10	0.1	10^{-6}	10^6
42	10^{-6}	1	10	0.1	10^{-3}	10^3
43	10^{-6}	1	10	0.5	10^{-6}	10^{20}
44	10^{-6}	1	10	0.5	10^{-6}	10^6
45	10^{-6}	1	10	0.5	10^{-3}	10^3
46	10^{-6}	1	100	10^{-2}	10^{-6}	10^{20}
47	10^{-6}	1	100	10^{-2}	10^{-6}	10^6
48	10^{-6}	1	100	10^{-2}	10^{-3}	10^3
49	10^{-6}	1	100	0.1	10^{-6}	10^{20}
50	10^{-6}	1	100	0.1	10^{-6}	10^6
51	10^{-6}	1	100	0.1	10^{-3}	10^3
52	10^{-6}	1	100	0.5	10^{-6}	10^{20}
53	10^{-6}	1	100	0.5	10^{-6}	10^6
54	10^{-6}	1	100	0.5	10^{-3}	10^3
55	10^{-6}	10	2	10^{-2}	10^{-6}	10^{20}
56	10^{-6}	10	2	10^{-2}	10^{-6}	10^6
57	10^{-6}	10	2	10^{-2}	10^{-3}	10^3
58	10^{-6}	10	2	0.1	10^{-6}	10^{20}
59	10^{-6}	10	2	0.1	10^{-6}	10^6
60	10^{-6}	10	2	0.1	10^{-3}	10^3
61	10^{-6}	10	2	0.5	10^{-6}	10^{20}
62	10^{-6}	10	2	0.5	10^{-6}	10^6
63	10^{-6}	10	2	0.5	10^{-3}	10^3
64	10^{-6}	10	10	10^{-2}	10^{-6}	10^{20}
65	10^{-6}	10	10	10^{-2}	10^{-6}	10^6
66	10^{-6}	10	10	10^{-2}	10^{-3}	10^3

(Continua na página seguinte.)

Combinação	μ^0	ρ^1	γ	r	μ_{\min}	μ_{\max}
67	10^{-6}	10	10	0.1	10^{-6}	10^{20}
68	10^{-6}	10	10	0.1	10^{-6}	10^6
69	10^{-6}	10	10	0.1	10^{-3}	10^3
70	10^{-6}	10	10	0.5	10^{-6}	10^{20}
71	10^{-6}	10	10	0.5	10^{-6}	10^6
72	10^{-6}	10	10	0.5	10^{-3}	10^3
73	10^{-6}	10	100	10^{-2}	10^{-6}	10^{20}
74	10^{-6}	10	100	10^{-2}	10^{-6}	10^6
75	10^{-6}	10	100	10^{-2}	10^{-3}	10^3
76	10^{-6}	10	100	0.1	10^{-6}	10^{20}
77	10^{-6}	10	100	0.1	10^{-6}	10^6
78	10^{-6}	10	100	0.1	10^{-3}	10^3
79	10^{-6}	10	100	0.5	10^{-6}	10^{20}
80	10^{-6}	10	100	0.5	10^{-6}	10^6
81	10^{-6}	10	100	0.5	10^{-3}	10^3
82	1	10^{-3}	2	10^{-2}	10^{-6}	10^{20}
83	1	10^{-3}	2	10^{-2}	10^{-6}	10^6
84	1	10^{-3}	2	10^{-2}	10^{-3}	10^3
85	1	10^{-3}	2	0.1	10^{-6}	10^{20}
86	1	10^{-3}	2	0.1	10^{-6}	10^6
87	1	10^{-3}	2	0.1	10^{-3}	10^3
88	1	10^{-3}	2	0.5	10^{-6}	10^{20}
89	1	10^{-3}	2	0.5	10^{-6}	10^6
90	1	10^{-3}	2	0.5	10^{-3}	10^3
91	1	10^{-3}	10	10^{-2}	10^{-6}	10^{20}
92	1	10^{-3}	10	10^{-2}	10^{-6}	10^6
93	1	10^{-3}	10	10^{-2}	10^{-3}	10^3
94	1	10^{-3}	10	0.1	10^{-6}	10^{20}
95	1	10^{-3}	10	0.1	10^{-6}	10^6
96	1	10^{-3}	10	0.1	10^{-3}	10^3
97	1	10^{-3}	10	0.5	10^{-6}	10^{20}
98	1	10^{-3}	10	0.5	10^{-6}	10^6
99	1	10^{-3}	10	0.5	10^{-3}	10^3
100	1	10^{-3}	100	10^{-2}	10^{-6}	10^{20}
101	1	10^{-3}	100	10^{-2}	10^{-6}	10^6
102	1	10^{-3}	100	10^{-2}	10^{-3}	10^3
103	1	10^{-3}	100	0.1	10^{-6}	10^{20}
104	1	10^{-3}	100	0.1	10^{-6}	10^6
105	1	10^{-3}	100	0.1	10^{-3}	10^3
106	1	10^{-3}	100	0.5	10^{-6}	10^{20}
107	1	10^{-3}	100	0.5	10^{-6}	10^6
108	1	10^{-3}	100	0.5	10^{-3}	10^3
109	1	1	2	10^{-2}	10^{-6}	10^{20}
110	1	1	2	10^{-2}	10^{-6}	10^6
111	1	1	2	10^{-2}	10^{-3}	10^3
112	1	1	2	0.1	10^{-6}	10^{20}
113	1	1	2	0.1	10^{-6}	10^6
114	1	1	2	0.1	10^{-3}	10^3
115	1	1	2	0.5	10^{-6}	10^{20}

(Continua na página seguinte.)

Combinação	μ^0	ρ^1	γ	r	μ_{\min}	μ_{\max}
116	1	1	2	0.5	10^{-6}	10^6
117	1	1	2	0.5	10^{-3}	10^3
118	1	1	10	10^{-2}	10^{-6}	10^{20}
119	1	1	10	10^{-2}	10^{-6}	10^6
120	1	1	10	10^{-2}	10^{-3}	10^3
121	1	1	10	0.1	10^{-6}	10^{20}
122	1	1	10	0.1	10^{-6}	10^6
123	1	1	10	0.1	10^{-3}	10^3
124	1	1	10	0.5	10^{-6}	10^{20}
125	1	1	10	0.5	10^{-6}	10^6
126	1	1	10	0.5	10^{-3}	10^3
127	1	1	100	10^{-2}	10^{-6}	10^{20}
128	1	1	100	10^{-2}	10^{-6}	10^6
129	1	1	100	10^{-2}	10^{-3}	10^3
130	1	1	100	0.1	10^{-6}	10^{20}
131	1	1	100	0.1	10^{-6}	10^6
132	1	1	100	0.1	10^{-3}	10^3
133	1	1	100	0.5	10^{-6}	10^{20}
134	1	1	100	0.5	10^{-6}	10^6
135	1	1	100	0.5	10^{-3}	10^3
136	1	10	2	10^{-2}	10^{-6}	10^{20}
137	1	10	2	10^{-2}	10^{-6}	10^6
138	1	10	2	10^{-2}	10^{-3}	10^3
139	1	10	2	0.1	10^{-6}	10^{20}
140	1	10	2	0.1	10^{-6}	10^6
141	1	10	2	0.1	10^{-3}	10^3
142	1	10	2	0.5	10^{-6}	10^{20}
143	1	10	2	0.5	10^{-6}	10^6
144	1	10	2	0.5	10^{-3}	10^3
145	1	10	10	10^{-2}	10^{-6}	10^{20}
146	1	10	10	10^{-2}	10^{-6}	10^6
147	1	10	10	10^{-2}	10^{-3}	10^3
148	1	10	10	0.1	10^{-6}	10^{20}
149	1	10	10	0.1	10^{-6}	10^6
150	1	10	10	0.1	10^{-3}	10^3
151	1	10	10	0.5	10^{-6}	10^{20}
152	1	10	10	0.5	10^{-6}	10^6
153	1	10	10	0.5	10^{-3}	10^3
154	1	10	100	10^{-2}	10^{-6}	10^{20}
155	1	10	100	10^{-2}	10^{-6}	10^6
156	1	10	100	10^{-2}	10^{-3}	10^3
157	1	10	100	0.1	10^{-6}	10^{20}
158	1	10	100	0.1	10^{-6}	10^6
159	1	10	100	0.1	10^{-3}	10^3
160	1	10	100	0.5	10^{-6}	10^{20}
161	1	10	100	0.5	10^{-6}	10^6
162	1	10	100	0.5	10^{-3}	10^3

(Fim da Tabela A.1.)

Apêndice B

Resultados numéricos

Neste apêndice, apresentamos a tabela com os resultados numéricos obtidos ao testar o Algoritmo 2.2 com as penalidades P_0 , P_1 e P_2 , com a melhor combinação encontrada de parâmetros iniciais na resolução do conjunto de todos os 86 problemas-teste da coleção CUTEr que envolvem restrições de desigualdade e de caixa. Como discutido na Seção 3.2, foram comparadas as seguintes instâncias do algoritmo:

- (P0, 111) – Algoritmo 2.2 com função penalidade (2.32) com combinação de parâmetros iniciais 111;
- (P0, 154) – Algoritmo 2.2 com função penalidade (2.32) com combinação de parâmetros iniciais 154;
- (P1, 147) – Algoritmo 2.2 com função penalidade (2.33) com combinação de parâmetros iniciais 147;
- (P1, 66) – Algoritmo 2.2 com função penalidade (2.33) com combinação de parâmetros iniciais 66;
- (P2, 69) Algoritmo 2.2 com função penalidade (2.34) com combinação de parâmetros iniciais 69;

onde a combinação de parâmetros iniciais está descrita na Tabela A.1.

A primeira coluna da Tabela B.1 indica o nome do problema como aparece na coleção CUTEr. A segunda coluna mostra o número de variáveis do problema, a terceira coluna mostra o número de restrições associadas ao problema além da restrição de caixa, a quarta e quinta coluna indicam as instâncias implementadas associadas à combinação de parâmetro com melhor desempenho. As últimas duas colunas mostram o tempo de CPU e o valor ótimo obtido por cada instância no problema indicado.

Problema	n	m	pen	comb	CPU	f^*
AIRPORT	84	42	P0	111	0.2021726221	4.7952217D + 04
AIRPORT	84	42	P0	154	0.4331070781	4.7952702D + 04
AIRPORT	84	42	P1	147	0.1880517602	4.7952491D + 04
AIRPORT	84	42	P1	66	0.1508493870	4.7952311D + 04
AIRPORT	84	42	P2	69	0.2789773643	4.7952628D + 04
AVGASA	8	10	P0	111	0.0020801302	-4.6319277D + 00
AVGASA	8	10	P0	154	0.0009680600	-4.6318599D + 00
AVGASA	8	10	P1	147	0.0007720490	-4.6318802D + 00
AVGASA	8	10	P1	66	0.0014400904	-4.6319259D + 00
AVGASA	8	10	P2	69	0.0008160510	-4.6321105D + 00
AVGASB	8	10	P0	111	0.0020281260	-4.4832170D + 00
AVGASB	8	10	P0	154	0.0011080699	-4.4831552D + 00
AVGASB	8	10	P1	147	0.0009880612	-4.4832001D + 00
AVGASB	8	10	P1	66	0.0019521217	-4.4832418D + 00
AVGASB	8	10	P2	69	0.0007840490	-4.4832334D + 00
BIGGSC4	4	13	P0	111	0.0013720850	-2.4499945D + 01
BIGGSC4	4	13	P0	154	0.0005440340	-2.4500039D + 01
BIGGSC4	4	13	P1	147	0.0007880501	-2.4500008D + 01
BIGGSC4	4	13	P1	66	0.0018721170	-2.4500000D + 01
BIGGSC4	4	13	P2	69	0.0008880548	-2.4500037D + 01
BURKEHAN	1	1	P0	111	0.0000000000	-3.0268161D - 07
BURKEHAN	1	1	P0	154	0.0000000000	-4.6852329D - 10
BURKEHAN	1	1	P1	147	0.0000000000	-4.0972992D - 05
BURKEHAN	1	1	P1	66	0.0000000000	-3.5410021D - 05
BURKEHAN	1	1	P2	69	0.0000000000	-4.5454682D - 05
CANTILVR	5	1	P0	111	0.0010960681	1.3400479D + 00
CANTILVR	5	1	P0	154	0.0008320520	1.3400204D + 00
CANTILVR	5	1	P1	147	0.0008800551	1.3399568D + 00
CANTILVR	5	1	P1	66	0.0009160569	1.3399565D + 00
CANTILVR	5	1	P2	69	0.0008680551	1.3399530D + 00
CRESC100	6	200	P0	111	19.4212150574	1.0519730D + 00
CRESC100	6	200	P0	154	7.0188384056	5.6925546D - 01
CRESC100	6	200	P1	147	4.1850614548	5.6945051D - 01
CRESC100	6	200	P1	66	8.5977382660	5.6761517D - 01
CRESC100	6	200	P2	69	5.2455277443	5.6889125D - 01
CRESC4	6	8	P0	111	1.6000989676	1.9606762D + 00
CRESC4	6	8	P0	154	0.2648565173	8.7189792D - 01
CRESC4	6	8	P1	147	0.1077267677	8.7189847D - 01
CRESC4	6	8	P1	66	0.4120254517	6.9052940D - 01
CRESC4	6	8	P2	69	0.0844172835	8.7186424D - 01
CRESC50	6	100	P0	111	7.6188764572	1.0808155D + 00
CRESC50	6	100	P0	154	9.5821971893	6.0419056D - 01
CRESC50	6	100	P1	147	4.0838561058	5.9772933D - 01
CRESC50	6	100	P1	66	2.3961486816	7.8596217D - 09
CRESC50	6	100	P2	69	8.7357454300	6.0143083D - 01
DEMBO7	16	21	P0	111	3.1241948605	2.7992475D + 02
DEMBO7	16	21	P0	154	2.6801674366	2.3200484D + 02
DEMBO7	16	21	P1	147	0.6490405798	1.7491045D + 02
DEMBO7	16	21	P1	66	3.4042129517	1.7591538D + 02
DEMBO7	16	21	P2	69	3.7118325233	1.7482449D + 02

Tabela B.1: Resultados numéricos

Problema	n	m	pen	comb	CPU	f^*
EQC	9	3	P0	111	0.0003400210	-8.2954771D + 02
EQC	9	3	P0	154	0.0000564035	-8.2954771D + 02
EQC	9	3	P1	147	0.0000560035	-8.2954771D + 02
EQC	9	3	P1	66	0.0000560035	-8.2954771D + 02
EQC	9	3	P2	69	0.0000556035	-8.2954771D + 02
HATFLDH	4	13	P0	111	0.0011920739	-2.4499945D + 01
HATFLDH	4	13	P0	154	0.0012400781	-2.4500312D + 01
HATFLDH	4	13	P1	147	0.0010920679	-2.4500020D + 01
HATFLDH	4	13	P1	66	0.0018361151	-2.4500000D + 01
HATFLDH	4	13	P2	69	0.0007360463	-2.4500009D + 01
HIMMELBI	100	12	P0	111	0.0825611576	-1.7355696D + 03
HIMMELBI	100	12	P0	154	0.2448152900	-1.7355696D + 03
HIMMELBI	100	12	P1	147	0.2448152900	-1.7355694D + 03
HIMMELBI	100	12	P1	66	0.0256455988	-1.7355696D + 03
HIMMELBI	100	12	P2	69	0.2726170421	-1.7355696D + 03
HIMMELP2	2	1	P0	111	0.0001888118	-6.2053936D + 01
HIMMELP2	2	1	P0	154	0.0001892118	-6.2053936D + 01
HIMMELP2	2	1	P1	147	0.0001924120	-6.2053936D + 01
HIMMELP2	2	1	P1	66	0.0001884118	-6.2053936D + 01
HIMMELP2	2	1	P2	69	0.0001828114	-6.2053936D + 01
HIMMELP3	2	2	P0	111	0.0000508032	-5.9013178D + 01
HIMMELP3	2	2	P0	154	0.0000504031	-5.9013178D + 01
HIMMELP3	2	2	P1	147	0.0000504031	-5.9013178D + 01
HIMMELP3	2	2	P1	66	0.0000504032	-5.9013178D + 01
HIMMELP3	2	2	P2	69	0.0000484030	-5.9013178D + 01
HIMMELP4	2	3	P0	111	0.0000524033	-5.9013178D + 01
HIMMELP4	2	3	P0	154	0.0000520033	-5.9013178D + 01
HIMMELP4	2	3	P1	147	0.0000520032	-5.9013178D + 01
HIMMELP4	2	3	P1	66	0.0000524032	-5.9013178D + 01
HIMMELP4	2	3	P2	69	0.0000496031	-5.9013178D + 01
HIMMELP5	2	3	P0	111	0.0003120190	-5.9013178D + 01
HIMMELP5	2	3	P0	154	0.0001456091	-5.9013178D + 01
HIMMELP5	2	3	P1	147	0.0001452090	-5.9013178D + 01
HIMMELP5	2	3	P1	66	0.0000808051	-5.9013178D + 01
HIMMELP5	2	3	P2	69	0.0001688106	-5.9013178D + 01
HIMMELP6	2	5	P0	111	0.0000904057	-5.9013178D + 01
HIMMELP6	2	5	P0	154	0.0001540096	-5.9013178D + 01
HIMMELP6	2	5	P1	147	0.0001544097	-5.9013178D + 01
HIMMELP6	2	5	P1	66	0.0000848053	-5.9013178D + 01
HIMMELP6	2	5	P2	69	0.0001752110	-5.9013178D + 01
HS101	7	6	P0	111	0.1349284202	1.8097438D + 03
HS101	7	6	P0	154	0.8590536714	1.8106470D + 03
HS101	7	6	P1	147	0.0680882484	1.8094979D + 03
HS101	7	6	P1	66	0.1285679638	1.8094979D + 03
HS101	7	6	P2	69	0.0819571391	1.8095915D + 03
HS102	7	6	P0	111	0.0942178816	9.1187693D + 02
HS102	7	6	P0	154	2.0953316689	9.1316606D + 02
HS102	7	6	P1	147	0.0480389930	9.1184330D + 02
HS102	7	6	P1	66	0.0968620479	9.1184330D + 02
HS102	7	6	P2	69	0.0610598437	9.1184328D + 02

(Continua na página seguinte.)

Problema	n	m	pen	comb	CPU	f^*
HS103	7	6	P0	111	0.0731925741	5.4365642D + 02
HS103	7	6	P0	154	4.2162704468	1.6082409D + 02
HS103	7	6	P1	147	0.0418786146	5.4362669D + 02
HS103	7	6	P1	66	0.1356083751	5.4362954D + 02
HS103	7	6	P2	69	0.0804610327	5.4363086D + 02
HS104	8	6	P0	111	0.0080000004	4.2867533D + 00
HS104	8	6	P0	154	0.8800550103	1.1081368D + 01
HS104	8	6	P1	147	0.0120009780	4.2012753D + 00
HS104	8	6	P1	66	0.0050763171	3.9511577D + 00
HS104	8	6	P2	69	0.0033682107	3.9511006D + 00
HS105	8	1	P0	111	0.0362942703	1.0446117D + 03
HS105	8	1	P0	154	0.0347101726	1.0446117D + 03
HS105	8	1	P1	147	0.0367542952	1.0446117D + 03
HS105	8	1	P1	66	0.0354062133	1.0446117D + 03
HS105	8	1	P2	69	0.0355062112	1.0446117D + 03
HS106	8	6	P0	111	0.2025726438	7.0491782D + 03
HS106	8	6	P0	154	0.9200572968	2.1000000D + 03
HS106	8	6	P1	147	0.3559022546	1.3534262D + 04
HS106	8	6	P1	66	0.1734108776	1.3556128D + 04
HS106	8	6	P2	69	0.1740508229	1.4799484D + 04
HS116	13	15	P0	111	1.7881109715	9.7537815D + 01
HS116	13	15	P0	154	2.5961630344	5.2330879D + 01
HS116	13	15	P1	147	2.7441711426	1.2169980D + 02
HS116	13	15	P1	66	2.7881741524	9.5085637D + 01
HS116	13	15	P2	69	3.7202329636	1.9037277D + 02
HS117	15	5	P0	111	0.0112847053	3.2348693D + 01
HS117	15	5	P0	154	0.0326060355	3.2348676D + 01
HS117	15	5	P1	147	0.0344141573	3.2348685D + 01
HS117	15	5	P1	66	0.0365742892	3.2348679D + 01
HS117	15	5	P2	69	0.0307099149	3.2348678D + 01
HS118	15	29	P0	111	0.0096046003	6.6482049D + 02
HS118	15	29	P0	154	0.0281177592	6.6482045D + 02
HS118	15	29	P1	147	0.0048123016	6.6482054D + 02
HS118	15	29	P1	66	0.0519552417	6.6482045D + 02
HS118	15	29	P2	69	0.0034042129	6.6482044D + 02
HS13	2	1	P0	111	0.0011320700	9.2184778D - 01
HS13	2	1	P0	154	0.0007160450	9.4476856D - 01
HS13	2	1	P1	147	0.0008840560	9.3412371D - 01
HS13	2	1	P1	66	0.0008840552	9.2531913D - 01
HS13	2	1	P2	69	0.0011040689	9.3507894D - 01
HS16	2	2	P0	111	0.0011880740	2.5000000D - 01
HS16	2	2	P0	154	0.0004440270	2.5000000D - 01
HS16	2	2	P1	147	0.0004440280	2.5000000D - 01
HS16	2	2	P1	66	0.0002732171	2.5000000D - 01
HS16	2	2	P2	69	0.0004560280	2.5000000D - 01
HS17	2	2	P0	111	0.0027361710	1.0000994D + 00
HS17	2	2	P0	154	0.0005880361	1.0000003D + 00
HS17	2	2	P1	147	0.0006920443	1.0000008D + 00
HS17	2	2	P1	66	0.0015200954	1.0001996D + 00
HS17	2	2	P2	69	0.0004920306	9.9988223D - 01

(Continua na página seguinte.)

Problema	n	m	pen	comb	CPU	f^*
HS18	2	2	P0	111	0.0008360520	5.0000479D + 00
HS18	2	2	P0	154	0.0011880741	5.0000001D + 00
HS18	2	2	P1	147	0.0011840740	5.0000000D + 00
HS18	2	2	P1	66	0.0016361029	4.9999998D + 00
HS18	2	2	P2	69	0.0011440711	5.0000000D + 00
HS19	2	2	P0	111	0.0010840670	-6.9618260D + 03
HS19	2	2	P0	154	0.0006880430	-6.9618739D + 03
HS19	2	2	P1	147	0.0011680729	-6.9618432D + 03
HS19	2	2	P1	66	0.0010200641	-6.9618528D + 03
HS19	2	2	P2	69	0.0017441090	-6.9618432D + 03
HS21	2	1	P0	111	0.0000464029	-9.9960000D + 01
HS21	2	1	P0	154	0.0000464029	-9.9960000D + 01
HS21	2	1	P1	147	0.0000460029	-9.9960000D + 01
HS21	2	1	P1	66	0.0000472029	-9.9960000D + 01
HS21	2	1	P2	69	0.0000448028	-9.9960000D + 01
HS21MOD	7	1	P0	111	0.0000484031	-9.5960000D + 01
HS21MOD	7	1	P0	154	0.0000484030	-9.5960000D + 01
HS21MOD	7	1	P1	147	0.0000484030	-9.5960000D + 01
HS21MOD	7	1	P1	66	0.0000484030	-9.5960000D + 01
HS21MOD	7	1	P2	69	0.0000468030	-9.5960000D + 01
HS23	2	5	P0	111	0.0009720600	9.4721261D + 00
HS23	2	5	P0	154	0.0006400400	2.0000000D + 00
HS23	2	5	P1	147	0.0005440339	2.0000150D + 00
HS23	2	5	P1	66	0.0017641101	2.0000000D + 00
HS23	2	5	P2	69	0.0006520407	2.0000072D + 00
HS24	2	3	P0	111	0.0000508031	0.0000000D + 00
HS24	2	3	P0	154	0.0007040440	-9.9999880D - 01
HS24	2	3	P1	147	0.0007040440	-9.9999482D - 01
HS24	2	3	P1	66	0.1484892815	0.0000000D + 00
HS24	2	3	P2	69	0.0016641044	-1.0000005D + 00
HS30	3	1	P0	111	0.0004440270	1.0000338D + 00
HS30	3	1	P0	154	0.0000472030	1.0000000D + 00
HS30	3	1	P1	147	0.0000464029	1.0000000D + 00
HS30	3	1	P1	66	0.0000468029	1.0000000D + 00
HS30	3	1	P2	69	0.0000460028	1.0000000D + 00
HS31	3	1	P0	111	0.0005920370	6.0000000D + 00
HS31	3	1	P0	154	0.0008840550	6.0000000D + 00
HS31	3	1	P1	147	0.0004360270	6.0000000D + 00
HS31	3	1	P1	66	0.0005720359	6.0000000D + 00
HS31	3	1	P2	69	0.0003960249	5.9999936D + 00
HS33	3	2	P0	111	0.0007040440	-4.0000080D + 00
HS33	3	2	P0	154	0.0003120190	-3.9999995D + 00
HS33	3	2	P1	147	0.0002860179	-3.9999997D + 00
HS33	3	2	P1	66	0.0000000000	-6.0000000D + 00
HS33	3	2	P2	69	0.0002020127	-3.9999981D + 00
HS34	3	2	P0	111	0.0008680540	-8.3401800D - 01
HS34	3	2	P0	154	0.0010880680	-8.3396918D - 01
HS34	3	2	P1	147	0.0011480718	-8.3403158D - 01
HS34	3	2	P1	66	0.0006520409	-8.3403250D - 01
HS34	3	2	P2	69	0.0007760482	-8.3406725D - 01

(Continua na página seguinte.)

Problema	n	m	pen	comb	CPU	f^*
HS35	3	1	P0	111	0.0004880300	1.1112450D - 01
HS35	3	1	P0	154	0.0003120200	1.1117657D - 01
HS35	3	1	P1	147	0.0003520219	1.1111265D - 01
HS35	3	1	P1	66	0.0002808176	1.1111107D - 01
HS35	3	1	P2	69	0.0002372148	1.1110045D - 01
HS35I	3	1	P0	111	0.0004960300	1.1112450D - 01
HS35I	3	1	P0	154	0.0003120200	1.1117657D - 01
HS35I	3	1	P1	147	0.0003600230	1.1111265D - 01
HS35I	3	1	P1	66	0.0002852179	1.1111107D - 01
HS35I	3	1	P2	69	0.0002388149	1.1110045D - 01
HS35MOD	3	1	P0	111	0.0006600410	2.5005056D - 01
HS35MOD	3	1	P0	154	0.0003360210	2.5000026D - 01
HS35MOD	3	1	P1	147	0.0002300144	2.5001380D - 01
HS35MOD	3	1	P1	66	0.0000756047	2.5000119D - 01
HS35MOD	3	1	P2	69	0.0000740046	2.5000000D - 01
HS36	3	1	P0	111	0.0001472092	0.0000000D + 00
HS36	3	1	P0	154	0.0001752109	-3.3000000D + 03
HS36	3	1	P1	147	0.0001488093	-3.3000000D + 03
HS36	3	1	P1	66	0.0001476092	0.0000000D + 00
HS36	3	1	P2	69	0.0001472092	-3.3000000D + 03
HS37	3	2	P0	111	0.0029801859	-3.4560000D + 03
HS37	3	2	P0	154	0.0003640230	-3.4559998D + 03
HS37	3	2	P1	147	0.0004480279	-3.4560000D + 03
HS37	3	2	P1	66	0.0006040380	-3.4560000D + 03
HS37	3	2	P2	69	0.0005920367	-3.4559997D + 03
HS44	4	6	P0	111	0.0008840550	-1.5000002D + 01
HS44	4	6	P0	154	0.0006920429	-1.4999993D + 01
HS44	4	6	P1	147	0.0006960439	-1.4999905D + 01
HS44	4	6	P1	66	0.0000000000	-3.6046330D + 20
HS44	4	6	P2	69	0.0010520652	-1.5000000D + 01
CAMSHAPE	800	2404	P0	111	83.0491943359	-5.5017093D + 00
CAMSHAPE	800	2404	P0	154	324.7202758789	-5.7625825D + 00
CAMSHAPE	800	2404	P1	147	721.6051635742	-5.4216187D + 00
CAMSHAPE	800	2404	P1	66	491.9906005859	-6.1477874D + 00
CAMSHAPE	800	2404	P2	69	624.0631103516	-5.5208938D + 00
HS44NEW	4	6	P0	111	0.0007680480	-1.5000002D + 01
HS44NEW	4	6	P0	154	0.0003120191	-1.4999993D + 01
HS44NEW	4	6	P1	147	0.0003080189	-1.4999905D + 01
HS44NEW	4	6	P1	66	0.0040010214	-7.0843394D + 39
HS44NEW	4	6	P2	69	0.0003440210	-1.5000000D + 01
HS57	2	1	P0	111	0.0006120380	3.0647619D - 02
HS57	2	1	P0	154	0.0000944059	3.0646306D - 02
HS57	2	1	P1	147	0.0000948059	3.0646306D - 02
HS57	2	1	P1	66	0.0000948059	3.0646306D - 02
HS57	2	1	P2	69	0.0000936059	3.0646306D - 02
HS59	2	3	P0	111	0.0006280390	-6.7495047D + 00
HS59	2	3	P0	154	0.0013560849	-6.7495047D + 00
HS59	2	3	P1	147	0.0013560841	-6.7495047D + 00
HS59	2	3	P1	66	0.0010960682	-7.8027892D + 00
HS59	2	3	P2	69	0.0013440843	-6.7495047D + 00

(Continua na página seguinte.)

Problema	n	m	pen	comb	CPU	f^*
HS64	3	1	P0	111	0.0012360769	6.2996604D + 03
HS64	3	1	P0	154	0.0011280699	6.2997428D + 03
HS64	3	1	P1	147	0.0012640791	6.2998133D + 03
HS64	3	1	P1	66	0.0012520781	6.2998133D + 03
HS64	3	1	P2	69	0.0014400901	6.2998133D + 03
HS65	3	1	P0	111	0.0007880490	9.5354117D - 01
HS65	3	1	P0	154	0.0007000430	9.5353405D - 01
HS65	3	1	P1	147	0.0006960439	9.5353314D - 01
HS65	3	1	P1	66	0.0005360331	9.5352866D - 01
HS65	3	1	P2	69	0.0006720419	9.5352854D - 01
HS66	3	2	P0	111	0.0005680350	5.1818121D - 01
HS66	3	2	P0	154	0.0005160320	5.1817773D - 01
HS66	3	2	P1	147	0.0004840300	5.1826051D - 01
HS66	3	2	P1	66	0.0008720550	5.1816169D - 01
HS66	3	2	P2	69	0.0005400340	5.1816258D - 01
HS67	3	14	P0	111	0.0037122322	-1.1621187D + 03
HS67	3	14	P0	154	0.0025201577	-1.1621187D + 03
HS67	3	14	P1	147	0.0025361581	-1.1621187D + 03
HS67	3	14	P1	66	0.0013800869	-1.1621187D + 03
HS67	3	14	P2	69	0.0040962552	-1.1621187D + 03
HS70	4	1	P0	111	0.0124767795	7.4985251D - 03
HS70	4	1	P0	154	0.0124927815	7.4985251D - 03
HS70	4	1	P1	147	0.0125167845	7.4985251D - 03
HS70	4	1	P1	66	0.0124967806	7.4985251D - 03
HS70	4	1	P2	69	0.0124687804	7.4985251D - 03
HS72	4	2	P0	111	0.0021161321	7.2443016D + 02
HS72	4	2	P0	154	0.0015520970	7.2763500D + 02
HS72	4	2	P1	147	0.0015640982	7.2598442D + 02
HS72	4	2	P1	66	0.0017241082	7.2598442D + 02
HS72	4	2	P2	69	0.0018361154	7.2551442D + 02
HS76	4	3	P0	111	0.0009680600	-4.6818176D + 00
HS76	4	3	P0	154	0.0003520219	-4.6818165D + 00
HS76	4	3	P1	147	0.0003160200	-4.6818145D + 00
HS76	4	3	P1	66	0.0002796174	-4.6818498D + 00
HS76	4	3	P2	69	0.0002216139	-4.6818549D + 00
HS76I	4	3	P0	111	0.0009800610	-4.6818176D + 00
HS76I	4	3	P0	154	0.0003560221	-4.6818165D + 00
HS76I	4	3	P1	147	0.0003160199	-4.6818145D + 00
HS76I	4	3	P1	66	0.0002828176	-4.6818498D + 00
HS76I	4	3	P2	69	0.0002228140	-4.6818549D + 00
HS83	5	6	P0	111	0.0000000000	-3.1026428D + 04
HS83	5	6	P0	154	0.0203052685	-3.0665539D + 04
HS83	5	6	P1	147	0.0702963844	-3.0665547D + 04
HS83	5	6	P1	66	0.1919319928	-3.0665544D + 04
HS83	5	6	P2	69	0.0010080643	-3.0665540D + 04
HS84	5	6	P0	111	0.8334520459	-5.2803351D + 06
HS84	5	6	P0	154	0.1710906923	-5.2803351D + 06
HS84	5	6	P1	147	0.1698105633	-5.2803351D + 06
HS84	5	6	P1	66	0.0022921371	-5.2803351D + 06
HS84	5	6	P2	69	0.1575297564	-5.2803351D + 06

(Continua na página seguinte.)

Problema	n	m	pen	comb	CPU	f^*
HS85	5	38	P0	111	0.0790609419	-2.1886462D + 00
HS85	5	38	P0	154	0.7187650204	-2.2156031D + 00
HS85	5	38	P1	147	0.9643002152	-2.2156025D + 00
HS85	5	38	P1	66	0.0044282838	-2.2156070D + 00
HS85	5	38	P2	69	0.1673706025	-2.2156048D + 00
HS86	5	10	P0	111	0.0040000002	-3.2355132D + 01
HS86	5	10	P0	154	0.0068724304	-3.2348679D + 01
HS86	5	10	P1	147	0.0013400841	-3.2348680D + 01
HS86	5	10	P1	66	0.0019641218	-3.2348680D + 01
HS86	5	10	P2	69	0.0010760679	-3.2348912D + 01
HS93	6	2	P0	111	0.0000000000	0.0000000D + 00
HS93	6	2	P0	154	0.0000000000	0.0000000D + 00
HS93	6	2	P1	147	0.0000000000	0.0000000D + 00
HS93	6	2	P1	66	0.0000000000	0.0000000D + 00
HS93	6	2	P2	69	0.0040000002	0.0000000D + 00
HS95	6	4	P0	111	0.0010400651	1.5619525D - 02
HS95	6	4	P0	154	0.0009280579	1.5620287D - 02
HS95	6	4	P1	147	0.0010600659	1.5619525D - 02
HS95	6	4	P1	66	0.0003200200	1.5619525D - 02
HS95	6	4	P2	69	0.0007040439	1.5619525D - 02
HS96	6	4	P0	111	0.0010640661	1.5619525D - 02
HS96	6	4	P0	154	0.0009320589	1.5620287D - 02
HS96	6	4	P1	147	0.0010640662	1.5619525D - 02
HS96	6	4	P1	66	0.0003240199	1.5619525D - 02
HS96	6	4	P2	69	0.0007120442	1.5619525D - 02
HS97	6	4	P0	111	0.0033082070	4.0712458D + 00
HS97	6	4	P0	154	0.0084405271	3.1358091D + 00
HS97	6	4	P1	147	0.0084925322	3.1358091D + 00
HS97	6	4	P1	66	0.0089845601	3.1358091D + 00
HS97	6	4	P2	69	0.0471509472	4.0712464D + 00
HS98	6	4	P0	111	0.0032602029	4.0712458D + 00
HS98	6	4	P0	154	0.0084365271	3.1358091D + 00
HS98	6	4	P1	147	0.0084325280	3.1358091D + 00
HS98	6	4	P1	66	0.0020601291	3.1358091D + 00
HS98	6	4	P2	69	0.0459388718	4.0712464D + 00
HYDROELM	505	1008	P0	111	286.3218994141	-3.5823182D + 06
HYDROELM	505	1008	P0	154	68.4642944336	-3.5690147D + 06
HYDROELM	505	1008	P1	147	288.6580200195	-3.5811087D + 06
HYDROELM	505	1008	P1	66	209.2091064453	-3.5841366D + 06
HYDROELM	505	1008	P2	69	8.2213191986	-3.5820154D + 06
HYDROELS	169	336	P0	111	45.5788459778	-3.5837945D + 06
HYDROELS	169	336	P0	154	21.2453269958	-3.7560087D + 06
HYDROELS	169	336	P1	147	33.7501068115	-3.5822681D + 06
HYDROELS	169	336	P1	66	32.1620178223	-3.5323993D + 06
HYDROELS	169	336	P2	69	0.3335008323	-3.5822683D + 06
LOOTSMA	3	2	P0	111	0.0000000000	0.0000000D + 00
LOOTSMA	3	2	P0	154	0.1440089941	0.0000000D + 00
LOOTSMA	3	2	P1	147	0.7200450301	0.0000000D + 00
LOOTSMA	3	2	P1	66	0.7080439925	0.0000000D + 00
LOOTSMA	3	2	P2	69	0.2720170021	0.0000000D + 00

(Continua na página seguinte.)

Problema	n	m	pen	comb	CPU	f^*
MATRIX2	6	2	P_0	111	0.0005320330	1.2555170D - 05
MATRIX2	6	2	P_0	154	0.0000548035	2.8100989D - 16
MATRIX2	6	2	P_1	147	0.0000552035	2.8100989D - 16
MATRIX2	6	2	P_1	66	0.0004600281	1.1758043D - 28
MATRIX2	6	2	P_2	69	0.0000532033	2.0109967D - 12
OPTPRLOC	30	30	P_0	111	0.0996862352	-1.6413919D + 01
OPTPRLOC	30	30	P_0	154	3.9802484512	-1.6497020D + 01
OPTPRLOC	30	30	P_1	147	1.2324769497	-1.6419743D + 01
OPTPRLOC	30	30	P_1	66	0.0947619230	-1.6419771D + 01
OPTPRLOC	30	30	P_2	69	0.1182072461	-1.6419785D + 01
QC	9	4	P_0	111	0.0005960370	-9.5653773D + 02
QC	9	4	P_0	154	0.0003240200	-9.5653773D + 02
QC	9	4	P_1	147	0.0003520220	-9.5653773D + 02
QC	9	4	P_1	66	0.0002972186	-9.5653773D + 02
QC	9	4	P_2	69	0.0002852178	-9.5653773D + 02
QCNEW	9	3	P_0	111	0.0003440210	-8.0652185D + 02
QCNEW	9	3	P_0	154	0.0000568036	-8.0652185D + 02
QCNEW	9	3	P_1	147	0.0000564036	-8.0652185D + 02
QCNEW	9	3	P_1	66	0.0000564035	-8.0652185D + 02
QCNEW	9	3	P_2	69	0.0000548034	-8.0652185D + 02
S277-280	4	4	P_0	111	0.0003560220	5.0761908D + 00
S277-280	4	4	P_0	154	0.0003200200	5.0761912D + 00
S277-280	4	4	P_1	147	0.0002868179	5.0761911D + 00
S277-280	4	4	P_1	66	0.0013080824	5.0761898D + 00
S277-280	4	4	P_2	69	0.0014320894	5.0761497D + 00
SIMPLLLPA	2	2	P_0	111	0.0005960370	9.9999952D - 01
SIMPLLLPA	2	2	P_0	154	0.0001028064	1.0000000D + 00
SIMPLLLPA	2	2	P_1	147	0.0001032064	1.0000000D + 00
SIMPLLLPA	2	2	P_1	66	0.0002712170	1.0000000D + 00
SIMPLLLPA	2	2	P_2	69	0.0001460091	1.0000000D + 00
SIMPLLLPB	2	3	P_0	111	0.0005760360	1.1000777D + 00
SIMPLLLPB	2	3	P_0	154	0.0003320200	1.1000000D + 00
SIMPLLLPB	2	3	P_1	147	0.0003400211	1.1000000D + 00
SIMPLLLPB	2	3	P_1	66	0.0007520471	1.1000000D + 00
SIMPLLLPB	2	3	P_2	69	0.0002644166	1.1000000D + 00
STANCMIN	3	2	P_0	111	0.0007880490	4.2499817D + 00
STANCMIN	3	2	P_0	154	0.0012760801	4.2499941D + 00
STANCMIN	3	2	P_1	147	0.0010080629	4.2499961D + 00
STANCMIN	3	2	P_1	66	0.0014880929	4.2499989D + 00
STANCMIN	3	2	P_2	69	0.0010280643	4.2499016D + 00
SYNTHESES1	6	6	P_0	111	0.0013360830	7.5929890D - 01
SYNTHESES1	6	6	P_0	154	0.0007520469	7.5928268D - 01
SYNTHESES1	6	6	P_1	147	0.0007640481	7.5925000D - 01
SYNTHESES1	6	6	P_1	66	0.0013440840	7.5922168D - 01
SYNTHESES1	6	6	P_2	69	0.0007840490	7.5928216D - 01
TWOBARS	2	2	P_0	111	0.0010560660	1.5086523D + 00
TWOBARS	2	2	P_0	154	0.0003960240	1.5085653D + 00
TWOBARS	2	2	P_1	147	0.0004440280	1.5086500D + 00
TWOBARS	2	2	P_1	66	0.0008560540	1.5086510D + 00
TWOBARS	2	2	P_2	69	0.0004360270	1.5085224D + 00

(Continua na página seguinte.)

Problema	n	m	pen	comb	CPU	f^*
ZECEVIC2	2	2	$P0$	111	0.0007360460	$-4.1250002D + 00$
ZECEVIC2	2	2	$P0$	154	0.0002120132	$-4.1250000D + 00$
ZECEVIC2	2	2	$P1$	147	0.0002096131	$-4.1250000D + 00$
ZECEVIC2	2	2	$P1$	66	0.0004040251	$-4.1250000D + 00$
ZECEVIC2	2	2	$P2$	69	0.0002024127	$-4.1250000D + 00$
ZECEVIC3	2	2	$P0$	111	0.0009360580	$9.7309461D + 01$
ZECEVIC3	2	2	$P0$	154	0.0004080250	$9.7309380D + 01$
ZECEVIC3	2	2	$P1$	147	0.0004600290	$9.7309445D + 01$
ZECEVIC3	2	2	$P1$	66	0.0006480410	$9.7309449D + 01$
ZECEVIC3	2	2	$P2$	69	0.0004440270	$9.7309098D + 01$
ZECEVIC4	2	2	$P0$	111	0.0007320450	$7.5575114D + 00$
ZECEVIC4	2	2	$P0$	154	0.0003480220	$7.5575045D + 00$
ZECEVIC4	2	2	$P1$	147	0.0003480220	$7.5574627D + 00$
ZECEVIC4	2	2	$P1$	66	0.0004600290	$7.5574929D + 00$
ZECEVIC4	2	2	$P2$	69	0.0003920240	$7.5575075D + 00$
ZY2	3	2	$P0$	111	0.0008200510	$2.0000142D + 00$
ZY2	3	2	$P0$	154	0.0003000190	$2.0000005D + 00$
ZY2	3	2	$P1$	147	0.0002684168	$2.0000003D + 00$
ZY2	3	2	$P1$	66	0.0000000000	$0.0000000D + 00$
ZY2	3	2	$P2$	69	0.0002704169	$2.0000020D + 00$

(Fim da Tabela B.1.)

Referências Bibliográficas

- [1] R. Andreani, E. G. Birgin, J. M. Martínez, e M. Schuverdt. Augmented Lagrangian methods under the constant positive linear dependence constraint qualification. *Mathematical Programming*, 111(1-2):5–32, 2008.
- [2] R. Andreani, E. G. Birgin, J. M. Martínez, e M. Schuverdt. On Augmented Lagrangian methods with general lower-level constraints. *SIAM Journal on Optimization*, 18(4):1286–1309, 2008.
- [3] R. Andreani, J. M. Martínez, e M. L. Schuverdt. On the relation between constant positive linear dependence condition and quasinormality constraint qualification. *Journal of Optimization Theory and Applications*, 125(2):473–485, 2005.
- [4] A. Auslender, M. Teboulle, e S. Ben-Tiba. Interior proximal and multiplier methods based on second order homogeneous kernels. *Mathematics of Operations Research*, 24:645–668, 1999.
- [5] A. Ben-Tal e M. Zibulevsky. Penalty/barrier multiplier methods for convex programming problems. *SIAM Journal on Optimization*, 7:347–366, 1997.
- [6] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, New York, 1982.
- [7] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 1995.
- [8] E. G. Birgin, R. A. Castillo, e J. M. Martínez. Numerical comparison of augmented Lagrangian algorithms for nonconvex problems. *Computational Optimization and Applications*, 31:31–55, 2005.
- [9] E. G. Birgin, D. Fernandez, e J. M. Martínez. On the boundedness of penalty parameters in an augmented Lagrangian method with lower level constraints. *Optimization Methods and Software*, 27:1001–1024, 2012.
- [10] E. G. Birgin e J. M. Martínez. Large-scale active-set box-constrained optimization method with spectral projected gradients. *Computational Optimization and Applications*, 23(1):101–125, 2002.
- [11] E. G. Birgin, J. M. Martínez, e M. Raydan. Nonmonotone spectral projected gradient methods on convex sets. *SIAM Journal on Optimization*, pages 1196–1211, 2000.
- [12] E. G. Birgin, J. M. Martínez, e M. Raydan. SPG: Software for convex-constrained optimization, 2001.
- [13] E. G. Birgin e J. M. Martínez. Augmented Lagrangian method with nonmonotone penalty parameters for constrained optimization. *Computational Optimization and Applications*, 51:941–965, 2012.

- [14] E.G. Birgin, J.M. Martínez, e L.F. Prudente. Augmented Lagrangians with possible infeasibility and finite termination for global nonlinear programming. *Journal of Global Optimization*, 58(2):207–242, 2014.
- [15] R. A. Castillo. *Métodos de Langriano Aumentado usando Penalidades Generalizadas para Programação não linear*. Tese de doutorado, COPPE - Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil, 1998.
- [16] R. A. Castillo, C. C. Gonzaga, E. W. Karas, e L. C. Matioli. A unified approach to multiplier and proximal methods. Technical report, Universidade Federal do Paraná, 2013.
- [17] F. H. Clarke. *Optimization and nonsmooth analysis*. SIAM, Philadelphia, 1990.
- [18] E. D. Dolan e J. J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91(2):201–213, 2002.
- [19] J. Eckstein e P. J. S. Silva. A practical relative error criterion for augmented Lagrangians. *Mathematical Programming*, 141(1-2):319–348, 2013.
- [20] R. G. Eustáquio. Condições de otimalidade e de qualificação para problemas de programação não linear. Dissertação de mestrado em Métodos Numéricos em Engenharia, UFPR, 2007.
- [21] D. Fernández e M. Solodov. Local convergence of exact and inexact augmented Lagrangian methods under the second-order sufficient optimality condition. *SIAM Journal on Optimization*, 22(2):384–407, 2012.
- [22] M. R. Hestenes. Multiplier and gradient methods. *Journal of Optimization Theory and applications*, 4:303–320, 1969.
- [23] W. Hock e K. Schittkowski. Test examples for nonlinear programming codes. *Journal of Optimization Theory and Applications*, 30(1):127–129, 1980.
- [24] C. Humes e P. S. Silva. Strict convex regularization, proximal points and augmented Lagrangians. *RAIRO Operations Research*, 34:283–303, 2000.
- [25] A. Izmailov e M. Solodov. *Otimização: Métodos computacionais*, volume 2. IMPA, Rio de Janeiro, 2012.
- [26] A. Izmailov, M. Solodov, e E. Uskov. Global convergence of augmented Lagrangian methods applied to optimization problems with degenerate constraints, including problems with complementarity constraints. *SIAM Journal on Optimization*, 22(4):1579–1606, 2012.
- [27] K. C. Kiwiel. Proximal minimization methods with generalized Bregman function. *SIAM Journal on Control and Optimization*, 35:1142–1168, 1997.
- [28] B. Kort e D. Bertsekas. Combined primal–dual and penalty methods for convex programming. *SIAM Journal on Control and Optimization*, 14(2):268–294, 1976.
- [29] B. W. Kort e D. P. Bertsekas. Multiplier methods for convex programming. In *Decision and Control including the 12th Symposium on Adaptive Processes, 1973 IEEE Conference on*, volume 12, págs 428–432, 1973.

- [30] D. Luenberger e Y. Ye. *Linear and Nonlinear Programming*. Springer, 3^a edição, 2008.
- [31] J. M. Martínez. Otimização prática usando o Lagrangiano aumentado. IMECC – UNICAMP, 2009.
- [32] L. C. Matioli. *Uma nova metodologia para construção de funções de penalização para algoritmos de Lagrangiano aumentado*. PhD thesis, Universidade Federal de Santa Catarina, Florianópolis, Brasil, 2001.
- [33] H. Nakayama, H. Sayama, e Y. Sawaragi. A generalized Lagrangian function and multiplier method. *Journal of Optimization Theory and Applications*, 17(3-4):211–227, 1975.
- [34] J. Nocedal e S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer-Verlag, 1999.
- [35] R. Polyak e M. Teboulle. Nonlinear rescaling and proximal-like methods in convex optimization. *Mathematical Programming*, 76(2):265–284, 1997.
- [36] M. J. D. Powell. A method for nonlinear constraints in minimization problems. In R. Fletcher, editor, *Optimization*, págs. 283–298. Academic Press, New York, 1969.
- [37] L. Qi e Z. Wei. On the constant positive linear dependence condition and its application to SQP methods. *SIAM Journal on Optimization*, 35(2):183–238, 1993.
- [38] A. A. Ribeiro e E. W. Karas. *Otimização contínua: aspectos teóricos e computacionais*. Cengage Learning, 2013.
- [39] R. T. Rockafellar. The multiplier method of Hestenes and Powell applied to convex programming. *Journal of Optimization Theory and applications*, 12:555–562, 1973.
- [40] K. Schittkowski. An updated set of 306 test problems for nonlinear programming with validated optimal solutions - user's guide. Technical report, Department of Computer Science, University of Bayreuth, Alemanha, 2008.
- [41] P. Tseng e D. Bertsekas. On the convergence of the exponential multiplier method for convex programming. *Mathematical Programming*, 60(1-3):1–19, 1993.
- [42] A. E. Xavier. *Penalização hiperbólica*. Tese de doutorado, COPPE - Universidade Federal de Rio de Janeiro, Rio de Janeiro, Brasil, 1992.