

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação
Departamento de Engenharia de Sistemas

**Estudo e Implementação dos Métodos da Lagrangeana Aumentada e da
Barreira Modificada**

Autor: Iara da Cunha Ribeiro da Silva
Orientador: Prof. Dr. Anésio dos Santos Júnior

Dissertação de Mestrado apresentada à Faculdade
de Engenharia Elétrica e de Computação como
parte dos requisitos para obtenção do título de
Mestre em Engenharia Elétrica.

Aprovação em 01/07/2008

Banca Examinadora:

Prof. Dr. Anésio dos Santos Júnior (Orientador) FEEC/UNICAMP
Prof. Dr. Akebo Yamakami FEEC/UNICAMP
Prof. Dr. Esdras Penêdo de Carvalho DMA/UEM

Campinas, 2008

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

Si38e Silva, Iara da Cunha Ribeiro da
Estudo e implementação dos métodos da lagrangeana
aumentada e da barreira modificada / Iara da Cunha Ribeiro
da Silva. --Campinas, SP: [s.n.], 2008.

Orientador: Anésio dos Santos Júnior
Dissertação (Mestrado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

1. Programação não-linear. 2. Otimização matemática. 3.
Algoritmos. I. Santos Júnior, Anésio dos. II. Universidade
Estadual de Campinas. Faculdade de Engenharia Elétrica e
de Computação. III. Título.

Título em Inglês: Study and implementation of augmented lagrangian and modified
barrier methods.

Palavras-chave em Inglês: Nonlinear programming, Mathematical optimization,
Algorithms.

Área de concentração: Energia Elétrica.

Titulação: Mestre em Engenharia Elétrica

Banca examinadora: Akebo Yamakami, Esdras Penêdo de Carvalho.

Data da defesa: 1/7/2008.

Programa de Pós-Graduação: Engenharia Elétrica.

COMISSÃO JULGADORA - TESE DE MESTRADO

Candidata: Iara da Cunha Ribeiro da Silva

Data da Defesa: 1 de julho de 2008

Título da Tese: "Estudo e Implementação dos Métodos da Lagrangeana Aumentada e da Barreira Modificada"

Prof. Dr. Anésio dos Santos Júnior (Presidente):

Prof. Dr. Esdras Penedo de Carvalho:

Prof. Dr. Akebo Yamakami:

Resumo

Neste trabalho analisamos e comparamos extensões dos métodos clássicos de penalidades: métodos da Lagrangeana aumentada e da barreira logarítmica modificada. As penalidades podem ser classificadas como externa e interna ou também por penalidade e barreira. Os métodos de penalidade externa geram seqüências de soluções infactíveis e de penalidade interna seqüências de soluções factíveis. O método da Lagrangeana aumentada é uma combinação dos métodos de penalidade quadrática e dual Lagrange. Já o método da barreira modificada combina o método de barreira logarítmica com o método dual Lagrange. A estrutura desses métodos é bastante similar, ambos geram pontos factíveis e infactíveis. Esses métodos foram aplicados a problemas não-lineares com restrições de desigualdade e o desempenho dos algoritmos implementados é discutido neste trabalho.

Palavras-chave: método de penalidade, método de barreira, método da Lagrangeana aumentada, método da barreira logarítmica modificada.

Abstract

In this work we analyze and compare extensions of traditional penalty methods: augmented Lagrangian and modified logarithmic barrier methods. The penalties may be classified as external and internal or penalty and barrier. The external penalty method generates a sequence of unfeasible solutions and the internal penalty method produces a sequence of feasible solutions. The augmented Lagrangian method is a combination of quadratic penalty and Lagrange dual methods. Already the modified barrier method combines the logarithmic penalty and Lagrange dual methods. The structure of these methods is very similar, both generate feasible and unfeasible points. These methods have been applied to nonlinear problems with inequality restrictions and the performance of algorithms implemented is discussed in this work.

Keywords: penalty method, barrier method, augmented Lagrangian method, modified logarithmic barrier method.

Sumário

1	Introdução	1
2	Método de Penalidade e Método da Barreira	4
2.1	Método de Penalidade	5
2.2	Método de Barreira	11
3	Lagrangeana Aumentada	18
3.1	A Função Lagrangeana Aumentada - Restrições de igualdade	19
3.2	A Função Lagrangeana Aumentada - Restrições de desigualdade	23
3.3	Método da Lagrangeana Aumentada	25
4	Barreira Modificada	34
4.1	A Função Barreira Modificada	35
4.2	Método da Barreira Modificada	40
5	Resultados	44
5.1	Problemas Testes	45
5.2	Implementação	49
5.3	Resultados dos testes	51
6	Conclusão	56
	Referências bibliográficas	58
A	Definições	62
B	Método de Newton	65

Lista de Figuras

2.1	Comportamento dos pontos gerados pelos métodos penalidade e barreira. .	5
2.2	Representação gráfica das funções penalidade com $q = 1$ e $q = 2$ para a restrição $g(x) = -x + 3 \geq 0$	7
2.3	Em (a) temos a derivada de primeira ordem da função penalidade $\mathcal{P}_1$ e em (b) temos a derivada de $\mathcal{P}_2$	8
2.4	Representação gráfica da função penalidade quadrática com diferentes valores para o parâmetro c	9
2.5	Função auxiliar $\mathcal{F}_{\mathcal{P}}(x)$ para o problema (2.3) com diferentes parâmetros de penalidade.	10
2.6	Exemplos de conjuntos robustos e não robustos.	12
2.7	Representação gráfica da função barreira com diferentes parâmetros c	14
2.8	Função auxiliar $\mathcal{F}_{\mathcal{B}}(x)$ para o problema (2.8) com diferentes parâmetros de penalidade.	16
3.1	Gráficos da função Lagrangeana aumentada (3.9) com diferentes valores de parâmetros de penalidade e multiplicadores de Lagrange.	22
3.2	Passos principais do método da Lagrangeana aumentada.	26
3.3	Comportamento da função (3.26) quando $\lambda \rightarrow \lambda^*$	31
3.4	Comportamento da função (3.26) quando $c \rightarrow \infty$	32
4.1	Representação gráfica dos conjuntos Ω e Ω_c do problema 4.4.	36
4.2	Comportamento da função (4.6).	40
5.1	Representação Gráfica do Problema 1.	46
5.2	Representação Gráfica do Problema 2.	47
5.3	Representação Gráfica do Problema 3.	47

5.4	Representação Gráfica do Problema 4.	48
5.5	Representação Gráfica do Problema 5.	48
5.6	Fluxograma do método da Lagrangeana aumentada.	49
5.7	Fluxograma do método da barreira modificada.	50
A.1	Ilustração de um conjunto convexo e não-convexo.	62
A.2	Exemplos de funções convexas e não-convexa.	63
C.1	Seqüências de soluções primais geradas pelos métodos para o problema 1. .	68
C.2	Seqüências de soluções primais geradas pelos métodos para o problema 2. .	69
C.3	Seqüências de soluções primais geradas pelos métodos para o problema 3. .	70
C.4	Seqüências de soluções primais geradas pelos métodos para o problema 4. .	71
C.5	Função auxiliar $\mathcal{F}_{\mathcal{P}}(x)$ para o problema (2.3) com diferentes parâmetros de penalidade.	72

Lista de Tabelas

5.1	Soluções obtidas pelo método Lagrangeana aumentada.	52
5.2	Soluções obtidas pelo método barreira modificada.	53
5.3	Número de iterações interna e externa.	54
5.4	Erros absolutos das soluções primais e duais.	55

Notações

$\mathbb{R}^n$ – conjunto dos vetores $x = (x_1, x_2, \dots, x_n)^T$

$x^T y$ – produto escalar entre x e $y \in \mathbb{R}^n$, $x^T y = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$

$\|x\| = \sqrt{x^T x}$ (norma euclideana)

$\mathbb{Z}_+^*$ – conjunto dos números inteiros positivos

$f(x)$ – função objetivo

$h_j(x)$ – j -ésima restrição de igualdade

$g_i(x)$ – i -ésima restrição de desigualdade

$h(x)$ – vetor contendo as restrições de igualdade $h(x) = (h_1(x), h_2(x), \dots, h_p(x))^T$

$g(x)$ – vetor contendo as restrições de desigualdade $g(x) = (g_1(x), g_2(x), \dots, g_m(x))^T$

Ω – conjunto da região factível

Ω_c – conjunto da região factível relaxada

λ_j – multiplicador de Lagrange ou variável dual relacionado à j -ésima restrição de igualdade

μ_i – multiplicador de Lagrange ou variável dual relacionado à i -ésima restrição de desigualdade

λ – vetor contendo os multiplicadores de Lagrange relacionado às restrições de igualdade
 $\lambda = (\lambda_1, \dots, \lambda_p)^T$

μ – vetor contendo os multiplicadores de Lagrange relacionado às restrições de desigualdade $\mu = (\mu_1, \dots, \mu_m)^T$

C^k – denota o conjunto de funções $f : \mathbb{R}^n \rightarrow \mathbb{R}$ tais que todas as derivadas de ordem menor ou igual a k são contínuas

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix} - \text{gradiente de } f$$

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \dots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix} - \text{matriz Hessiana de } f$$

$\mathcal{L}(x, \lambda)$ – função Lagrangeana relacionado à problemas de otimização com restrições de igualdade

$\mathcal{L}(x, \mu)$ – função Lagrangeana relacionado à problemas de otimização com restrições de desigualdade

$\mathcal{P}(x)$ – função penalidade

$\mathcal{B}_{\mathcal{L}}(x)$ – função barreira logarítmica

$\mathcal{B}_{\mathcal{I}}(x)$ – função barreira inversa

$\mathcal{F}_{\mathcal{P}}(x)$ – função auxiliar de penalidade

$\mathcal{F}_{\mathcal{B}}(x)$ – função auxiliar de barreira

$\mathcal{L}_{\mathcal{A}}(x, \lambda)$ – função Lagrangeana aumentada relacionado à problemas de otimização com restrições de igualdade

$\mathcal{L}_{\mathcal{A}}(x, \mu)$ – função Lagrangeana aumentada relacionado à problemas de otimização com restrições de desigualdade

$\mathcal{M}_{\mathcal{L}}(x, \mu)$ – função barreira logarítmica modificada de Polyak

$\mathcal{M}_{\mathcal{I}}(x, \mu)$ – função barreira inversa modificada de Polyak

s_i – variável de excesso

c – parâmetro de penalidade ou relaxação no caso da função barreira modificada

β – constante que incrementa ou decrementa o parâmetro de penalidade ou relaxação

Capítulo 1

Introdução

Desde o surgimento de um dos métodos mais conhecidos da otimização, método do gradiente ou método da máxima descida, desenvolvido pelo matemático francês Augustin-Louis Cauchy (1789-1857), são mais de um século e meio de estudos e aprimoramentos de técnicas para solucionar problemas da programação não-linear. O desenvolvimento de métodos para otimização foi mais significativo na década de quarenta após o surgimento dos computadores. O primeiro método a destacar-se foi o método Simplex, criado por Dantzig, para otimização linear. Aproximadamente na década de cinquenta originaram-se os métodos para otimização não-linear restrita que visam solucionar esse problema resolvendo uma seqüência de problemas aproximados. Esses métodos são conhecidos como métodos de penalidades.

Os métodos de penalidades dividem-se em penalidade externa e interna. Esses dois métodos podem ser chamados, respectivamente, por penalidade e barreira clássicas. A estrutura algorítmica desses métodos é análoga, o que difere é a seqüência de problemas irrestritos que cada método gera.

O método de penalidade externa penaliza os pontos inactivíveis tornando-os mais próximos da região factível. Já o método de penalidade interna evita que os pontos factíveis tornem-se inactivíveis, penalizando os pontos que se aproximam da fronteira da factibilidade. Os métodos de barreira foram introduzidos por Frisch [Fri55] e Carrol [CF61]. Tanto a teoria quanto a implementação do método de barreira foram extensivamente estudadas na década de sessenta por Fiacco e McCormick [FM68]. Mas devido ao problema de mal condicionamento da Hessiana da função barreira, causado quando seu parâmetro de penalidade c torna-se muito grande, esses métodos deixaram de ser investigados durante

algum tempo. Esse mal condicionamento está relacionado com o número de condição da matriz Hessiana que se aproxima do infinito quando c cresce muito [BS93]. Os métodos de barreiras são os mais antigos métodos de pontos interiores e voltaram a ter destaque com os métodos de Karmarkar [Kar84] para programação linear.

O problema de mal condicionamento não aparece somente nos métodos de penalidade interna, mas também no método de penalidade externa [BS93, Lue86]. Diante disso, reformulações que melhorem o desempenho desses métodos foram propostas [Hes69, Pol92]. Esses métodos melhorados são o da Lagrangeana aumentada e o da barreira modificada.

O método da Lagrangeana aumentada foi desenvolvida por Hestenes [Hes69] e Powell [Pow69] e estudado por Rockafellar [Roc74] e Bertsekas [Ber76]. A Lagrangeana aumentada combina aspectos dos métodos de penalidade e dual Lagrange. Os dois métodos operados em conjunto eliminam algumas desvantagens existentes quando esses são utilizados individualmente. O método de barreira modificada introduzida por Polyak [Pol92] insere multiplicadores de Lagrange e uma relaxação nas restrições em relação ao método de barreira clássica. Essa relaxação permite ao método operar com pontos infactíveis. Tanto o método Lagrangeana aumentada quanto a barreira modificada são empregados em problemas de grande porte como, por exemplo, o problema do fluxo de potência ótimo [Car04, Sou06].

No trabalho de Nash, Polyak e Sofer [NPS94] foram comparados os métodos da barreira estabilizada com o método da barreira modificada. Neste artigo, o método da barreira modificada obteve melhor desempenho comparado com o método da barreira estabilizada. O método da barreira estabilizada opera somente com pontos factíveis e o método da barreira modificada opera tanto com pontos factíveis quanto infactíveis. Como este último método, o método da Lagrangeana aumentada também não faz distinção de gerar pontos factíveis ou infactíveis. Essa similaridade dos métodos da barreira modificada e da Lagrangeana aumentada nos motivou a analisar e comparar esses métodos.

Neste trabalho estudamos e implementamos os métodos primordiais da Lagrangeana aumentada e da barreira modificada, como mostrados nos trabalhos [Roc74] e [Pol92], respectivamente. Não usamos nenhuma técnica adicional que melhore o desempenho desses algoritmos justamente para efeito de comparação.

Esta dissertação foi estruturada da seguinte maneira:

No Capítulo 2 apresentaremos os métodos de penalidades clássicas: penalidade ex-

terna e interna. No Capítulo 3 veremos a extensão do método de penalidade externa, Lagrangeana aumentada. Mostraremos no Capítulo 4 a teoria do método da barreira modificada. Os testes e a maneira como foram implementados os métodos da Lagrangeana aumentada e da barreira modificada veremos no Capítulo 5. A conclusão deste trabalho apresentaremos no Capítulo 6.

Capítulo 2

Método de Penalidade e Método da Barreira

Neste capítulo apresentaremos os métodos primordiais utilizados na resolução de problemas de otimização restrita. Os métodos de penalidade e barreira, conhecidos respectivamente por método de penalidade externa e método de penalidade interna, nos ajudarão na compreensão dos métodos que veremos nos capítulos seguintes: método da Lagrangeana aumentada e método da barreira logarítmica modificada. Os métodos de penalidades (externa e interna) consistem em transformar um problema restrito em uma seqüência de problemas irrestritos parametrizados. Podemos resolver esses problemas irrestritos através de métodos como o de Newton, quase-Newton, gradiente conjugado entre outros. O método de penalidade externa, precursor do método dos multiplicadores ou método da Lagrangeana aumentada, visa atuar sobre as restrições não factíveis por meio de uma função penalidade, com o propósito de “pressioná-las” no sentido de sua factibilização. Já o método de penalidade interna evita que os pontos gerados “escapem” da região factível, utilizando algum tipo de função que penaliza os pontos que se aproximam da fronteira do conjunto. Na Figura 2.1 podemos ter uma noção de como esses métodos funcionam. Este capítulo foi elaborado tendo como referências [BS93], [Ber99], [Fle81], [Lue86], [MS95], [NW99] e [PSU93] com a finalidade de ilustrar o funcionamento desses métodos.

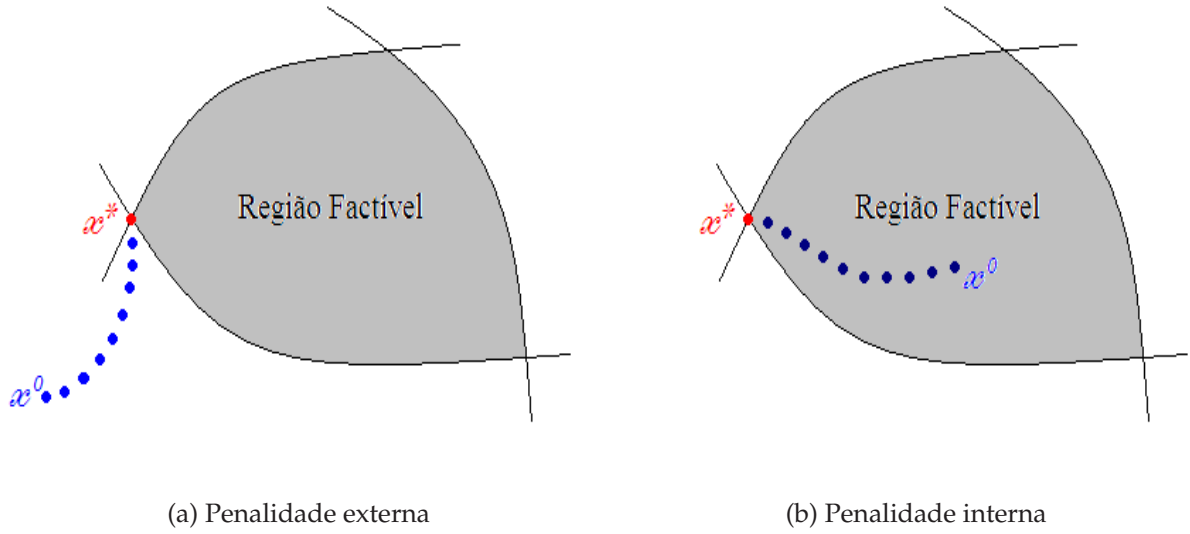


Figura 2.1: Comportamento dos pontos gerados pelos métodos penalidade e barreira.

2.1 Método de Penalidade

Consideremos o seguinte problema

$$\begin{aligned} &\text{Minimizar } f(x) \\ &\text{sujeito a } x \in \Omega, \end{aligned} \tag{2.1}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e Ω é o conjunto formado pelas restrições de igualdade e/ou desigualdade do problema (2.1). O método de penalidade visa resolver uma sequência de problemas irrestritos parametrizados, que chamaremos de subproblemas. Acrescentamos à função objetivo do problema (2.1) uma penalização obtendo o seguinte subproblema,

$$\text{Minimizar } \mathcal{F}_{\mathcal{P}}(x) = f(x) + c\mathcal{P}(x), \tag{2.2}$$

onde $\mathcal{F}_{\mathcal{P}}(x)$ é a função auxiliar de penalidade, c é uma constante positiva, denominada de parâmetro de penalidade, e $\mathcal{P}(x)$ é a função penalidade definida no $\mathbb{R}^n$. A função penalidade deve satisfazer as seguintes propriedades:

- $\mathcal{P}$ é contínua;

- $\mathcal{P}(x) \geq 0$ para todo $x \in \mathbb{R}^n$;
- $\mathcal{P}(x) = 0$ se e somente se $x \in \Omega$.

A seguir apresentamos alguns exemplos de funções de penalidades clássicas.

- Função penalidade para problemas com restrições de igualdade:

$$\mathcal{P}(x) = \sum_{j=1}^p |h_j(x)|^q, \quad q \in \mathbb{Z}_+^*.$$

- Função penalidade para problemas com restrições de desigualdade:

$$\mathcal{P}(x) = \sum_{i=1}^m (\max\{0, -g_i(x)\})^q, \quad q \in \mathbb{Z}_+^*.$$

- Função penalidade para problemas com restrições de desigualdade e igualdade:

$$\mathcal{P}(x) = \sum_{j=1}^p |h_j(x)|^q + \sum_{i=1}^m (\max\{0, -g_i(x)\})^q, \quad q \in \mathbb{Z}_+^*.$$

Na Figura 2.2 exemplificamos o comportamento das funções penalidades para restrições de desigualdade, ou seja, $\mathcal{P}(x) = \sum_{i=1}^m (\max\{0, -g_i(x)\})^q$ para $q = 1$ e $q = 2$. Para ilustrarmos utilizamos a seguinte restrição,

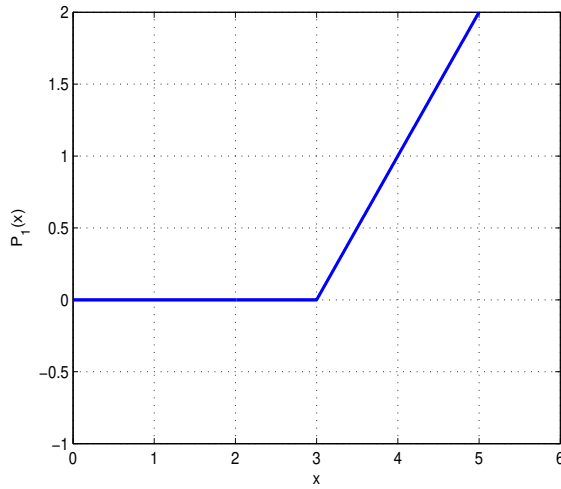
$$g(x) = -x + 3 \geq 0$$

A função penalidade para esta restrição é

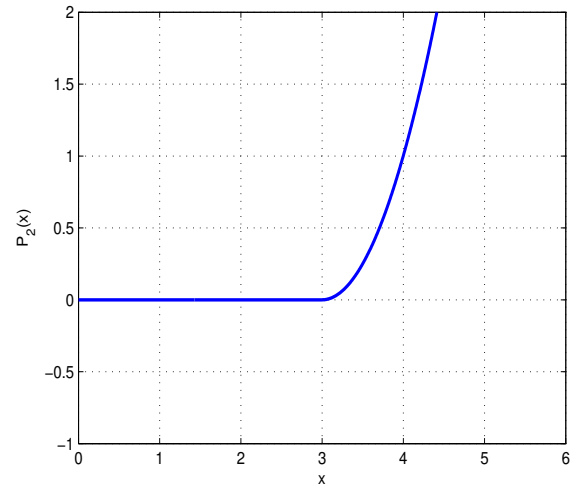
$$\mathcal{P}(x) = (\max\{0, x - 3\})^q.$$

Esta pode ser reescrita como,

$$\mathcal{P}(x) = \begin{cases} (x - 3)^q, & \text{se } x > 3, \\ 0, & \text{se } x \leq 3. \end{cases}$$



(a) $\mathcal{P}_1(x) = \max\{0, x - 3\}$



(b) $\mathcal{P}_2(x) = (\max\{0, x - 3\})^2$

Figura 2.2: Representação gráfica das funções penalidade com $q = 1$ e $q = 2$ para a restrição $g(x) = -x + 3 \geq 0$.

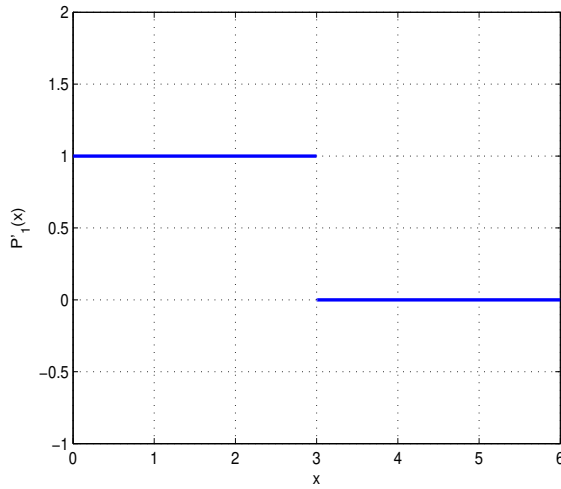
A função penalidade mais utilizada é a função penalidade quadrática¹, ou seja, $q=2$. É conveniente usarmos a função penalidade quadrática para assegurar a sua diferenciabilidade na fronteira da região factível. Vejamos o caso mencionado anteriormente para a restrição $g(x) = -x + 3 \geq 0$. Repare na figura 2.3 que a derivada da função penalidade com $q = 1$ não é contínua em $x = 3$. Já para $q = 2$ preserva-se a diferenciabilidade da função $\mathcal{P}_2$ e a continuidade de sua função derivada. O termo de penalidade quadrático é também conhecido por *função de penalidade Courant-Beltrami* [PSU93].

Agora discutiremos um outro exemplo utilizando as seguintes restrições:

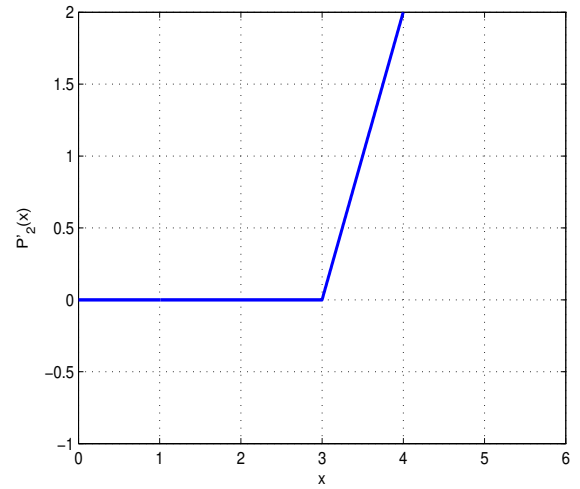
$$\begin{aligned} g_1(x) &= x - 2 \geq 0 \\ g_2(x) &= -x + 3 \geq 0. \end{aligned}$$

Observe na Figura 2.4 o comportamento da função penalidade quadrática para essas restrições, $\mathcal{P}(x) = (\max\{0, -x + 2\})^2 + (\max\{0, x - 3\})^2$, utilizando diferentes parâmetros de penalidades.

¹Neste trabalho seguiremos a descrição da função penalidade quadrática de algumas literaturas, [Ber99], [Lue86] e [NW99].



$$(a) P'_1(x) = \begin{cases} 1, & \text{se } x > 3 \\ 0, & \text{se } x \leq 3 \end{cases}$$



$$(b) P'_2(x) = \max\{0, 2(x - 3)\}$$

Figura 2.3: Em (a) temos a derivada de primeira ordem da função penalidade $\mathcal{P}_1$ e em (b) temos a derivada de $\mathcal{P}_2$.

Antes de apresentarmos o algoritmo do método da penalidade ilustraremos a função auxiliar $\mathcal{F}_{\mathcal{P}}(x, c)$ para o seguinte exemplo [BS93]:

$$\begin{aligned} &\text{Minimizar} && f(x) = x_1^2 + x_2^2 \\ &\text{sujeito a} && h(x) = x_1 + x_2 - 1 = 0 \end{aligned} \tag{2.3}$$

cuja solução é $x^* = \left(\frac{1}{2}, \frac{1}{2}\right)^T$.

A função auxiliar de penalidade para esse exemplo é

$$\mathcal{F}_{\mathcal{P}}(x) = x_1^2 + x_2^2 + \frac{c}{2}(x_1 + x_2 - 1)^2.$$

Na figura 2.5 temos as curvas de nível da função $\mathcal{F}_{\mathcal{P}}(x)$ para $c = 1$, $c = 10$, $c = 100$ e $c = 1000$. Note que para $c=100$ o ponto mínimo da função $\mathcal{F}_{\mathcal{P}}(x)$ está bem próxima da solução ótima do problema (2.3), $x^* = \left(\frac{1}{2}, \frac{1}{2}\right)^T$, e que para $c = 1000$ visivelmente não há nenhuma diferença entre o ponto mínimo da função $\mathcal{F}_{\mathcal{P}}(x)$ e a solução ótima x^* . Para alguns problemas, podemos encontrar a solução resolvendo um único subproblema (2.2)

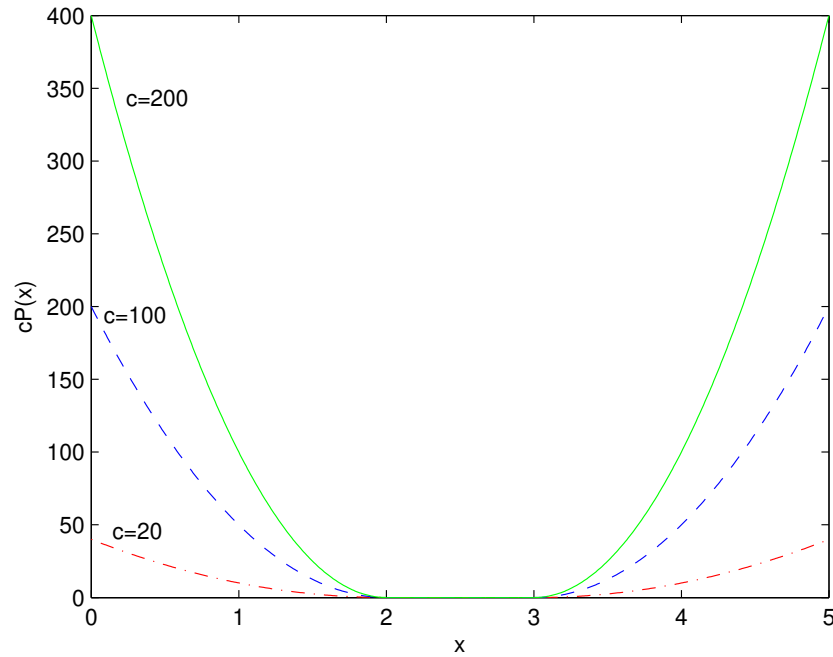


Figura 2.4: Representação gráfica da função penalidade quadrática com diferentes valores para o parâmetro c .

com um valor de c muito grande. Essa estratégia recebe o nome *shortcut* (atalho) em [Fle81] e nem sempre é eficiente.

Esquematizaremos o método de penalidade no Algoritmo 2.1.1:

Algoritmo 2.1.1 *Método de Penalidade*

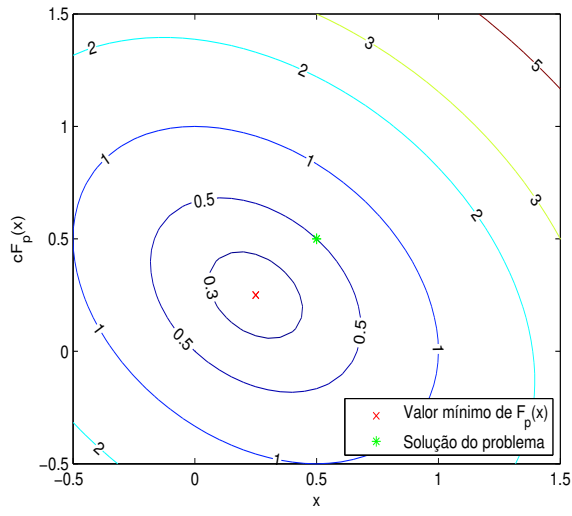
Dados $x^0 \in \mathbb{R}^n, c^0$ e $\beta > 1$

$k = 0$

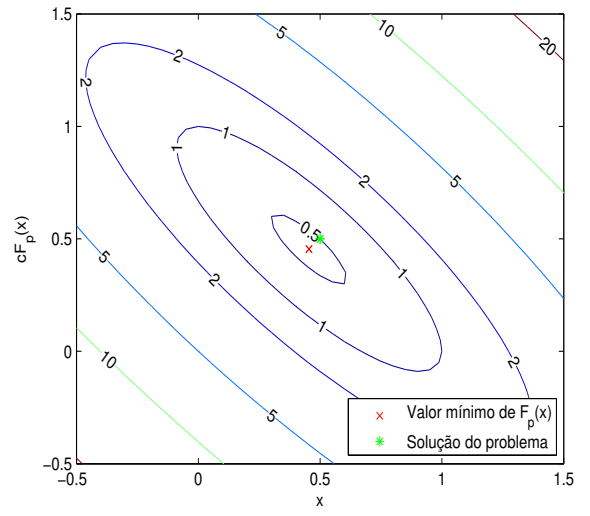
Repetir até a convergência

1. $x^{k+1} \in \operatorname{argmin} \{f(x) + c^k \mathcal{P}(x)\}$
2. $c^{k+1} = \beta c^k$
3. $k = k + 1$

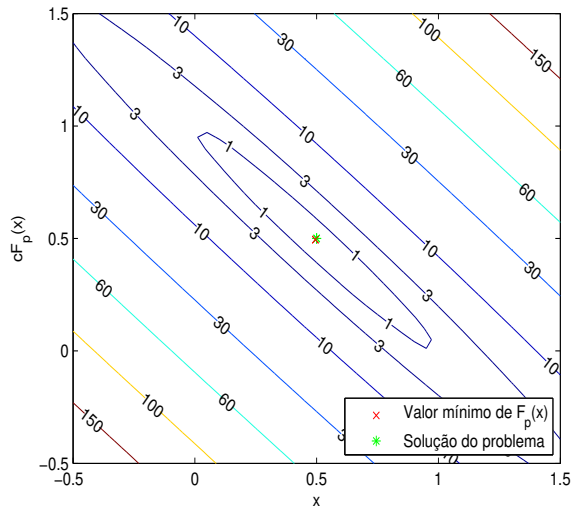
Os passos principais desse método são a resolução do subproblema (2.2) e a atualização do parâmetro de penalidade.



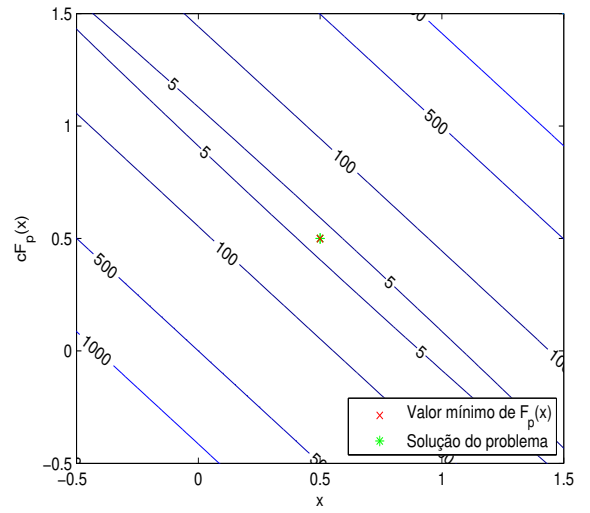
(a) $c = 1$



(b) $c = 10$



(c) $c = 100$



(d) $c = 1000$

Figura 2.5: Função auxiliar $\mathcal{F}_{\mathcal{P}}(x)$ para o problema (2.3) com diferentes parâmetros de penalidade.

Como em cada iteração k , $c^{k+1} > c^k$ significa que c^k sempre aumenta e só se consegue mostrar a convergência do Algoritmo 2.1.1 quando $c^k \rightarrow \infty$. Isso pode gerar problema de mal condicionamento.

A dificuldade do método de penalidade, descrito no Algoritmo 2.1.1, cresce com o aumento do parâmetro c . Quando $c \rightarrow \infty$ os subproblemas (2.2) tornam-se computacionalmente mais difíceis de serem resolvidos. Vejamos como isso acontece com o seguinte exemplo,

$$\begin{aligned} &\text{Minimizar} && f(x) = (x_1 + 1)^2 + (x_2 - 1)^2 \\ &\text{sujeito a} && x_1 = 0 \end{aligned} \tag{2.4}$$

cuja solução é $x^* = (0, 1)^T$. A função a ser minimizada do subproblema (2.2) para (2.4) é

$$\mathcal{F}_{\mathcal{P}}(x) = (x_1 + 1)^2 + (x_2 - 1)^2 + \frac{c}{2}x_1^2. \tag{2.5}$$

A Hessiana da função $\mathcal{F}_{\mathcal{P}}(x)$ (2.5) é

$$\nabla_{xx}^2 \mathcal{F}_{\mathcal{P}}(x) = \begin{pmatrix} 2 + c & 0 \\ 0 & 2 \end{pmatrix}.$$

Observamos que o número de condição da matriz $\nabla_{xx}^2 \mathcal{F}_{\mathcal{P}}(x)$ tende para infinito quando $c \rightarrow \infty$. Esse fato implica no mal condicionamento da matriz Hessiana $\nabla_{xx}^2 \mathcal{F}_{\mathcal{P}}(x)$ provocando uma instabilidade numérica no método. Veremos no Capítulo 3 que um dos problemas que a função Lagrangeana aumentada elimina é a instabilidade originada na matriz Hessiana presente no método de penalidade.

2.2 Método de Barreira

Consideremos o seguinte problema:

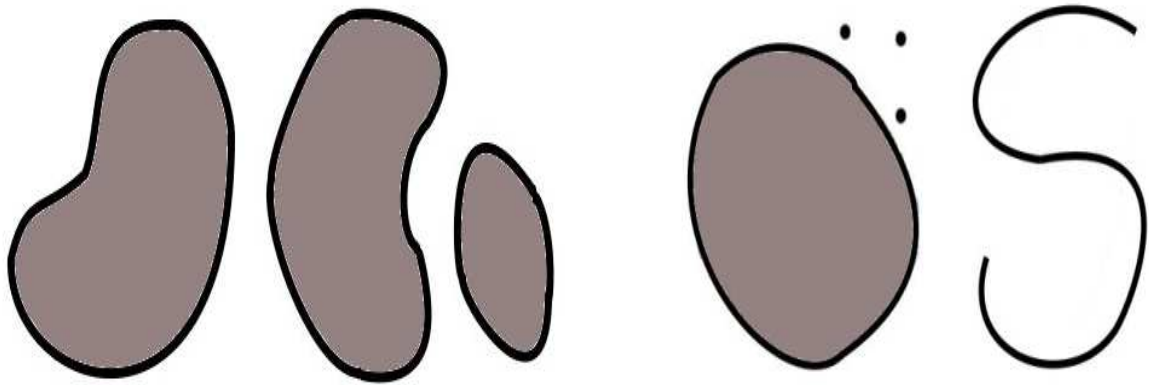
$$\begin{aligned} &\text{Minimizar} && f(x) \\ &\text{sujeito a} && x \in \Omega, \end{aligned} \tag{2.6}$$

onde $\Omega = \{x \mid g_i(x) \geq 0, i = 1, \dots, m\}$ e f e g_i são funções contínuas.

Seja $\overset{\circ}{\Omega}$ o interior do conjunto Ω definido como,

$$\overset{\circ}{\Omega} = \{x \in \Omega \mid g_i(x) > 0, i = 1, \dots, m\}.$$

Assumimos que $\overset{\circ}{\Omega}$ é um conjunto não-vazio e que qualquer ponto factível que não está em $\overset{\circ}{\Omega}$ pode ser aproximado arbitrariamente por um vetor de $\overset{\circ}{\Omega}$, isto é, dado qualquer x factível, $x \in \Omega$ e qualquer $\delta > 0$, existe $\tilde{x} \in \overset{\circ}{\Omega}$ tal que $\|\tilde{x} - x\| \leq \delta$. Esse tipo de conjunto é conhecido como *robusto*. Na Figura 2.6 temos exemplos de conjuntos robustos e não robustos.



(a) Robusto

(b) Não Robusto

Figura 2.6: Exemplos de conjuntos robustos e não robustos.

No método de barreira, consideramos inicialmente uma função $\mathcal{B}(x)$ que é definida no interior do conjunto Ω .

Uma *função barreira* $\mathcal{B}(x)$ definida no interior de Ω é tal que:

- $\mathcal{B}$ é contínua;
- $\mathcal{B}(x) \geq 0$;
- $\mathcal{B}(x) \rightarrow \infty$ quando x se aproxima da fronteira de Ω .

As funções barreiras mais citadas na literatura são a inversa, definida como,

$$\mathcal{B}_{\mathcal{I}}(x) = \sum_{i=1}^m \frac{1}{g_i(x)}$$

e a a logarítmica

$$\mathcal{B}_{\mathcal{L}}(x) = - \sum_{i=1}^m \ln g_i(x).$$

A função barreira inversa também denominada por barreira clássica foi estudada por Carrol [CF61] e a logarítmica foi estudada por Frisch [Fri55]. Note que ambas funções citadas acima são convexas se todas as restrições $-g_i(x)$ são convexas.

Para ilustrar a forma da função barreira consideremos as seguintes restrições,

$$\begin{aligned} g_1(x) &= 2 - x \geq 0 \\ g_2(x) &= 4 - x \geq 0. \end{aligned}$$

Para essas restrições, a função barreira inversa é representada como,

$$\mathcal{B}_{\mathcal{I}}(x) = \frac{1}{2-x} + \frac{1}{4-x}$$

e a logarítmica como,

$$\mathcal{B}_{\mathcal{L}}(x) = -\ln(2-x) - \ln(4-x).$$

Na Figura 2.7 apresentamos o gráfico da função barreira $B(x) = -\ln(2-x) - \ln(4-x)$ com diferentes parâmetros de penalidade c .

Notemos pela Figura 2.7 que o termo de barreira $\frac{1}{c}\mathcal{B}(x)$ se aproxima de zero para todos os pontos interiores $x \in \overset{\circ}{\Omega}$ quando $c \rightarrow \infty$.

O método de barreira consiste em resolver uma seqüência de problemas irrestritos

$$\begin{aligned} &\text{Minimizar } \mathcal{F}_{\mathcal{B}}(x) = f(x) + \frac{1}{c}\mathcal{B}(x) \\ &\text{sujeito a } x \in \overset{\circ}{\Omega}, \end{aligned} \tag{2.7}$$

onde $\mathcal{F}_{\mathcal{B}}(x)$ é a função auxiliar de barreira e c é o parâmetro de penalidade. A seqüência de soluções geradas $\{x^k\}$ do problema (2.7) está relacionada com uma seqüência de parâ-

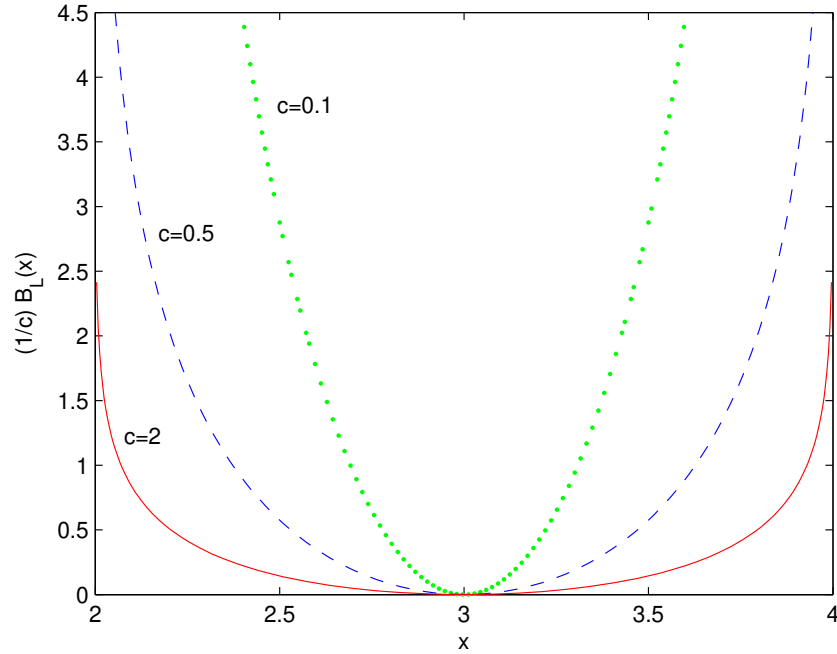


Figura 2.7: Representação gráfica da função barreira com diferentes parâmetros c .

metros $\{c^k\}$ onde

$$c^{k+1} > c^k > 0, \quad k = 0, 1, \dots, \quad c^k \rightarrow \infty.$$

Desde que a função barreira seja definida somente no interior de Ω , os sucessivos x^k 's gerados pelo problema (2.7) devem ser pontos interiores. Como no método de penalidade, os subproblemas (2.7) podem ser resolvidos por qualquer método para problemas irrestritos, como o método de Newton, quase-Newton, gradientes conjugados entre outros.

A função barreira logarítmica tem sido muito pesquisada em métodos que geram soluções pertencentes ao interior de Ω . Esses métodos são geralmente referidos como *métodos de pontos interiores* e têm sido amplamente aplicados a problemas de programação quadrática e linear seguindo o trabalho do matemático indiano Karmarkar [Kar84].

Agora analisaremos o comportamento da função auxiliar de barreira logarítmica $\mathcal{F}_B(x)$ para diferentes valores do parâmetro c . Para isto, utilizaremos o problema exemplo

[Ber99].

$$\begin{array}{ll} \text{Minimizar} & \frac{1}{2}(x_1^2 + x_2^2) \\ \text{sujeito a} & x_1 - 2 \geq 0 \end{array} \quad (2.8)$$

cujas a solução ótima é $x^* = (2, 0)^T$.

A função auxiliar de barreira para esse problema tem a seguinte forma:

$$\mathcal{F}_B(x) = \frac{1}{2}(x_1^2 + x_2^2) - \frac{1}{c} \ln(x_1 - 2). \quad (2.9)$$

Na Figura 2.8 temos os gráficos da função $\mathcal{F}_B(x)$ para os valores de parâmetros, $c = 5$, $c = 10$, $c = 150$ e $c = 500$. Note que para $c = 500$ o ponto mínimo da função $\mathcal{F}_B(x)$ está bem próximo da solução ótima do problema (2.8). A estratégia *shortcut*, vista na Seção 2.1 para a penalização externa, pode ser empregada também para a penalização interna utilizando um c muito grande. Infelizmente essa estratégia nem sempre funciona.

Como ocorre no método da penalidade, a dificuldade em resolver os subproblemas (2.7) cresce com o aumento do parâmetro c . Essa dificuldade se reflete em instabilidade numérica e está relacionada com o número de condição da Hessiana da função $\mathcal{F}_B(x)$. Utilizaremos o exemplo (2.8) para ver como isso ocorre. A Hessiana da função (2.9) é

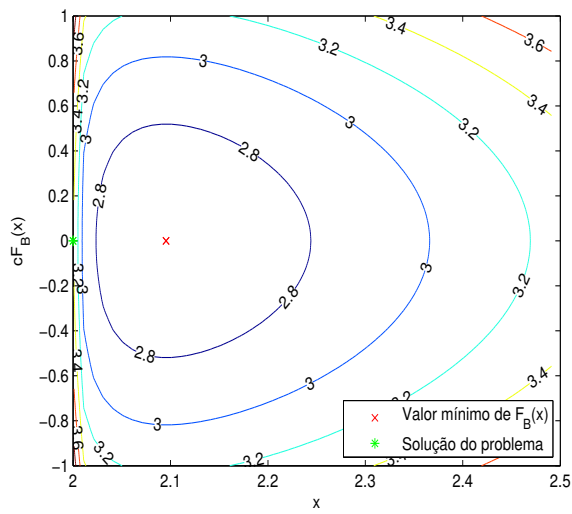
$$\nabla_{xx}^2 \mathcal{F}_B(x) = \begin{pmatrix} 1 + \frac{1}{c(x_1 - 2)^2} & 0 \\ 0 & 1 \end{pmatrix}.$$

Os pontos estacionários de (2.9) para $x_1 \geq 2$ são da forma $x = \left(1 + \sqrt{1 + \frac{1}{c}}, 0\right)^T$. Substituindo esse ponto na Hessiana $\nabla_{xx}^2 \mathcal{F}_B(x)$ obtemos,

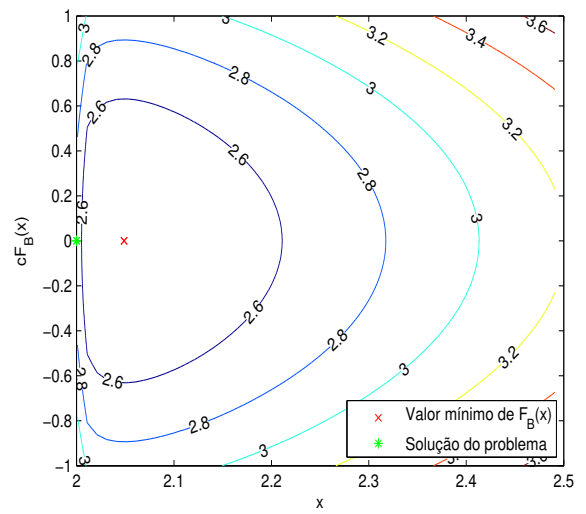
$$\nabla_{xx}^2 \mathcal{F}_B(x) = \begin{pmatrix} 1 + \frac{1}{2c + 1 - 2\sqrt{c^2 + c}} & 0 \\ 0 & 1 \end{pmatrix}.$$

Quando $\lim_{c \rightarrow \infty} 1 + \frac{1}{2c + 1 - 2\sqrt{c^2 + c}} = \infty$, o número de condição da matriz $\nabla_{xx}^2 \mathcal{F}_B(x)$ também tende ao infinito. Assim podemos comprovar a dificuldade em resolver os subproblemas (2.7) à medida que c cresce.

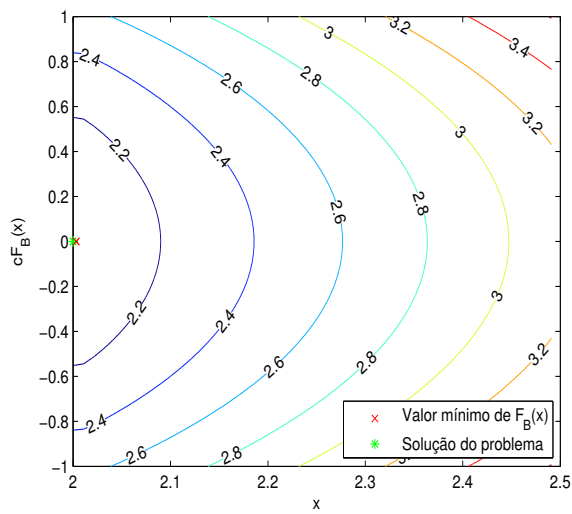
No algoritmo (2.2.1) descrevemos o método de barreira. Os passos principais desse



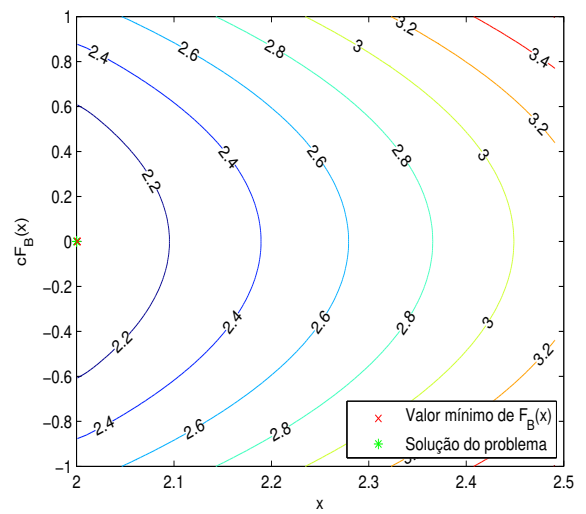
(a) $c = 5$



(b) $c = 10$



(c) $c = 150$



(d) $c = 500$

Figura 2.8: Função auxiliar $\mathcal{F}_B(x)$ para o problema (2.8) com diferentes parâmetros de penalidade.

método são encontrar uma solução x^k para o subproblema (2.7) e atualizar c^k aumentando seu valor por um fator β .

Algoritmo 2.2.1 *Método de Barreira*

Dados $x^0 \in \overset{\circ}{\Omega}, c^0$ e $\beta > 1$

$k = 0$

Repetir até a convergência

1. $x^{k+1} \in \operatorname{argmin}_{x \in \overset{\circ}{\Omega}} \{f(x) + \frac{1}{c^k} \mathcal{B}(x)\}$
2. $c^{k+1} = \beta c^k$
3. $k = k + 1$

Praticamente as mesmas propriedades de convergência são válidas para o método de penalidade quanto para o método de barreira.

Os métodos de penalidade interna e externa são similares em alguns aspectos. Observe que o Algoritmo 2.2.1 é bem parecido com o Algoritmo 2.1.1 visto na seção anterior. Tendo em vista que os objetivos de ambos os métodos são parecidos, penalizar para aproximar-se da solução, a forma das funções utilizadas nos métodos de penalidade e de barreira são diferentes. Um penaliza pontos inefectíveis e outro penaliza pontos que estão perto da fronteira da região factível.

Neste capítulo apresentamos os métodos clássicos de penalidades externa e interna. Esses combinados com outros métodos, como por exemplo, o método dual Lagrange geram outros métodos mais eficazes. No próximo capítulo veremos o método da Lagrangeana aumentada que é uma extensão do método de penalidade externa.

Capítulo 3

Lagrangeana Aumentada

O método da Lagrangeana aumentada e da barreira modificada, este último veremos no Capítulo 4, são oriundos dos métodos clássicos de penalidades. Os métodos de penalidades externa e interna geram, respectivamente, seqüência de soluções primais infactíveis e factíveis. Para que esses métodos convirjam é preciso que o parâmetro de penalidade c cresça infinitamente. Isso pode causar uma instabilidade numérica nesses métodos.

Hestenes [Hes69], Powell [Pow69] e Polyak [Pol92] introduziram extensões dos métodos clássicos de penalidades. Isso foi motivado devido a algumas desvantagens encontradas nos métodos de penalidade e da barreira. Hestene e Powell criaram a Lagrangeana aumentada, uma extensão da penalidade externa. Polyak introduziu a barreira modificada baseada na penalidade interna, permitindo a esse método gerar pontos infactíveis. Uma das diferenças dessas extensões com os métodos clássicos de penalidade, é que a Lagrangeana aumentada e barreira modificada além de gerarem soluções aproximadas primais geram também soluções aproximadas duais.

Daqui em diante, focalizaremos nossa atenção, neste capítulo, na Lagrangeana aumentada.

O método de Lagrangeana aumentada, chamado também de métodos dos multiplicadores, foi apresentando à comunidade acadêmica em uma conferência de 1968 [Roc73b] pelos matemáticos Hestenes [Hes69] e Powell [Pow69]. Cada um desenvolveu seu trabalho de forma independente.

O método dos multiplicadores foi aplicado inicialmente à problemas de otimização com restrições de igualdade. Outros pesquisadores como Miele, Cragg, Levy, Moseley dentre outros [MCIL71, MCL71, MMLC72] trabalharam nesses métodos modificando-

os computacionalmente ou desenvolvendo novas técnicas de atualizações dos multiplicadores. Mas todas essas técnicas foram desenvolvidas para problemas com restrições de igualdade. Rockafellar [Roc74] foi o primeiro a propor um algoritmo análogo ao método dos multiplicadores de Hestenes e Powell para problemas de desigualdade.

Para facilitar a explanação do método da Lagrangeana aumentada apresentaremos primeiro sua aplicação em problemas com restrições de igualdade e depois estenderemos sua aplicação para os problemas com restrições de desigualdade.

3.1 A Função Lagrangeana Aumentada - Restrições de igualdade

Consideremos o seguinte problema de otimização com restrições de igualdade

$$\begin{aligned} &\text{Minimizar} && f(x) \\ &\text{sujeito a} && h_j(x) = 0, \quad j = 1, \dots, p \end{aligned} \quad (3.1)$$

onde $n > p$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h_j : \mathbb{R}^n \rightarrow \mathbb{R}$, $j = 1, \dots, p$ e $f, h_1, \dots, h_p \in C^1$.

Inserindo perturbações ξ_j nas restrições de igualdade de (3.1), obtemos o seguinte problema perturbado:

$$\begin{aligned} &\text{Minimizar} && f(x) \\ &\text{sujeito a} && h_j(x) = \xi_j, \quad j = 1, \dots, p. \end{aligned} \quad (3.2)$$

Aplicando a penalização quadrática em (3.2) obtemos,

$$\text{Minimizar} \quad f(x) + \frac{c}{2} \sum_{j=1}^p [h_j(x) - \xi_j]^2, \quad (3.3)$$

Notemos que com as perturbações ξ_j , $j = 1, \dots, p$, c não precisa tender ao infinito. Expandindo a função objetivo do problema (3.3) obtemos,

$$f(x) + \frac{c}{2} \sum_{j=1}^p h_j^2(x) - \sum_{j=1}^p c\xi_j h_j(x) + \frac{c}{2} \sum_{j=1}^p \xi_j^2. \quad (3.4)$$

Considerando $\lambda_j = c\xi_j$ para $j = 1, \dots, p$ e ignorando a constante $\frac{c}{2} \sum_{j=1}^p \xi_j^2$ ficamos com

a seguinte função que é denominada *função da Lagrangeana aumentada*:

$$\mathcal{L}_{\mathcal{A}}(x, \lambda) = f(x) - \sum_{j=1}^p \lambda_j h_j(x) + \frac{c}{2} \sum_{j=1}^p h_j^2(x). \quad (3.5)$$

Podemos ver que a função (3.5) é a função auxiliar de penalidade $\mathcal{F}_{\mathcal{P}}(x)$ aplicada ao seguinte problema:

$$\begin{aligned} &\text{Minimizar} && f(x) - \sum_{j=1}^p \lambda_j h_j(x) \\ &\text{sujeito a} && h_j(x) = 0 \quad j = 1, \dots, p. \end{aligned}$$

Observamos que a função objetivo desse problema é a Lagrangeana do problema (3.1) sujeita às mesmas restrições de igualdade. Isso significa que a Lagrangeana aumentada escrita em (3.5) pode ser interpretada como a função auxiliar de penalidade $\mathcal{F}_{\mathcal{P}}(x)$ associada ao problema (3.1), ou seja,

$$\mathcal{F}_{\mathcal{P}}(x) = f(x) - \sum_{j=1}^p \lambda_j h_j(x) + \frac{c}{2} \sum_{j=1}^p h_j^2(x) = \mathcal{L}_{\mathcal{A}}(x, \lambda, c).$$

Desse modo, podemos observar também que a função (3.5) difere da função Lagrangeana do problema (3.1) pelo termo quadrático. Por outro lado, o termo envolvendo os λ_j 's diferenciam a função definida em (3.5) da função auxiliar de penalidade associada ao problema (3.1). Assim, podemos dizer que a função Lagrangeana aumentada $\mathcal{L}_{\mathcal{A}}$ é uma combinação das funções Lagrangeana $\mathcal{L}$ e penalidade $\mathcal{P}$.

Seja $x^* = (x_1^*, x_2^*, \dots, x_n^*)^T$ e $\lambda^* = (\lambda_1^*, \lambda_2^*, \dots, \lambda_p^*)^T$ a solução ótima primal e dual, respectivamente, do problema (3.1). Observe que o gradiente da Lagrangeana aumentada é igual a zero no ponto (x^*, λ^*) para qualquer c ,

$$\nabla \mathcal{L}_{\mathcal{A}}(x^*, \lambda^*) = \nabla f(x^*) - \sum_{j=1}^p \lambda_j^* \nabla h_j(x^*) + c \sum_{j=1}^p h_j(x^*) \nabla h_j(x^*) = 0. \quad (3.6)$$

Isso não necessariamente ocorre com a função de penalidade quadrática. Sabendo que

x^* é a solução de (3.1) e aplicando a penalidade quadrática a esse problema,

$$\mathcal{F}_{\mathcal{P}}(x) = f(x) + \frac{c}{2} \sum_{j=1}^p h_j(x)^2, \quad (3.7)$$

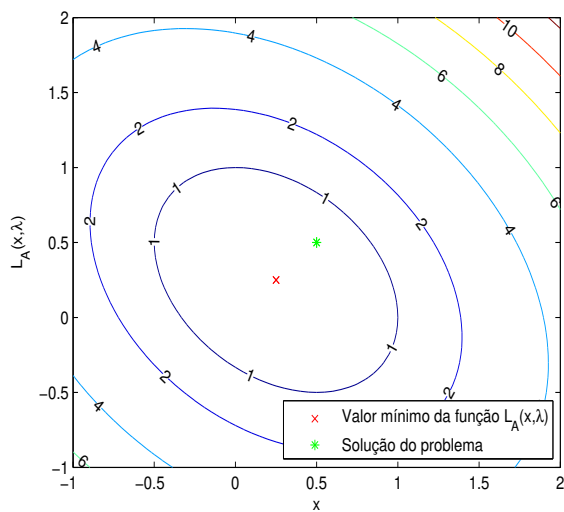
note que o gradiente da função $\mathcal{F}_{\mathcal{P}}$ só será igual a zero se $\nabla f(x^*)$ também for,

$$\nabla_x \mathcal{F}_{\mathcal{P}}(x^*) = \nabla f(x^*) + c \sum_{j=1}^p h_j(x^*) \nabla h_j(x^*) = \nabla f(x^*). \quad (3.8)$$

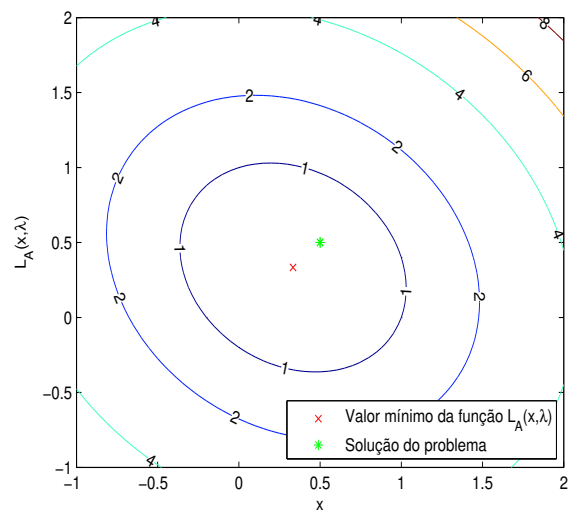
Para ilustrar o comportamento da função Lagrangeana aumentada utilizaremos o exemplo (2.3) da Seção 2.1, onde as soluções ótimas do primal e dual são respectivamente, $x^* = (\frac{1}{2}, \frac{1}{2})^T$ e $\lambda^* = 1$. A função Lagrangeana aumentada para esse problema é

$$\mathcal{L}_{\mathcal{A}}(x, \lambda) = x_1^2 + x_2^2 - \lambda(x_1 + x_2 - 1) + \frac{c}{2}(x_1 + x_2 - 1)^2. \quad (3.9)$$

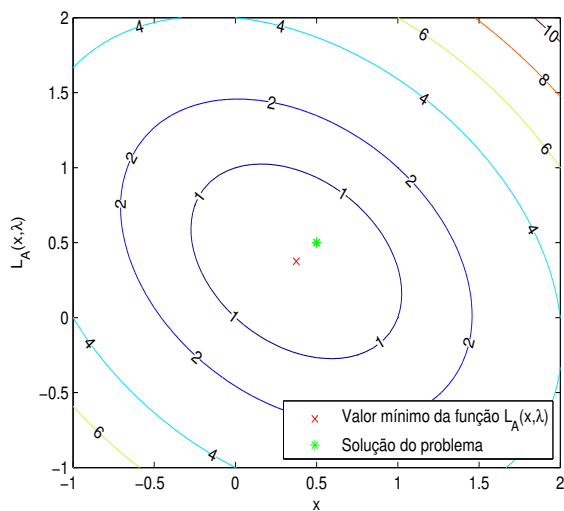
Na Figura (3.1) apresentamos o gráfico da função (3.9) para diferentes valores de λ e c : (a) $\mathcal{L}_{\mathcal{A}}(x, 0)$, $c = 1$; (b) $\mathcal{L}_{\mathcal{A}}(x, \frac{1}{2})$, $c = \frac{1}{2}$; (c) $\mathcal{L}_{\mathcal{A}}(x, \frac{1}{2})$, $c = 1$ e (d) $\mathcal{L}_{\mathcal{A}}(x, \frac{3}{5})$, $c = 2$. Para os valores $c = 1$ e $\lambda = 0$ a função $\mathcal{L}_{\mathcal{A}}(x, 0)$ coincide com a função auxiliar de penalidade $\mathcal{F}_{\mathcal{P}}(x)$, com parâmetro de penalidade igual a 1. Os minimizadores de cada caso são respectivamente, $x = (\frac{1}{4}, \frac{1}{4})^T$, $x = (\frac{1}{3}, \frac{1}{3})^T$, $x = (\frac{3}{8}, \frac{3}{8})^T$ e $x = (\frac{4}{9}, \frac{4}{9})^T$. Note que para os valores $c = 2$ e $\lambda = \frac{3}{5}$ o ponto mínimo de $\mathcal{L}_{\mathcal{A}}(x, \lambda)$ está bem próximo da solução ótima do problema (2.3).



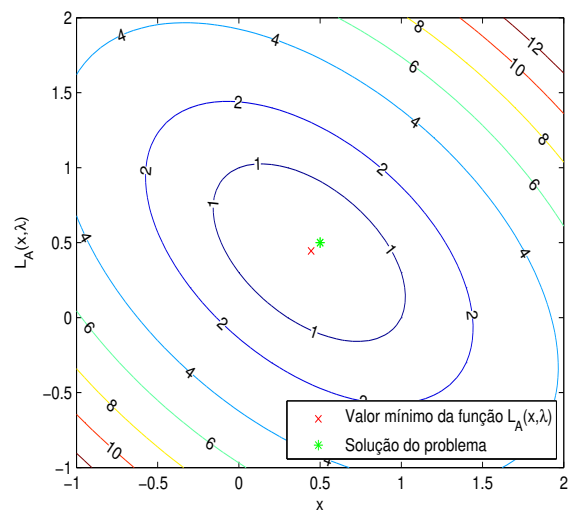
(a) $\lambda = 0$ e $c = 1$



(b) $\lambda = \frac{1}{2}$ e $c = \frac{1}{2}$



(c) $\lambda = \frac{1}{2}$ e $c = 1$



(d) $\lambda = \frac{3}{5}$ e $c = 2$

Figura 3.1: Gráficos da função Lagrangeana aumentada (3.9) com diferentes valores de parâmetros de penalidade e multiplicadores de Lagrange.

3.2 A Função Lagrangeana Aumentada - Restrições de desigualdade

Da mesma forma como abordamos o problema com restrições de igualdade, conduziremos a definição da função Lagrangeana aumentada para o problema com restrições de desigualdade. Consideremos o seguinte problema:

$$\begin{aligned} &\text{Minimizar} && f(x) \\ &\text{sujeito a} && g_i(x) \geq 0, \quad i = 1, \dots, m \end{aligned} \quad (3.10)$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$ e $f, g_1, \dots, g_m \in C^1$.

Vamos representar o problema (3.10) transformando as restrições de desigualdade em restrições de igualdade inserindo variáveis auxiliares de excesso como,

$$\begin{aligned} &\text{Minimizar} && f(x) \\ &\text{sujeito a} && g_i(x) - s_i = 0, \quad i = 1, \dots, m \\ &&& s_i \geq 0, \quad i = 1, \dots, m. \end{aligned} \quad (3.11)$$

Consideremos um subproblema de (3.11) para o qual são ignoradas temporariamente as restrições $s_i \geq 0$, $i = 1, \dots, m$,

$$\begin{aligned} &\text{Minimizar} && f(x) \\ &\text{sujeito a} && g_i(x) - s_i = 0, \quad i = 1, \dots, m. \end{aligned}$$

Definimos então uma função Lagrangeana aumentada auxiliar para esse subproblema como:

$$\mathcal{L}_{\mathcal{A}}'(x, \mu, s) = f(x) - \sum_{i=1}^m \mu_i [g_i(x) - s_i] + \frac{c}{2} \sum_{i=1}^m [g_i(x) - s_i]^2. \quad (3.12)$$

A Lagrangeana aumentada (3.12) nos permite calcular a condição de estacionariedade em relação às variáveis de excesso s_i , ou seja,

$$\frac{\partial}{\partial s_i} \mathcal{L}_{\mathcal{A}}'(x, \mu, s) \Big|_{s_i = \tilde{s}_i} = 0.$$

Dessa condição concluímos que

$$\mu_i - c[g_i(x) - \tilde{s}_i] = 0. \quad (3.13)$$

Da equação (3.13) obtemos a solução $\tilde{s}_i$ para o mínimo da Lagrangeana aumentada auxiliar $\mathcal{L}_{\mathcal{A}}'$,

$$\tilde{s}_i = -\frac{\mu_i}{c} + g_i(x). \quad (3.14)$$

Considerando agora a restrição sobre a variável de excesso $s_i \geq 0$ no ponto de mínimo irrestrito (3.14) da Lagrangeana aumentada, obtemos:

$$\min_s \{\mathcal{L}_{\mathcal{A}}'(x, \mu, s)\} = f(x) - \sum_{i=1}^m \begin{cases} \frac{\mu_i^2}{2c} & , \text{se } \tilde{s}_i \geq 0 \\ \mu_i g_i(x) - \frac{c}{2} g_i(x)^2 & , \text{se } \tilde{s}_i < 0 \end{cases}. \quad (3.15)$$

Assim definimos a função Lagrangeana aumentada associada ao problema (3.1) como

$$\mathcal{L}_{\mathcal{A}}(x, \mu, s) = \min_s \{\mathcal{L}_{\mathcal{A}}'(x, \mu, s)\}$$

que resulta na expressão:

$$\mathcal{L}_{\mathcal{A}}(x, \mu) = f(x) - \sum_{i=1}^m \begin{cases} \frac{\mu_i^2}{2c} & , \text{se } g_i(x) \geq -\frac{\mu_i}{c} \\ \mu_i g_i(x) - \frac{c}{2} g_i(x)^2 & , \text{caso contrário.} \end{cases} \quad (3.16)$$

Para o problema com restrição de desigualdade (2.8) da Seção 2.2, a função Lagrangeana aumentada tem a seguinte forma,

$$\mathcal{L}_{\mathcal{A}}(x, \mu) = \frac{1}{2}(x_1^2 + x_2^2) - \begin{cases} \frac{\mu^2}{2c}, & \text{se } x_1 \geq 2 - \frac{\mu}{c} \\ \mu(x_1 - 2) - \frac{c}{2}(x_1 - 2)^2, & \text{caso contrário.} \end{cases}$$

Nesta seção apresentamos os passos para a construção da função (3.16). Esta função será utilizada para encontrar a solução primal no método da Lagrangeana aumentada que veremos na Seção 3.3.

3.3 Método da Lagrangeana Aumentada

Consideremos o problema com restrições de desigualdade (3.10). A aproximação primal desse problema é

$$\text{Minimizar } f(x) + \frac{c}{2} \sum_{i=1}^m \max\{0, -g_i(x)\}^2, \quad c \rightarrow \infty. \quad (3.17)$$

A aproximação dual, baseada nas condições de KKT, substitui o problema (3.10) pela seqüência de problemas da forma,

$$\text{Minimizar } f(x) - \sum_{i=1}^m \mu_i g_i(x). \quad (3.18)$$

Os μ_i 's de (3.18) constituem uma seqüência de maximizadores para o problema dual ordinário [Roc73a],

$$\max \Phi(\mu) \triangleq \inf \mathcal{L}_o(x, \mu) \quad (3.19)$$

onde

$$\mathcal{L}_o(x, \mu) = \begin{cases} f(x) - \sum_{i=1}^m \mu_i g_i(x), & \text{se } \mu \geq 0 \\ -\infty, & \text{caso contrário.} \end{cases}$$

A desvantagem do método de penalidade é que quando $c \rightarrow \infty$, o problema (3.17) pode-se tornar instável. Já para o método dual, caso a função objetivo f não seja estritamente convexa, a seqüência gerada $\{x^k\}$ do problema (3.18) não é necessariamente a seqüência do problema primal (3.10). A Lagrangeana aumentada mescla esses dois métodos para eliminar as desvantagens inerentes aos métodos de penalidade e dual.

Na figura 3.2 temos os passos principais do método da Lagrangeana aumentada.

Como comentado no início deste capítulo, o método Lagrangeana aumentada apresenta dois ciclos: interno e externo. O ciclo interno envolve a minimização da função Lagrangeana aumentada onde o parâmetro de penalidade e os multiplicadores de Lagrange são fixos. Nesse ciclo pode-se usar qualquer método para minimização irrestrita. Neste trabalho optamos pelo método de Newton:

$$\nabla^2 \mathcal{L}_A(x^k, \lambda^k) \Delta x = -\nabla \mathcal{L}_A(x^k, \lambda^k) \quad (3.20)$$

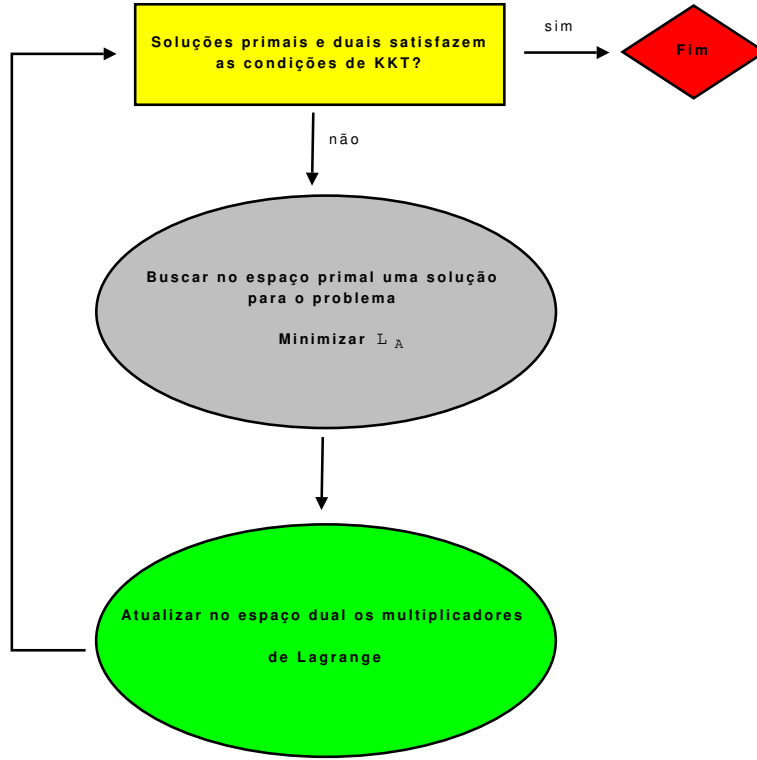


Figura 3.2: Passos principais do método da Lagrangeana aumentada.

onde Δx é a direção de descida [BS93, DS93, Lue86, PSU93]. Esse ciclo acaba quando $\nabla \mathcal{L}_A(x^k, \lambda^k) \leq \epsilon_1$, onde ϵ_1 é um valor pequeno. No Apêndice B temos um pouco da teoria sobre o método de Newton.

O ciclo externo constitui-se nas atualizações dos multiplicadores de Lagrange e do parâmetro de penalidade c . Primeiramente, vamos ver a atualização dos multiplicadores de Lagrange para os problemas com restrições de igualdade. Faremos isso para facilitar o entendimento para os problemas com restrições de desigualdade.

Calculando o gradiente da função(3.5) obtemos:

$$\nabla \mathcal{L}_A(x, \lambda) = \nabla f(x) - \sum_{j=1}^m [\lambda_j - ch_j(x)] \nabla h_j(x). \quad (3.21)$$

Comparando (3.21) com a condição de estacionaridade do problema (3.1)

$$\nabla \mathcal{L}(x, \lambda) = \nabla f(x) - \sum_{j=1}^p \lambda_j \nabla h_j(x),$$

temos a seguinte regra de atualização:

$$\lambda_i^{k+1} = \lambda_i^k - c^k h_i(x^{k+1}). \quad (3.22)$$

Reescrevendo (3.22) de forma vetorial temos,

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

Observe que essa atualização nada mais é que a aplicação do método do gradiente¹ ao problema,

$$\text{Maximizar } \mathcal{L}_{\mathcal{A}}(x^{k+1}, \lambda) = f(x^{k+1}) - \lambda^T h(x^{k+1}) + \frac{c^k}{2} \|h(x^{k+1})\|^2 \quad (3.23)$$

em relação à variável dual λ . A direção de subida é dada pelo vetor $-c^k h(x^{k+1})$.

Para problemas com desigualdade a regra de atualização dos multiplicadores de Lagrange é obtida de maneira análoga ao problema (3.1).

Calculando o gradiente da função (3.16) obtemos:

$$\nabla \mathcal{L}_{\mathcal{A}}(x, \mu) = \nabla f(x) - \sum_{i=1}^m \begin{cases} 0 & , \text{ se } g_i(x) \geq -\frac{\mu_i}{c} \\ [\mu_i - c g_i(x)] \nabla g_i(x) & , \text{ caso contrário.} \end{cases} \quad (3.24)$$

Comparando (3.24) com a condição de estacionaridade do problema (3.10)

$$\nabla \mathcal{L}(x, \mu) = \nabla f(x) - \sum_{i=1}^m \mu_i \nabla g_i(x),$$

¹Esse método é também conhecido como método de subida quando o problema for de maximizar uma função objetivo e método de descida caso queira encontrar o valor mínimo de uma função [BS93, DS93, Lue86, PSU93].

obtemos a seguinte atualização para os multiplicadores de Lagrange:

$$\mu_i^{k+1} = \begin{cases} 0 & , \text{ se } g_i(x) \geq -\frac{\mu_i}{c} \\ \mu_i^k - c^k g_i(x^{k+1}) & , \text{ caso contrário.} \end{cases} \quad (3.25)$$

O x^{k+1} utilizado nas atualizações (3.22) e (3.25) é a solução aproximada obtida no ciclo interno.

É no ciclo externo também que ocorre a atualização do parâmetro de penalidade c . Apresentaremos agora, algumas atualizações do parâmetro c encontradas em alguns trabalhos da literatura. No artigo de Hestenes [Hes69] não se dá muita ênfase nas regras de atualização, mas ressalta-se que não é difícil obter um critério o qual assegure a convergência do método. Vale lembrar que neste artigo, Hestenes sugere um método para problemas com restrições de igualdade. Powell demonstrou em [Pow69], que se a condição suficiente de segunda ordem for satisfeita, o algoritmo do método dos multiplicadores aplicado ao problema (3.1) converge localmente com um raio linear, sem precisar que $c \rightarrow \infty$. No trabalho de Miele, Moseley, Levy e Coggins [MMLC72] são sugeridas duas técnicas de atualização de c . Uma foi desenvolvida para ser usada com o método do gradiente ordinário

$$c^{k+1} = \frac{1}{2 \sum_{j=1}^p \left[\frac{\partial h_j(x^{k+1})}{\partial x^{k+1}} \right]^2}$$

e a outra com os métodos de gradientes conjugados e de quase-linearização modificado

$$c^{k+1} = \sum_{j=1}^p \frac{\lambda_j h_j(x^{k+1})}{h_j(x^{k+1})^2}.$$

Esses métodos mencionados são utilizados no ciclo interno. Uma outra abordagem para atualização vem do algoritmo LANCELOT² implementado por Conn, Gould e Toint [CGT92]. A atualização de c funciona assim: quando x^{k+1} está longe da factibilidade mantém-se os multiplicadores de Lagrange inalterados, $\lambda^{k+1} = \lambda^k$, e aumenta-se o valor

²Esse algoritmo é designado a resolver problemas com restrições de igualdade e canalizadas, isto é, $l_j \leq x_j \leq u_j, j = 1, \dots, p$, onde l_j é o limite inferior (*lower*) e u_j é o limite superior (*upper*). A função Lagrangeana aumentada do LANCELOT só incorpora as restrições de igualdade.

de c^{k+1} , ou seja,

$$c^{k+1} = \beta c^k, \text{ onde } \beta > 1 \quad (3.26)$$

para penalizar a função $\mathcal{L}_A$. Agora quando x^k está bem próximo da factibilidade atualizam-se os multiplicadores pela fórmula (3.22) e não se aumenta c^{k+1} . Para problemas com restrições de desigualdade Rockafellar [Roc73a] mostrou que se $-g_i(x), i = 1, \dots, m$, são convexas e não necessariamente diferenciáveis, o método converge globalmente com um valor de c fixo.

Neste trabalho implementamos o método da Lagrangeana aumentada com c variando conforme a regra 3.26 e com um valor $c > 1$ fixo.

Na Seção 2.1, vimos que as soluções geradas pelos subproblemas (2.2), aproximam-se da solução ótima quando o parâmetro de penalidade tende a infinito. Isso pode ocasionar problemas de mal-condicionamento e instabilidade numérica associados à função $\mathcal{F}_P$. Diferente do método de penalidade, o método Lagrangeana aumentada atualiza dois parâmetros: a penalidade e os multiplicadores de Lagrange. O que o método faz realmente é aproximar os multiplicadores de Lagrange, λ e/ou μ , dos multiplicadores ótimos, λ^* e/ou μ^* . Assim não precisa que o parâmetro c tenda a infinito para que o método convirja para a solução ótima.

Explicaremos agora, o que ocorre entre o método de penalidade e Lagrangeana aumentada. Para isso, utilizaremos a função auxiliar de penalidade quadrática para o problema (3.1)

$$\mathcal{F}_P(x, c^k) = f(x^k) + \frac{c^k}{2} \sum_{j=1}^p h_j(x)^2. \quad (3.27)$$

Omitiremos a explicação para os problemas com restrições de desigualdade. Mas o que ocorre entre os dois problemas, tanto com os de restrições de igualdade quanto com os de desigualdade, é muito parecido. Derivando a função (3.27) em relação a x , obtemos

$$\nabla_x \mathcal{F}_P(x, c^k) = \nabla_x f(x) + c^k \sum_{j=1}^p h_j(x) \nabla_x h_j(x). \quad (3.28)$$

Comparando com a condição de estacionaridade

$$\nabla_x \mathcal{L}(x, \lambda^k) = \nabla_x f(x) - \sum_{j=1}^p \lambda_j \nabla_x h_j(x)$$

obtemos a seguinte aproximação

$$\lambda_j \approx -c^k h_j(x^{k+1}) \quad (3.29)$$

onde x^{k+1} é a solução inexata do problema de minimizar a função (3.27), ou seja, satisfaz apenas as condições de primeira ordem de KKT.

O teorema 17.2, página 497, de [NW99], nos diz que o limite do lado direito de (3.29) quando $k \rightarrow \infty$ é λ_j^* ,

$$\lim_{k \rightarrow \infty} -c^k h_j(x^{k+1}) = \lambda_j^*. \quad (3.30)$$

Observamos em (3.30) que a factibilidade somente é obtida quando $c^k \rightarrow \infty$. Já no método Lagrangeana aumentada temos de (3.22) a seguinte aproximação

$$h_j(x^{k+1}) = -\frac{1}{c^k}(\lambda_j^* - \lambda_j^k). \quad (3.31)$$

Note que se λ_j^k estiver muito próximo de λ_j^* , a infactibilidade de x^k é bem menor. Isso faz com que c não precise tender ao infinito para que esse método convirja para a solução ótima. Dessa maneira, a função Lagrangeana aumentada elimina o problema de mal condicionamento existente no método de penalidade.

Ilustraremos agora o comportamento da função $\mathcal{L}_{\mathcal{A}}(x, \lambda, c)$ quando $\lambda \rightarrow \lambda^*$ e $c \rightarrow \infty$ através do seguinte exemplo,

$$\begin{array}{ll} \text{Minimizar} & x_1^2 + x_2^2 \\ \text{sujeito a} & x_2 - 2 = 0 \end{array} \quad (3.32)$$

com solução ótima em $x^* = (0, 2)^T$ e $\lambda = 4$. A Lagrangeana aumentada é

$$\mathcal{L}_{\mathcal{A}}(x, \lambda) = x_1^2 + x_2^2 - \lambda(x_2 - 2) + \frac{c}{2}(x_2 - 2)^2. \quad (3.33)$$

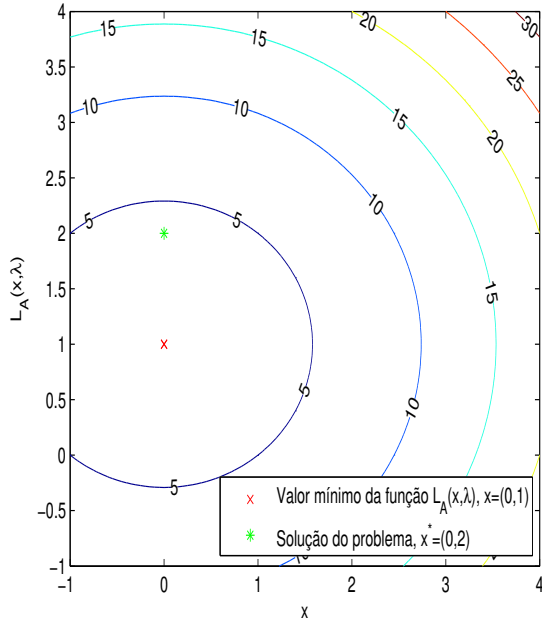
Utilizando a condição de estacionariedade obtemos

$$x_1(\lambda, c) = 0 \quad x_2(\lambda, c) = \frac{\lambda + 2c}{2 + c}.$$

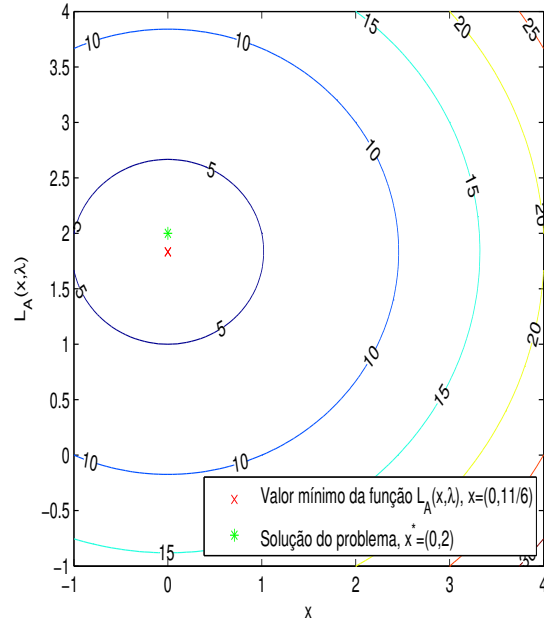
Para $c > 0$ temos

$$\lim_{\lambda \rightarrow \lambda^*} x_1(\lambda, c) = 0 \quad \text{e} \quad \lim_{\lambda \rightarrow \lambda^*} x_2(\lambda, c) = 2.$$

Na Figura 3.3 podemos ver que quando $\lambda \rightarrow \lambda^*$, x fica mais próximo da solução ótima, ou seja, $x \rightarrow x^*$.



(a) $\lambda = 1$ e $c = 1$



(b) $\lambda = \frac{7}{2}$ e $c = 1$

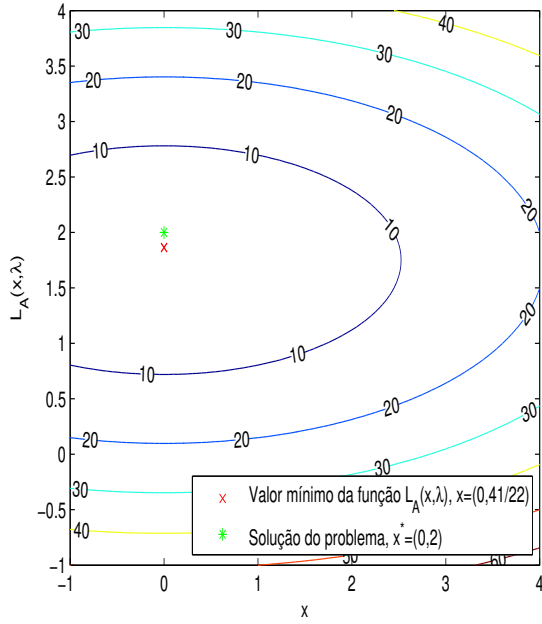
Figura 3.3: Comportamento da função (3.26) quando $\lambda \rightarrow \lambda^*$.

Agora vamos estabelecer λ constante e c variável,

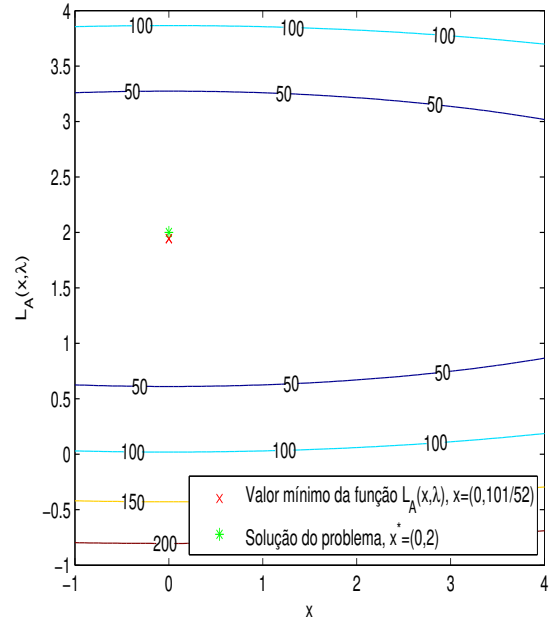
$$\lim_{c \rightarrow \infty} x_1(\lambda, c) = 0 \quad \text{e} \quad \lim_{c \rightarrow \infty} x_2(\lambda, c) = 2.$$

Como acontece também com o método da penalidade, a Figura 3.4 mostra que quando $c \rightarrow \infty$ temos que $x \rightarrow x^*$.

Para começar o algoritmo, alguma suposição para o parâmetro e os multiplicadores são necessárias. Geralmente, atribui-se o vetor nulo como vetor inicial para o multiplicador de Lagrange. O método é finalizado quando as condições de KKT do problema são satisfeitas. Apresentamos nos algoritmos 3.3.1 e 3.3.2, respectivamente, os métodos da Lagrangeana aumentada para os problemas com restrições de igualdade e desigualdade.



(a) $\lambda = 1$ e $c = 10$



(b) $\lambda = 1$ e $c = 50$

Figura 3.4: Comportamento da função (3.26) quando $c \rightarrow \infty$.

Algoritmo 3.3.1 (Lagrangeana Aumentada - Para o Problema (3.1))

Dados $x^0 \in \mathbb{R}^n$, $\lambda^0 \in \mathbb{R}^p$ e $c^0 > 0$

$k = 0$

Passo 1: $x^{k+1} \in \operatorname{argmin} \{\mathcal{L}_{\mathcal{A}}(x, \lambda^k)\}$

Passo 2: Atualizar λ_j , $i = 1, \dots, p$

$$\lambda_j^{k+1} = \lambda_j^k + c^k h_j(x^{k+1})$$

Passo 3: Atualizar c^{k+1}

Passo 4: Testar as condições de KKT.

Se satisfeitas, **pare**.

Senão, volte para o **Passo 1**

Algoritmo 3.3.2 (Lagrangeana Aumentada - Para o problema (3.10))

Dados $x^0 \in \mathbb{R}^n$, $\mu^0 \in \mathbb{R}^m$ e $c^0 > 0$

$k = 0$

Passo 1: $x^{k+1} \in \operatorname{argmin} \{\mathcal{L}_A(x, \mu^k)\}$

Passo 2: Atualizar μ_i , $i = 1, \dots, m$

Se $g_i(x^{k+1}) \geq -\mu_i^k/c$, $\mu_i^{k+1} = 0$

Senão, $\mu_i^{k+1} = \mu_i^k + c^k g_i(x^{k+1})$

Passo 3: Atualizar c^{k+1}

Passo 4: Testar as condições de KKT.

Se satisfeitas, **pare**.

Senão, volte para o **Passo 1**

Neste capítulo vimos a extensão do método de penalidade quadrática. No próximo, veremos a extensão do método da barreira logarítmica: método da barreira logarítmica modificada.

Capítulo 4

Barreira Modificada

O método da barreira modificada foi introduzido por Polyak [Pol92] em 1992. Este método é uma extensão do método de barreira clássica, mas diferentemente deste aceita pontos infactíveis. Os métodos barreiras modificadas combinam as melhores propriedades das funções lagrangeanas e barreiras clássicas, e eliminam as deficiências existentes dessas. A Hessiana da função barreira torna-se mal condicionada quando os pontos gerados se aproximam da solução ótima. Isto não acontece com a Hessiana da função barreira modificada. A convergência do método da barreira modificada é finita enquanto que o do método de barreira clássica é assintótica. O método da barreira modificada permite que a solução gerada possa estar na fronteira da região factível. Já o método da barreira clássica a solução pode estar próxima à fronteira mas nunca alcançá-la. Uma qualidade importante é a representação explícita dos multiplicadores de Lagrange. Para obter a solução, o parâmetro de barreira do método barreira modificada não precisa crescer indefinidamente, mas isso só é válido se os multiplicadores de Lagrange estiverem convergindo para a solução dual ótima do problema. Para iniciar o método da barreira modificado não precisamos de um ponto factível como no método da barreira clássica. O método de barreira modificada desenvolvido por Polyak trata apenas de restrições de desigualdades. Mas extensões desse método foram criadas podendo ser aplicados a problemas com restrições de igualdade e desigualdade, como é o caso do método primal-dual barreira modificada de Sousa [Sou06].

4.1 A Função Barreira Modificada

Considere o seguinte problema,

$$\begin{aligned} &\text{Minimizar} && f(x) \\ &\text{sujeito a} && g_i(x) \geq 0, \quad i = 1, \dots, m \end{aligned} \tag{4.1}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$ e $f, g_1, \dots, g_m \in C^2$.

As funções barreiras modificadas citadas por Polyak [Pol92], como o nome já diz, são modificações das funções de barreiras clássicas de Frisch e Carrol.

$$\text{Função de Frisch: } \varphi(x) = \begin{cases} f(x) - \frac{1}{c} \sum_{i=1}^m \ln g_i(x) & , \text{ se } x \in \Omega \\ \infty & , \text{ se } x \notin \Omega \end{cases}$$

$$\text{Função de Carrol: } \chi(x) = \begin{cases} f(x) - \frac{1}{c} \sum_{i=1}^m \frac{1}{g_i(x)} & , \text{ se } x \in \Omega \\ \infty & , \text{ se } x \notin \Omega \end{cases}$$

Essas funções conhecidas também como função logarítmica e inversa, respectivamente, são as mais famosas das funções barreiras. A desvantagem de usá-las é que essas, tal como suas derivadas, não estão definidas na solução ótima x^* do problema e estas funções crescem infinitamente quando $x \rightarrow x^*$. A solução para isso, proposta por Polyak foi relaxar o conjunto das restrições factíveis.

As funções barreiras modificadas definidas por Polyak associadas ao problema (4.1) são:

- **Função barreira logarítmica modificada de Polyak**

$$\mathcal{M}_{\mathcal{L}}(x, \mu) = \begin{cases} f(x) - \frac{1}{c} \sum_{i=1}^m \mu_i \ln [cg_i(x) + 1] & , \text{ se } x \in \Omega \\ \infty & , \text{ se } x \notin \Omega \end{cases} \tag{4.2}$$

- **Função barreira inversa modificada de Polyak**

$$\mathcal{M}_{\mathcal{I}}(x, \mu) = \begin{cases} f(x) - \frac{1}{c} \sum_{i=1}^m \frac{1}{cg_i(x) + 1} & , \text{ se } x \in \Omega \\ \infty & , \text{ se } x \notin \Omega \end{cases} \tag{4.3}$$

Antes de apresentarmos o desenvolvimento dessas funções vamos observar o exemplo abaixo retirado do artigo de Polyak [Pol92]. Nesse exemplo inserimos algumas restrições que limitam x_1 superiormente e inferiormente. Isso foi feito para limitarmos a região factível. Vamos considerar o seguinte problema:

$$\begin{aligned}
 &\text{Minimizar} && f(x) = x_1^2 - x_2^2 \\
 &\text{sujeito a} && g_1(x) = 2 - x_2 \geq 0 \\
 &&& g_2(x) = x_2 \geq 0 \\
 &&& g_3(x) = 4 - x_1 \geq 0 \\
 &&& g_4(x) = 2 + x_1 \geq 0.
 \end{aligned} \tag{4.4}$$

Nesse exemplo vamos considerar apenas o conjunto factível do problema (4.4). Definiremos o conjunto factível como $\Omega = \{x \in \mathbb{R}^n \mid 2 - x_2 \geq 0, x_2 \geq 0, 4 - x_1 \geq 0, 2 + x_1 \geq 0\}$. Expandindo por um fator c (neste exemplo vamos considerar $c = 1$) e deslocando por um fator unitário cada uma das restrições do problema (4.4), temos o seguinte conjunto $\Omega_c = \{x \in \mathbb{R}^n \mid c(2 - x_2) + 1 \geq 0, cx_2 + 1 \geq 0, c(4 - x_1) + 1 \geq 0, c(2 + x_1) + 1 \geq 0\}$. Na Figura 4.1 temos a representação gráfica dos conjuntos Ω e Ω_c .

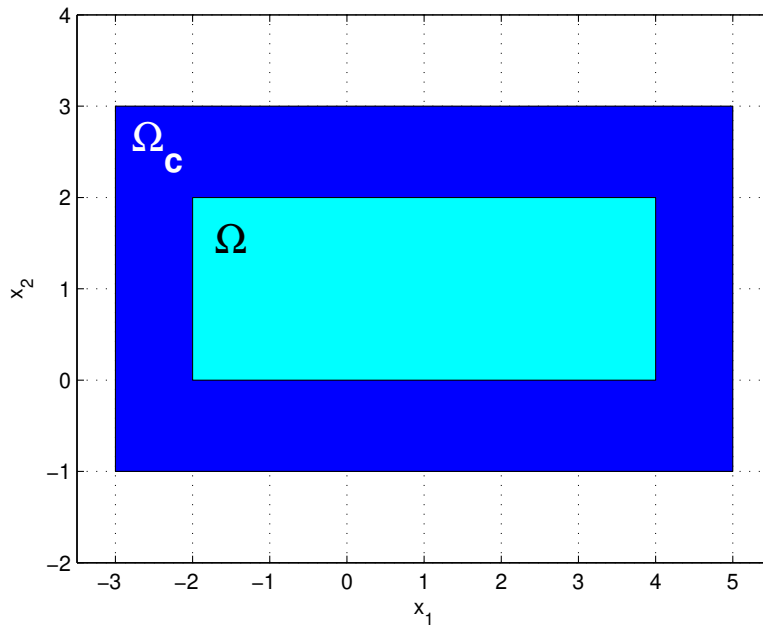


Figura 4.1: Representação gráfica dos conjuntos Ω e Ω_c do problema 4.4.

Note que Ω está contido em Ω_c . Independente da escolha das funções barreira modificada Polyak, seja a logarítmica ou a inversa, é no conjunto Ω_c que o método trabalha. Veja na Figura 4.1, que o método barreira modificada permite gerar pontos infactíveis em relação ao conjunto Ω . Eis aí o porquê de comparar com o método Lagrangeana aumentada. Observe também que quando aumentamos o parâmetro c expandimos ainda mais o conjunto Ω_c . A função barreira modificada é considerada uma Lagrangeana interna, pois mesmo que produza pontos infactíveis de Ω , o método não gera pontos que não pertençam a Ω_c .

Vamos agora conferir como chegamos na função barreira modificada. O desenvolvimento da função (4.3) é análoga à função (4.2).

Desenvolvimento da Função de barreira logarítmica modificada de Polyak

1. Multiplicamos por $c > 0$ as desigualdades do problema 4.1

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{sujeito a} & cg_i(x) \geq 0 \quad i = 1, \dots, m \end{array}$$

2. Adicionamos uma unidade as restrições

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{sujeito a} & cg_i(x) + 1 \geq 1, \quad i = 1, \dots, m \end{array}$$

3. Aplicamos a função logarítmica e dividimos as desigualdades por $c > 0$

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{sujeito a} & \frac{1}{c} \ln [cg_i(x) + 1] \geq 0, \quad i = 1, \dots, m \end{array} \tag{4.5}$$

No problema (4.5) a condição de existência para a função logarítmica é:

$$cg_i(x) + 1 \geq 0 \quad i = 1, \dots, m.$$

Isso implica

$$g_i(x) \geq -\frac{1}{c} \quad i = 1, \dots, m.$$

Assim as soluções do problema (4.5) se aproximam da solução de (4.1) quando c cresce indefinidamente.

4. Ao problema (4.5) associa-se uma função Lagrangeana

$$\mathcal{M}_{\mathcal{L}}(x, \mu, c) = f(x) - \frac{1}{c} \sum_{i=1}^m \mu_i \ln(cg_i(x) + 1)$$

Assim obtemos a função barreira logarítmica modificada (4.2). A obtenção da função barreira inversa modificada de Polyak obtém-se de forma análoga à da função (4.2).

Uma vantagem das funções barreiras modificadas é que elas e suas derivadas são definidas na solução ótima x^* . Já nas funções barreiras clássicas isso não ocorre. Lógico, se x^* for ponto interior da região factível, tanto as funções quanto suas derivadas são definidas neste ponto. Voltando ao exemplo (4.4), os valores da função da barreira clássica de Frisch deste problema e sua derivada são indefinidos,

$$\mathcal{F}_{\mathcal{BF}}(x^*) = x_1^{*2} - x_2^{*2} - \frac{1}{c}(\ln(2 - x_2^*) + \ln(x_2^*) + \ln(4 - x_1^*) + \ln(2 + x_1^*)) = -4 - \frac{1}{c}(\ln 0 + 4 \ln 2)$$

$$\nabla \mathcal{F}_{\mathcal{BF}}(x^*) = \begin{pmatrix} 2x_1^* - \frac{1}{c} \left(-\frac{1}{4 - x_1^*} + \frac{1}{2 + x_1^*} \right) \\ -2x_2^* - \frac{1}{c} \left(-\frac{1}{2 - x_2^*} + \frac{1}{x_2^*} \right) \end{pmatrix} = \begin{pmatrix} -\frac{1}{4c} \\ -4 - \frac{1}{c} \left(-\frac{1}{0} + \frac{1}{2} \right) \end{pmatrix}.$$

Além das funções barreira clássicas e suas derivadas serem indefinidas na solução do problema, $\mathcal{F}_{\mathcal{BF}}$ cresce infinitamente quando $x \rightarrow x^*$.

Agora vamos calcular o valor da função barreira logarítmica modificada de Polyak nas respectivas soluções ótimas primais e duais, $x^* = (0, 2)^T$ e $\mu^* = (4, 0, 0, 0)^T$ para o problema (4.4),

$$\mathcal{M}_{\mathcal{L}}(x^*, \mu^*) = x_1^{*2} - x_2^{*2} - \frac{1}{c} \left\{ \mu_1^* \ln [c(2 - x_2^*) + 1] + \mu_2^* \ln [cx_2^* + 1] + \mu_3^* \ln [c(4 - x_1^*) + 1] + \mu_4^* \ln [c(2 + x_1^*) + 1] \right\} = -4$$

$$\nabla \mathcal{M}_{\mathcal{L}}(x^*, \mu^*) = \begin{pmatrix} 2x_1^* + \frac{\mu_3^*}{c(4 - x_1^*) + 1} - \frac{\mu_4^*}{c(2 + x_1^*) + 1} \\ -2x_2^* + \frac{\mu_1^*}{c(2 - x_2^*) + 1} - \frac{\mu_2^*}{cx_2^* + 1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Note que os valores obtidos acima de $\mathcal{M}_{\mathcal{L}}$ e $\nabla \mathcal{M}_{\mathcal{L}}$ coincidem com os valores da função

objetivo e com o gradiente da Lagrangeana do problema (4.4) nos pontos $x^* = (0, 2)^T$ e $\mu^* = (4, 0, 0, 0)^T$, dado qualquer c ,

$$f(x^*) = -4$$

$$\nabla \mathcal{F}(x^*, \mu^*) = \begin{pmatrix} 2x_1^* + \mu_3^* - \mu_4^* \\ -2x_2^* + \mu_1^* - \mu_2^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

onde $\mathcal{L}(x, \mu) = x_1^2 - x_2^2 - \mu_1(2 - x_2) - \mu_2(4 - x_1) - \mu_3(4 - x_1) - \mu_4(2 + x_1)$.

Isso que acabamos de ver faz parte das propriedades da função barreira modificada. Sejam x^* e μ^* , respectivamente, a solução primal e dual do problema (4.1), as funções barreira modificadas têm as seguintes propriedades para $c > 0$,

P1. $\mathcal{M}_{\mathcal{L}}(x^*, \mu^*) = \mathcal{M}_{\mathcal{I}}(x^*, \mu^*, c) = f(x^*)$;

P2. $\nabla_x \mathcal{M}_{\mathcal{L}}(x^*, \mu^*) = \nabla_x \mathcal{M}_{\mathcal{I}}(x^*, \mu^*) = \nabla f(x) - \sum_i^m \mu_i \nabla g_i(x) = 0$.

Se (4.1) for um problema convexo e se μ^* for conhecido, então para qualquer $c > 0$ podemos resolvê-lo minimizando $\mathcal{M}_{\mathcal{L}}$ ou $\mathcal{M}_{\mathcal{I}}$ através de algum método para programação irrestrita.

P3. $x^* = \operatorname{argmin} \mathcal{M}_{\mathcal{L}}(x, \mu^*) = \operatorname{argmin} \mathcal{M}_{\mathcal{I}}(x, \mu^*)$.

Vamos agora analisar o comportamento da função barreira modificada para o problema (2.8), página 15. A função barreira logarítmica modificada de Polyak para este problema é

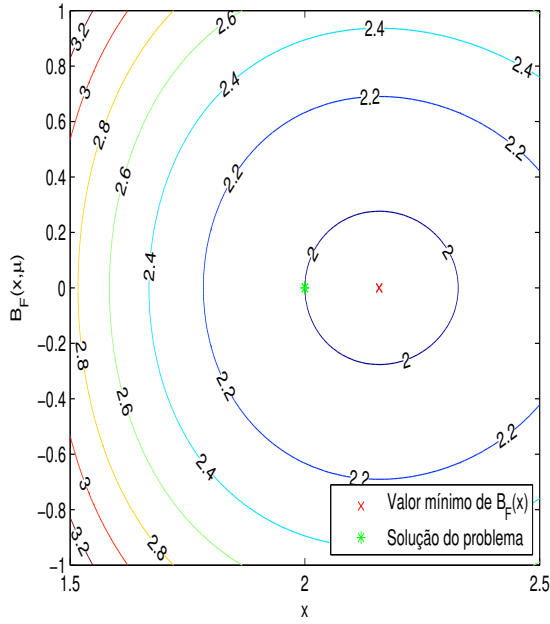
$$\mathcal{M}_{\mathcal{L}}(x, \mu) = \frac{1}{2}(x_1^2 + x_2^2) - \frac{1}{c} \ln [c(x_1 - 2) + 1]. \quad (4.6)$$

Na Figura 4.2 temos o gráfico da função $\mathcal{M}_{\mathcal{L}}(x, \mu)$ para valores diferentes de c e μ .

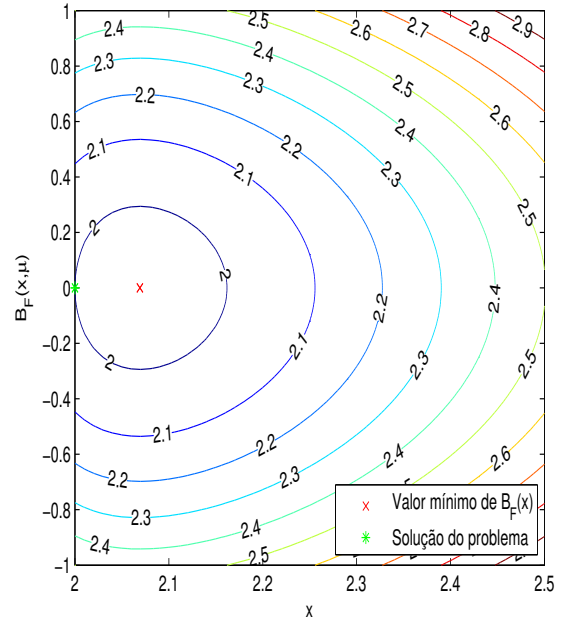
Observando essa figura podemos notar que (4.6) é definida no ponto ótimo $x^* = (2, 0)$, ao contrário da função barreira clássica

$$\mathcal{F}_{\mathcal{B}}(x) = \frac{1}{2}(x_1^2 + x_2^2) - \frac{1}{c} \ln(x_1 - 2).$$

onde não é definida em $x_1 = 2$.



(a) $\mu = \frac{5}{2}$ e $c = 1$



(b) $\mu = \frac{7}{2}$ e $c = 10$

Figura 4.2: Comportamento da função (4.6).

4.2 Método da Barreira Modificada

Polyak apresenta em seu artigo [Pol92] mais de uma versão para o método da barreira modificada. Ele apresenta versões tanto para problemas de programação convexa quanto para problemas não-convexos. A versão que mantém o parâmetro de barreira c constante tem um raio de convergência linear. Já aumentando o parâmetro passo a passo obtém-se raio de convergência superlinear. Os métodos apresentados por ele são: PPV (*Permanent Parameter Version*), APV (*Adjusted Parameter Version*) e VPV (*Varying Parameter Version*).

Os dois primeiros métodos, PPV e APV, são para problemas de programação convexa. O primeiro, como o nome já diz, mantém o parâmetro de barreira constante e no método APV c aumenta a cada passo. O método VPV é uma versão do APV para problemas não-convexos. Esses métodos diferem em seus ciclos externos, pois o interno todos possuem iguais, ou seja, a solução primal de uma determinada iteração é obtida resolvendo

o seguinte subproblema:

$$\begin{aligned} &\text{Minimizar } \mathcal{M}_{\mathcal{L}}(x, \mu) \\ &x \in \mathbb{R}^n. \end{aligned} \tag{4.7}$$

A seguir apresentamos a versão mais simples de Polyak, o algoritmo PPV. As outras versões podem ser encontradas em [Pol92].

Algoritmo 4.2.1 (PPV - Para problemas de programação convexa)

Dados $x^0 \in \text{int } \Omega_c$, $\mu^0 = e = (1, \dots, 1)^T \in \mathbb{R}^m$ e $c \geq \bar{c} > 0$

$k = 0$

Passo 1: $x^{k+1} = \text{argmin } \{\mathcal{M}_{\mathcal{L}}(x, \mu^k) : x \in \mathbb{R}^n\}$

Passo 2: Atualizar μ_i , $i = 1, \dots, m$

Passo 3: Testar as condições de KKT.

Se satisfeitas, **pare**.

Senão, volte para o **Passo 1**

A forma como os métodos PPV, APV e VPV são apresentados não é computacionalmente viável. Uma maneira de torná-la viável é substituir o procedimento infinito $x^{k+1} = \text{argmin } \{\mathcal{M}_{\mathcal{L}}(x, \mu^k) : x \in \mathbb{R}^n\}$ por um procedimento finito. Para isso, Polyak utilizou o método de Newton.

Como visto no método da Lagrangeana aumentada, o método da barreira modificada também se constitui em dois passos principais:

1. $\text{argmin } \{\mathcal{M}_{\mathcal{L}}(x, \mu^k) : x \in \mathbb{R}^n\}$;
2. Atualizar os multiplicadores de Lagrange μ_i , $i = 1, \dots, m$.

A minimização de $\mathcal{M}_{\mathcal{L}}(x, \mu)$ ocorre no espaço primal em relação à variável x enquanto que as atualizações dos μ_i 's ocorrem no espaço dual.

Para obter a regra de atualização dos multiplicadores de Lagrange, voltemos ao subproblema do método da barreira modificada $\text{argmin } \{\mathcal{M}_{\mathcal{L}}(x, \mu^k) : x \in \mathbb{R}^n\}$. Se x^{k+1} é solução do subproblema para um dado μ^k e c^k então

$$\nabla_x \mathcal{M}_{\mathcal{L}}(x^{k+1}, \mu^k, c^k) = \nabla f(x^{k+1}) - \sum_{i=1}^m \frac{\mu_i^k}{c^k g(x^{k+1}) + 1} \nabla g_i(x^{k+1}) = 0. \tag{4.8}$$

Comparando (4.8) com o gradiente função Lagrange do problema (4.1)

$$\nabla \mathcal{L}(x, \mu) = \nabla f(x) - \sum_{i=1}^m \mu_i \nabla g_i(x)$$

obtemos a atualização dos multiplicadores

$$\mu_i^{k+1} = \frac{\mu_i^k}{c^k g_i(x^{k+1}) + 1}, \quad i = 1, \dots, m. \quad (4.9)$$

A atualização do parâmetro de penalidade apresentada com maior frequência em trabalhos que utilizam o método de barreira modificada [BS06, CV03, VF97, VB98] é a regra que multiplica c por um fator positivo,

$$c^{k+1} = \beta c^k, \quad \beta > 1.$$

Para problemas de programação convexa foi provado por Jensen e Polyak [JP94] que as soluções geradas pelo método barreira modificada converge para a solução ótima para qualquer parâmetro de penalidade positivo fixo, ou seja,

$$c^{k+1} = c^0, \quad k = 0, 1, \dots$$

Mesmo que o método da barreira modificada permita pontos inactivéis, os pontos gerados deverão pertencer ao conjunto Ω_c . No algoritmo (4.2.2) há duas condições para controlar as soluções geradas. A primeira é verificar se x^k é factível em relação ao conjunto Ω_c , ou seja,

$$x \in \Omega_c \quad (4.10)$$

e o segundo critério é do decréscimo suficiente. Existem vários critérios como a condição de Wolfe, a condição de Armijo (também denominada de condição de curvatura) e condição de Goldstein [NW99]. No artigo de Polyak [Pol92] utilizou-se a seguinte condição Goldstein-Armijo [DS93],

$$\mathcal{M}_{\mathcal{L}}(x + \alpha \Delta x, \mu) \leq \mathcal{M}_{\mathcal{L}}(x, \mu) + \rho \alpha (\nabla_x \mathcal{M}_{\mathcal{L}}(x, \mu), \Delta x). \quad (4.11)$$

onde α é o tamanho do passo e ρ é uma constante pertencente ao intervalo $(0, 1)$. Em

[Pol92] se utiliza $\rho = \frac{1}{3}$.

Algoritmo 4.2.2 (Método da Barreira Logarítmica Modificada para problemas não-convexos)

Dados $x^0 \in \text{int } \Omega$, $\mu^0 = e = (1, \dots, 1)^T \in \mathbb{R}^m$ e $c^0 \geq \bar{c} > 0$
 $k = 0$

Passo 1: $x = x^k$, $\mu = \mu_k$, $c = c^k$

Passo 2: $x^{k+1} = \text{argmin } \{\mathcal{F}(x, \mu^k) : x \in \mathbb{R}^n\}$
 $\alpha = 1$

Passo 3: Verificar $x^k + \alpha \Delta x \in \Omega_c$ e condição de Armijo (4.11)

Passo 4: Se $x + \alpha \Delta x \in \Omega_c$ e Armijo satisfeito e $\alpha = 1$

$x^{k+1} = x^k + \Delta x$ e vá para o **Passo 5**

Se $x + \alpha \Delta x \in \Omega_c$ e Armijo satisfeito e $\alpha < 1$

$x^{k+1} = x^k + \Delta x$ e vá para o **Passo 2**

Passo 5: Atualizar μ_{k+1} , conformae (4.8)

Atualizar c_{k+1}

Passo 6: Testar as condições de KKT.

Se satisfeitas, **pare**.

Senão, $k = k + 1$ e volte para o **Passo 1**

O algoritmo 4.2.2 vale para problemas convexos e não-convexos.

O método apresentado neste capítulo e do anterior (métodos barreira logarítmica modificada e Lagrangeana aumentada) foram implementados, testados e comparados. Os resultados obtidos veremos no capítulo a seguir.

Capítulo 5

Resultados

Diferentemente do paper de Nash, Polyak e Sofer [NPS94], apresentaremos neste capítulo os resultados dos métodos Lagrangeana aumentada e barreira modificada para alguns problemas de programação não-linear. Em [NPS94] vemos a comparação do método barreira estabilizada¹ com o da barreira modificada. Esse primeiro método citado é o método da barreira clássica, porém, com alguns recursos adicionais que eliminem o mal condicionamento da Hessiana da função $\mathcal{F}_B(c)$ quando o parâmetro tende ao infinito. A atualização do parâmetro de penalidade utilizada em [NPS94] é a regra simples $c^{k+1} = \beta c^k$. Para cada método foi fixado um valor diferente de β . Dos trinta e três problemas testados apenas nove tiveram um desempenho melhor com o método da barreira estabilizada. O desempenho dos métodos nesse paper caracteriza-se pelo número de iterações e avaliações das funções auxiliares e suas derivadas.

Em nosso trabalho utilizamos a mesma regra de atualização do parâmetro citada anteriormente. Mas diferente de [NPS94] fixamos valores iguais de β para os dois métodos. O que queremos mostrar aqui, é a seqüência gerada pelos métodos e analisar qual dos métodos alcançou a melhor solução tanto no primal quanto no dual.

Os testes foram realizados com problemas no $\mathbb{R}^2$. Para que possamos mostrar o comportamento dos pontos gerados pelos métodos achamos adequado implementarmos em problemas pequenos. Como não queremos saber somente o número de iterações e avaliações do gradiente de $\mathcal{F}_B(c)$, nada melhor que um problema de pequeno porte para sabermos o quão precisa é a solução gerada.

¹Esse método utiliza a técnica de substituir a direção de Newton por uma aproximação numericamente estável no ciclo interno quando o parâmetro c está muito grande [NPS94, NS93]

Iniciaremos este capítulo apresentando os problemas testes e suas representações gráficas para depois mostrarmos os resultados obtidos.

5.1 Problemas Testes

Apresentaremos nesta seção os cinco problemas de programação não-linear definidos no $\mathbb{R}^2$ onde foram aplicados os métodos da Lagrangeana aumentada e da barreira modificada. Esses problemas foram retirados do livro [Sch87]. Eis aqui a expressão matemática desses problemas bem como sua representação gráfica no $\mathbb{R}^2$.

Problema 1:

$$\begin{aligned} \text{Minimizar} \quad & f(x) = (x_1 - 2)^2 + (x_2 - 1)^2 \\ \text{sujeito a} \quad & g_1(x) = -x_1^2 + x_2 \geq 0 \\ & g_2(x) = x_1 - x_2^2 \geq 0 \end{aligned}$$

$$\text{Solução: } x^* = (1, 1)^T, \mu^* = (\frac{4}{3}, \frac{4}{3})^T, f(x^*) = 1$$

Problema 2:

$$\begin{aligned} \text{Minimizar} \quad & f(x) = x_2 \\ \text{sujeito a} \quad & g_1(x) = x_2 - x_1^2 \geq 0 \\ & g_2(x) = x_2 \geq 0 \end{aligned}$$

$$\text{Solução: } x^* = (0, 0)^T, \mu^* = (\frac{1}{2}, \frac{1}{2})^T, f(x^*) = 0$$

Problema 3:

$$\begin{aligned} \text{Minimizar} \quad & f(x) = x_2 \\ \text{sujeito a} \quad & g_1(x) = -2x_1^2 + x_1^3 + x_2 \geq 0 \\ & g_2(x) = -2(1 - x_1)^2 + (1 - x_1)^3 + x_2 \geq 0 \end{aligned}$$

$$\text{Solução: } x^* = (\frac{1}{2}, \frac{3}{8})^T, \mu^* = (\frac{1}{2}, \frac{1}{2})^T, f(x^*) = \frac{3}{8}$$

Problema 4:

$$\begin{aligned}
&\text{Minimizar} && f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \\
&\text{sujeito a} && g_1(x) = \frac{1}{3}x_1 + x_2 + 0.1 \geq 0 \\
&&& g_2(x) = -\frac{1}{3}x_1 + x_2 + 0.1 \geq 0
\end{aligned}$$

Solução: $x^* = (1, 1)^T, \mu^* = (0, 0)^T, f(x^*) = 0$

Problema 5

$$\begin{aligned}
&\text{Minimizar} && f(x) = x_2 \\
&\text{sujeito a} && g_1(x) = x_2 - x_1^2 \geq 0 \\
&&& g_2(x) = x_1 \geq 0
\end{aligned}$$

Solução: $x^* = (0, 0)^T, \mu^* = (1, 0)^T, f(x^*) = 0$

A seguir temos os gráficos dos problemas apresentados anteriormente com suas respectivas soluções primais.

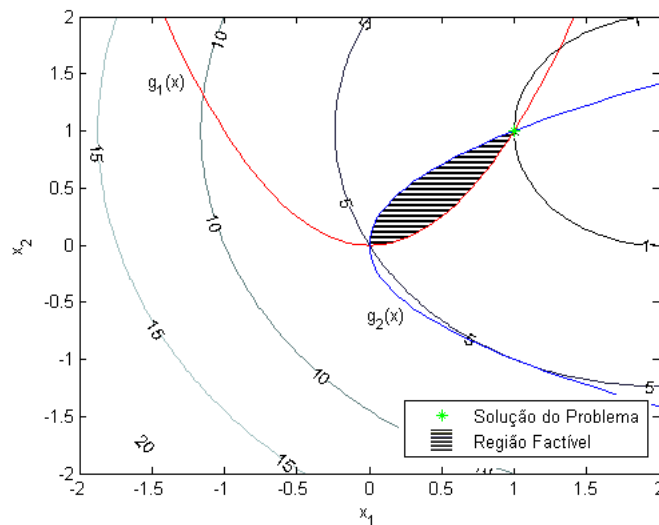


Figura 5.1: Representação Gráfica do Problema 1.

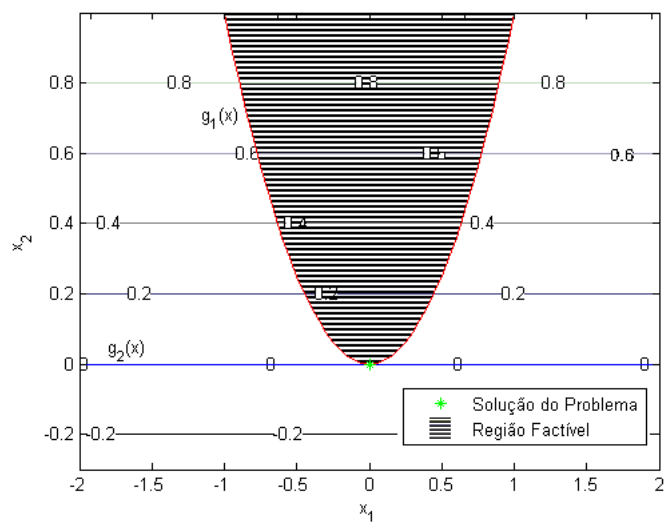


Figura 5.2: Representação Gráfica do Problema 2.

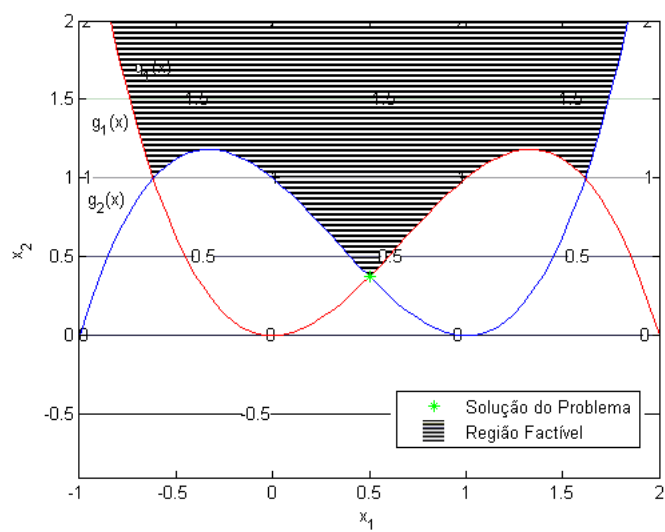


Figura 5.3: Representação Gráfica do Problema 3.

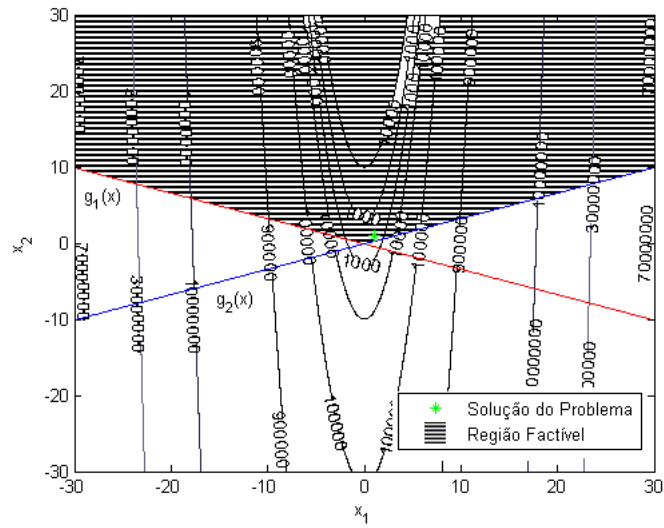


Figura 5.4: Representação Gráfica do Problema 4.

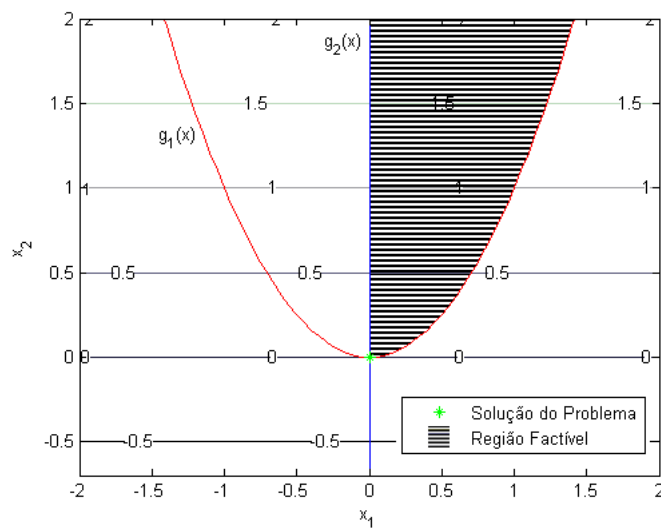


Figura 5.5: Representação Gráfica do Problema 5.

5.2 Implementação

Os métodos da Lagrangeana aumentada e da barreira modificada foram implementados no Matlab 7.2.0.232 (R2006a). Os algoritmos utilizados na implementação desses métodos são os apresentados na Seções 3.3 e 4.2. Nas Figuras 5.6 e 5.7 temos os fluxogramas dos métodos implementados.

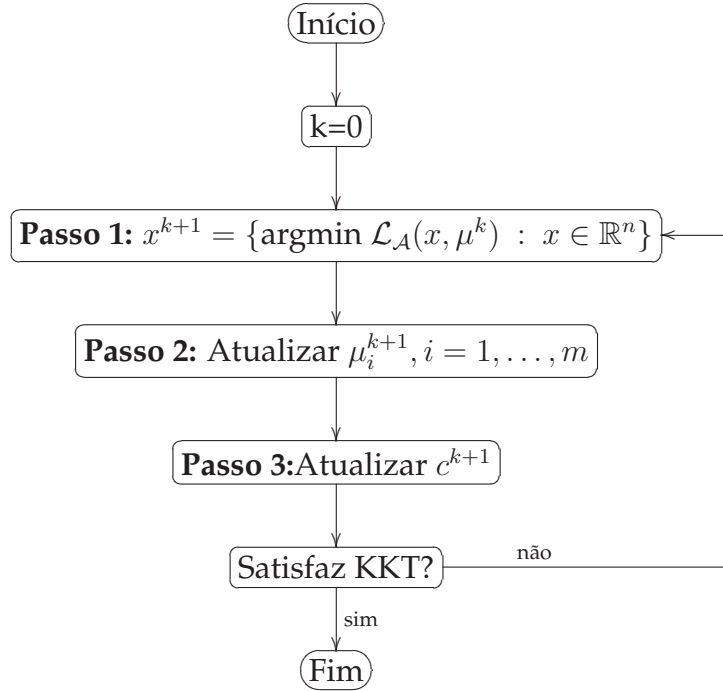


Figura 5.6: Fluxograma do método da Lagrangeana aumentada.

Em ambos os métodos utilizamos como critério de parada as condições de KKT (ver Apêndice A). Determinamos em nossa implementação como sendo zero o valor $tol = 1e - 4$. Assim as condições de KKT serão satisfeitas quando

$$\begin{aligned}
 &\nabla \mathcal{L}(x, \mu) < tol \\
 &g_i(x) > tol, \quad i = 1, \dots, m \\
 &\mu_i > tol, \quad i = 1, \dots, m \\
 &g_i(x)\mu_i < tol, \quad i = 1, \dots, m.
 \end{aligned}$$

Foram realizados testes onde o parâmetro c variou com o decorrer das iterações e para um c fixo. Os valores de c e β foram iguais para os dois métodos. Os valores estabelecidos,

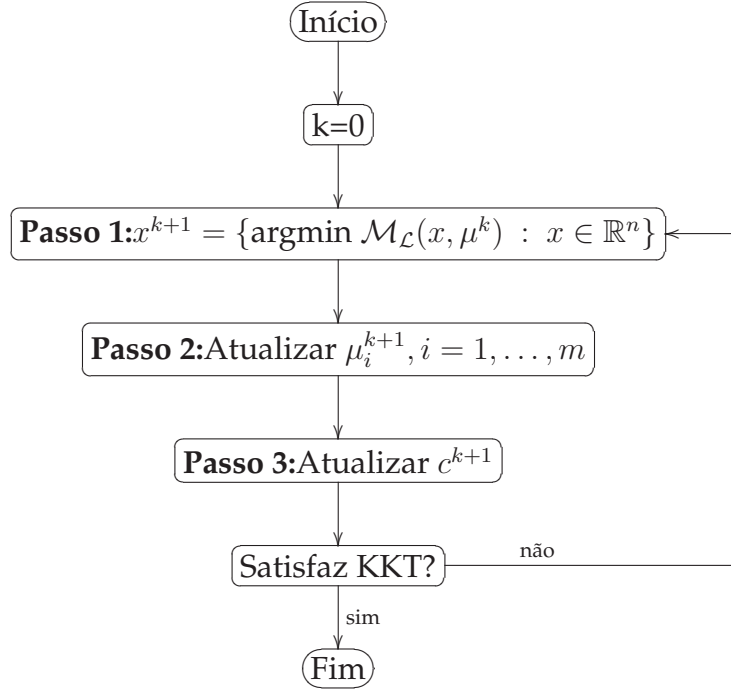


Figura 5.7: Fluxograma do método da barreira modificada.

respectivamente são:

- para c não fixo: $c^0 = 1e - 3$ e $\beta = 3$;
- para c fixo: $c = 1$ e $\beta = 1$.

Para encontrar a solução primal de ambos os métodos utilizamos o método de Newton (ver Apêndice B). A atualização das variáveis duais no método da Lagrangeana aumentada é dada pela fórmula (3.25), página 28. Já para o método da barreira modificada a atualização é dada por (4.9), página 42. O parâmetro de penalidade é atualizado pela regra,

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

onde $c > 0$ e $\beta > 1$.

Para iniciar esses métodos devemos dar estimativas das soluções primais e duais. Tanto para o método da Lagrangeana aumentada (LA) quanto para o barreira da modificada (BM) não há restrições em que região iniciar. Neste caso, podemos iniciar tanto com pontos factíveis quanto infactíveis. As aproximações iniciais das variáveis duais são

$\mu_i = 0, \quad i = 1, \dots, m$ para o método Lagrangeana aumentada e $\mu_i = 1, \quad i = 1, \dots, m$ para a barreira modificada. Em alguns casos quando a solução do problema não converge com o método da Lagrangeana aumentada iniciamos com $\mu_i = 1, \quad i = 1, \dots, m$.

5.3 Resultados dos testes

Apresentaremos agora os resultados obtidos dos métodos da Lagrangeana aumentada e da barreira modificada. Os testes foram realizados levando-se em conta quatro abordagens:

- 1ª abordagem:** variando c e soluções iniciais x^0 factíveis;
- 2ª abordagem:** variando c e soluções iniciais x^0 infactíveis;
- 3ª abordagem:** valor fixo de c e soluções iniciais x^0 factíveis;
- 4ª abordagem:** valor fixo c e soluções iniciais x^0 infactíveis.

No Apêndice C as figuras C.1, C.2, C.3, C.4 e C.5 mostram as seqüências geradas pelos métodos utilizando as abordagens mencionadas anteriormente. Na primeira e na segunda abordagens utilizamos $c^0 = 1e - 3$ e $\beta = 3$.

As soluções obtidas para cada método podemos conferir nas tabelas 5.1 e 5.2. Na Tabela 5.3 temos o número de iterações interna e externa para cada método. A iteração externa compreende os passos 1, 2 e 3 dos algoritmos 5.6 e 5.7. O número de iteração interna está relacionada com o subproblema do passo 1. Na Tabela 5.4 apresentamos os erros das soluções geradas.

Os erros que utilizamos nas análises foram o relativo e o absoluto. Em alguns problemas onde a solução ótima real primal ou dual foi em $(0, 0)^T$ utilizamos o erro absoluto. Em (5.1) e (5.2) podemos ver os erros, respectivamente, absoluto e relativo de uma solução obtida $\bar{x}$,

$$\frac{\|\bar{x} - x^*\|}{\|x^*\|} \quad (5.1)$$

$$\|\bar{x} - x^*\|. \quad (5.2)$$

Tabela 5.1: Soluções obtidas pelo método Lagrangeana aumentada.

Problema 1		
Abordagens	x_{LA}	μ_{LA}
1 ^a	(1.0000000e+000,1.0000000e+000)	(1.3333333e+000,6.6666672e-001)
2 ^a	(1.0000000e+000,1.0000000e+000)	(1.3333333e+000,6.6666672e-001)
3 ^a	(1.0000509e+000,1.0000443e+000)	(1.3331388e+000,6.6651420e-001)
4 ^a	(1.9999707e+000 9.9997972e-001)	(2.2516063e-005 3.1536572e-005)
Problema 2		
Abordagens	x_{LA}	μ_{LA}
1 ^a	-	-
2 ^a	(8.3276870e-010,-2.9811027e-005)	(5.0000009e-001,5.0000009e-001)
3 ^a	-	-
4 ^a	(1.7958553e-007,5.1595836e-007)	(5.0000003e-001,5.0000003e-001)
Problema 3		
Abordagens	x_{LA}	μ_{LA}
1 ^a	-	-
2 ^a	(4.9999982e-001,3.7497323e-001)	(5.0000009e-001,5.0000009e-001)
3 ^a	-	-
4 ^a	(4.9999995e-001,3.7500008e-001)	(4.9999996e-001,4.9999998e-001)
Problema 4		
Abordagens	x_{LA}	μ_{LA}
1 ^a	(1.0000000e+000,9.9999999e-001)	(0.0000000e+000,0.0000000e+000)
2 ^a	(1.0000000e+000,1.0000000e+000)	(0.0000000e+000,0.0000000e+000)
3 ^a	(1.0000000e+000,9.9999999e-001)	(0.0000000e+000,0.0000000e+000)
4 ^a	(9.9959999e-001,9.9999999e-001)	(0.0000000e+000,0.0000000e+000)
Problema 5		
Abordagens	x_{LA}	μ_{LA}
1 ^a	(1.1584773e-006 -1.0550885e-008)	(1.0000019e+000 2.3213686e-006)
2 ^a	(1.9722690e-007 -6.3472433e-005)	(1.0000002e+000 0.0000000e+000)
3 ^a	(6.6824939e-005,5.8610640e-009)	(1.0000000e+000,1.3364987e-004)
4 ^a	(1.8495844e-007,-3.8785623e-007)	(1.0000001e+000,0.0000000e+000)

x_{LA} e μ_{LA} soluções geradas pelo método da Lagrangeana aumentada

- : método não convergiu

Tabela 5.2: Soluções obtidas pelo método barreira modificada.

Problema 1		
1 ^a	(9.9995463e-001,9.9994383e-001)	(1.3335371e+000,6.6686229e-001)
2	(9.9995463e-001,9.9994383e-001)	(1.3335371e+000,6.6686229e-001)
3	(9.9996207e-001, 9.9995319e-001)	(1.3335034e+000,6.6682970e-001)
4	(1.0000509e+000 1.0000443e+000)	(1.3331388e+000 6.6651420e-001)
Problema 2		
1 ^a	(3.0788989e-014,7.3044462e-008)	(5.0000000e-001,5.0000000e-001)
2 ^a	(1.1252921e-018,8.7536348e-012)	(5.0000000e-001,5.0000000e-001)
3 ^a	(-2.3743763e-011,-5.5577520e-005)	(5.0000000e-001,5.0000000e-001)
4 ^a	-	-
Problema 3		
1 ^a	(5.0000025e-001,3.7500002e-001)	(5.0000055e-001,4.9999945e-001)
2 ^a	(5.0000105e-001,3.7508824e-001)	(5.0000001e-001,5.0000000e-001)
3 ^a	(5.0000178e-001,3.7500714e-001)	(5.0000506e-001,4.9999503e-001)
4 ^a	-	-
Problema 4		
1 ^a	(1.0000069e+000,1.0000138e+000)	(4.7088482e-007,7.5873872e-006)
2 ^a	(1.0000346e+000,1.0000694e+000)	(2.6648571e-006,3.7768799e-005)
3 ^a	(1.0001653e+000,1.0003316e+000)	(1.2997027e-006,1.0971731e-004)
4 ^a	(1.0001653e+000,1.0003316e+000)	(1.2997013e-006,1.0971718e-004)
Problema 5		
1 ^a	(3.3179145e-003,1.0991941e-005)	(9.9999914e-001,6.6341886e-003)
2 ^a	(3.3179145e-003,1.0991941e-005)	(9.9999914e-001,6.6341886e-003)
3 ^a	-	-
4 ^a	-	-

x_{BM} e μ_{BM} soluções geradas pelo método da barreira modificada

- : método não convergiu

Tabela 5.3: Número de iterações interna e externa.

Método Lagrangeana Aumentada			Método Barreira Modificada		
Abordagem 1	$inter_e$	$inter_i$	Abordagem 1	$inter_e$	$inter_i$
Problema 1	14	(3)3,4,5,6,6,4,5,5,3,3,2,2	Problema 1	11	3,(3)2,3,3,(3)4,3,3
Problema 2	-	-	Problema 2	3	3,3,2
Problema 3	-	-	Problema 3	7	3,2,(3)1,2,2
Problema 4	1	9	Problema 4	11	(4)21,8,7,7,5,4,3,2
Problema 5	12	3,(4)2,3,3,4,4,3,3,2	Problema 5	13	3,(4)2,3,3,4,5,6,(3)5
Abordagem 2	$inter_e$	$inter_i$	Abordagem 2	$inter_e$	$inter_i$
Problema 1	14	4,3,3,4,5,6,6,4,5,5,3,3,2,2	Problema 1	11	3,(3)2,3,3,(3)4,3,3
Problema 2	2	5,2	Problema 2	7	9,3,(3)1,2,2
Problema 3	2	7,2	Problema 3	2	3,2
Problema 4	1	3	Problema 4	11	(8)21,5,4,3
Problema 5	2	6,2	Problema 5	13	3,(4)2,3,3,4,5,6,(3)5
Abordagem 3	$inter_e$	$inter_i$	Abordagem 3	$inter_e$	$inter_i$
Problema 1	36	7,6,5,6,(5)7,(4)6,(8)5,(4)4,(7)3,2,3,2,3	Problema 1	30	4,(3)3,(24)2,1,2
Problema 2	-	-	Problema 2	2	20,3
Problema 3	-	-	Problema 3	2	7,2
Problema 4	1	9	Problema 4	14	10,8,(3)4,(5)3,(3)2,1
Problema 5	22	4,(6)3,(15)2	Problema 5	-	-
Abordagem 4	$inter_e$	$inter_i$	Abordagem 4	$inter_e$	$inter_i$
Problema 1	36	9,6,5,6,(5)7,(4)6,(8)5,(4)4,(7)3,2,3,2,3,	Problema 1	-	-
Problema 2	2	8,2	Problema 2	-	-
Problema 3	2	3,2	Problema 3	-	-
Problema 4	1	4	Problema 4	14	9,8,(3)4,(5)3,(3)2,1
Problema 5	2	4,2	Problema 5	-	-

$inter_e$: Número de iterações externas

$inter_i$: Sequência das sucessivas iterações internas em cada iteração externa

(a)b: essa notação significa a sucessivas iterações externas que foram obtidas b como iteração interna

Tabela 5.4: Erros absolutos das soluções primais e duais.

Método Lagrangeana Aumentada			Método Barreira Modificada		
Abordagem 1	$Erro_1$	$Erro_2$	Abordagem 1	$Erro_1$	$Erro_2$
Problema 1	0.000000000000	4.219005e-008	Problema 1	5.105637e-005	1.894867e-004
Problema 2	-	-	Problema 2	7.346934e-008*	0.000000000000
Problema 3	-	-	Problema 3	4.012779e-007	1.099999e-006
Problema 4	7.071068e-009	0.000000000000*	Problema 4	1.090986e-005	7.601985e-006*
Problema 5	1.158525e-006*	2.999792e-006	Problema 5	0.00331793271*	0.00663418865
Abordagem 2	$Erro_1$	$Erro_2$	Abordagem 2	$Erro_1$	$Erro_2$
Problema 1	0.000000000000	4.219005e-008	Problema 1	5.105637e-005	1.894867e-004
Problema 2	2.981102e-005*	1.800000e-007	Problema 2	8.753635e-012*	0.000000000000
Problema 3	4.283296e-005	1.800000e-007	Problema 3	1.411939e-004	1.414213e-008
Problema 4	0.000000000000	0.000000000000*	Problema 4	5.483393e-005	3.786269e-005*
Problema 5	6.347274e-005*	1.999999e-007	Problema 5	0.00331793271*	0.00663418865
Abordagem 3	$Erro_1$	$Erro_2$	Abordagem 3	$Erro_1$	$Erro_2$
Problema 1	4.771425e-005	1.658016e-004	Problema 1	4.260200e-005	1.580384e-004
Problema 2	-	-	Problema 2	5.557752e-005*	0.000000000000
Problema 3	-	-	Problema 3	1.177365e-005	1.003040e-005
Problema 4	7.071068e-009	0.000000000000*	Problema 4	2.619949e-004	1.097250e-004*
Problema 5	6.682494e-005*	1.336499e-004	Problema 5	-	-
Abordagem 4	$Erro_1$	$Erro_2$	Abordagem 4	$Erro_1$	$Erro_2$
Problema 1	4.771425e-005	1.658016e-004	Problema 1	-	-
Problema 2	5.463186e-007*	6.000000e-008	Problema 2	-	-
Problema 3	1.509437e-007	6.324555e-008	Problema 3	-	-
Problema 4	1.000000e-008	0.000000000000*	Problema 4	2.619949e-004	1.097249e-004*
Problema 5	4.296871e-007*	1.000000e-007	Problema 5	-	-

$Erro_1$: Erro relativo à solução primal

$Erro_2$: Erro relativo à solução dual

* Valores que correspondem a erros absolutos

Capítulo 6

Conclusão

Nesta dissertação apresentamos um pouco da teoria de alguns dos principais métodos para resolver problemas não-lineares com restrições: penalidade, barreira, Lagrangeana aumentada e barreira modificada. Nosso intuito foi comparar esses dois últimos métodos mencionados. Foi apresentado um estudo introdutório dos métodos de penalidade e barreira com o propósito de derivar os métodos de Lagrangeana aumentada e barreira logarítmica modificada.

Sabemos que tanto o método de penalidade quanto o método da barreira apresentam problemas de mal condicionamento quando o parâmetro c cresce. Mas o primeiro método possui uma vantagem sobre o segundo: a função penalidade é definida na solução ótima do problema enquanto a função barreira cresce indefinidamente quando os pontos se aproximam da solução ótima.

No trabalho de Nash, Polyak e Sofer [NPS94] foi realizada uma comparação entre os métodos barreira estabilizada e a da barreira logarítmica modificada. Este último teve um desempenho muito melhor em relação ao primeiro. Neste trabalho procuramos comparar o método da barreira modificada com o método da Lagrangeana aumentada, pois ambos buscam soluções através de atualizações sistemáticas nas variáveis primais e duais

Apesar de não trabalharmos com problemas reais, ao compararmos os métodos da Lagrangeana aumentada e da barreira modificada nos preocupamos em analisar a consistência das soluções obtidas. Analisando o número de iterações e o erro absoluto e/ou relativo das soluções podemos inferir qual método é eficaz. No nosso caso, como a estrutura dos dois métodos só se diferencia na resolução dos subproblemas e nas regras de atualização da variável dual, não achamos necessário a comparação do número de

avaliações da função e gradientes.

Nos estudos de caso que apresentamos, o método da Lagrangeana aumentada exigiu menos iterações que o método da barreira logarítmica modificada. Entretanto, no único problema não-convexo, o problema 3, o método da barreira logarítmica modificada convergiu na primeira, na segunda e na terceira abordagens. Já o método da Lagrangeana aumentada, para o mesmo problema, falhou na primeira e na terceira abordagens.

Sugerimos para trabalhos futuros a aplicação do método barreira modificada em problemas de grande porte como é o caso de problemas de fluxo de potência e fazer algumas alterações na resolução dos subproblemas e atualização do parâmetro de penalidade para torná-lo mais eficiente.

Referências Bibliográficas

- [And98] N. Andrei. Penalty-barrier algorithms for nonlinear optimization. preliminary computational results. *Studies in Informatics and Control*, 7(1):15–36, 1998.
- [Ber76] D. P. Bertsekas. Multiplier methods: A survey. *Automatica*, 12:133–145, 1976.
- [Ber99] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, Massachusetts, 2 edition, 1999.
- [BS93] M. S. Bazaraa and C. M. Shetty. *Nonlinear Programming: Theory and Algorithms*. John Wiley & Sons, New York, 2 edition, 1993.
- [BS06] C. Baptista and V. A. Sousa. A função barreira modificada e o problema de fluxo de potência Ótimo. *TEMA Tend. Mat. Apl. Comput.*, 7(1):21–30, 2006.
- [Car04] E. P. Carvalho. *Lagrangeana Aumentada e Barreira Combinadas com o Método do Gradiente Reduzido na Solução do Fluxo de Potência Ótimo*. PhD thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, SP, 2004.
- [CF61] C. W. Carroll and A. V. Fiacco. The created response surface technique for optimizing nonlinear restrained systems. *Operator Research*, 9:169–185, 1961.
- [CGT92] A. R. Conn, N. I. M. Gould, and P. L. Toint. *LANCELOT: a Fortran Package for Large-Scale Nonlinear Optimization (Release A)*, volume 17. Springer-Verlag, 1992.
- [Cos90] G. R. M. Costa. *O método Dual-Newton Aplicado ao Fluxo de Carga Ótimo*. PhD thesis, Departamento de Sistemas e Controle de Energia, Faculdade de En-

genharia Elétrica, Universidade Estadual de Campinas, Campinas,SP, Junho 1990.

- [CV03] T. W. C Chen and V. S. Vassiliadis. Solution of general nonlinear optimization problems using the penalty/modified barrier method with the use of exact hessians. *Computers Chem. Eng*, 27:501–525, 2003.
- [DS93] J. E. Dennis and R. B. Schnabel. *Numerical Methods. for Unconstrained Optimization and Nonlinear Equations*. Prentice Hall, Englewood Cliffs, 1993.
- [Fle81] R. Fletcher. *Practical Methods of Optimization*, volume 2. John Wiley & Sons, 1981.
- [FM68] A. V. Fiacco and G. P McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley & Sons, New York, 1968.
- [Fri55] K. R. Frisch. *The logarithmic Potencial method of Convex Programming*, Memorandum. University Institute of Economics, Oslo, Norway, 1955.
- [Hes69] M. R. Hestenes. Multiplier and gradient methods. *Journal of Optimization Theory and Applications*, 4(5), 1969.
- [JP94] D. L. Jensen and R. A. Polyak. The convergence of a modified barrier method for convex programming. *IBM J. RES. DEVELOP.*, 38(3):307–321, 1994.
- [Kar84] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica* 4, 54:373–395, 1984.
- [Kel99] C. T. Kelley. *Iterative Methods for Optimization*. SIAM, Philadelphia, 1999.
- [Lue86] D. G. Luenberger. *Linear and nonlinear programming*. Addison - Wesley Publishing Company, New York, 2 edition, 1986.
- [MCIL71] A. Miele, E. E. Cragg, R. R. Iver, and A. V. Levy. Use of the augmented penalty function in mathematical programming problems, part 1. *Journal of Optimization Theory and Applications*, 8(2):115130, 1971.

- [MCL71] A. Miele, E. E. Cragg, and A. V. Levy. Use of the augmented penalty function in mathematical programming problems, part 2. *Journal of Optimization Theory and Applications*, 8(2):131–153, 1971.
- [MMLC72] A. Miele, P. E. Moseley, A. V. Levy, and G. M. Coggins. On the method of multipliers for mathematical programming problems. *Journal of Optimization Theory and Applications*, 10(1):1–33, 1972.
- [MS95] J. M. Martínez and S. A. Santos. *20º Colóquio Brasileiro de Matemática: Métodos Computacionais de Otimização*. IMPA, Rio de Janeiro, 1995.
- [NPS94] S. G. Nash, R. Polyak, and A. Sofer. Numerical comparison of barrier and modified-barrier methods for large-scale bound-constrained optimization. *Large Scale Optimization*, pages 319–338, 1994.
- [NS93] S. G. Nash and A. Sofer. A barrier method for large-scale constrained optimization. *ORSA Journal on Computing* 5, pages 40–53, 1993.
- [NW99] J. Nocedal and J. S. Wright. *Numerical Optimization*. Springer-Verlag, 1999.
- [Pol92] R. Polyak. Modified barrier functions. *Mathematical Programming*, 54(2):177–222, 1992.
- [Pow69] M. J. D. Powell. *A method for nonlinear constraints in minimization problems, Optimization*. Academic Press, London, 1969.
- [PSU93] A. L. Peressini, F. E. Sullivan, and J. J. Jr. Uhl. *The Mathematics of Nonlinear Programming*. Springer-Verlag, New York, 1993.
- [Roc73a] R. T. Rockafellar. A dual approach solving nonlinear programming problem by unconstrained optimization. *Mathematical Programming*, pages 354–373, 1973.
- [Roc73b] R. T. Rockafellar. The multiplier method of hestenes and powell applied to convex programming. *Mathematical Programming*, 12(555-562), 1973.
- [Roc74] R. T. Rockafellar. Augmented lagrange multiplier functions and duality in nonconvex programming. *SIAM J. Control*, 12:268–285, 1974.

- [Sch87] K. Schittkowski. *More Test Examples for Nonlinear Programming Codes*, volume 282. Springer-Verlag, 1987.
- [Sou06] V. A. Sousa. *Resolução do Problema de Fluxo de Potência Ótimo Reativo Via Método da Função Lagrangeana Barreira Modificada*. PhD thesis, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, SP, 2006.
- [VB98] V. S. Vassiliadis and S. A. Brooks. Application of modified barrier method in large-scale quadratic programming problems. *Computers Chem. Eng.*, 22(9):1197–1205, 1998.
- [VF97] V. S. Vassiliadis and C. A. Floudas. The modified barrier function approach for large-scale optimization. *Computers Cem. Eng.*, 21(9):855–874, 1997.

Apêndice A

Definições

Convexidade

Definição 1 (Conjunto Convexo) Um conjunto $\mathcal{C} \subset \mathbb{R}^n$ é **convexo** se para todo x e y em $\mathcal{C}$, o segmento que liga x e y está contido em $\mathcal{C}$, ou seja,

$$\mathcal{C} \subset \mathbb{R}^n \text{ é convexo se } \forall x, y \in \mathcal{C}, \{\lambda x + (1 - \lambda)y \mid \lambda \in [0, 1]\} \subset \mathcal{C}.$$

A Figura A.1 ilustra dois exemplos de conjuntos, um convexo e outro não-convexo no $\mathbb{R}^2$.

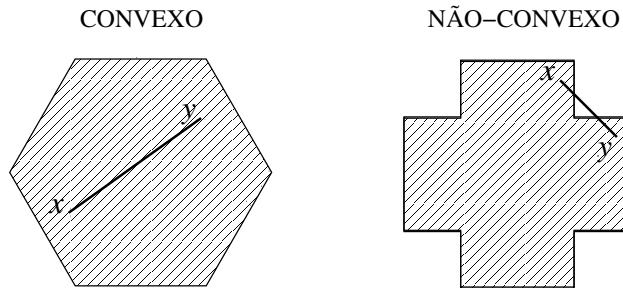


Figura A.1: Ilustração de um conjunto convexo e não-convexo.

Definição 2 (Função Convexa) Uma função f definida em um conjunto convexo $\mathcal{C} \in \mathbb{R}^n$ é **convexa** se para todo $x, y \in \mathcal{C}$, $\lambda \in [0, 1]$, $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$. Se para todo $\lambda \in (0, 1)$ e $x \neq y$, $f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$, diz-se que f é uma função **estritamente convexa**.

A figura A.2 mostra exemplos de uma função convexa, estritamente convexa e não-convexa no $\mathbb{R}$.



Figura A.2: Exemplos de funções convexas e não-convexas.

Condições de KKT

Consideremos o seguinte problema de programação não linear com restrições de igualdade

$$\begin{aligned} &\text{Minimizar} && f(x) \\ &\text{sujeito a} && h_j(x) = 0, \quad i = 1, \dots, p. \end{aligned} \tag{A.1}$$

A função Lagrangeana associada ao problema (A.1) é

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{j=1}^p \lambda_j h_j(x)$$

onde λ_j 's são os multiplicadores de Lagrange. As variáveis x e λ também são chamadas, respectivamente, de variáveis primais e duais.

Suponhamos que x^* e y^* sejam soluções ótimas do problema (A.1). As condições necessárias de primeira ordem para esse problema são,

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0 \quad (\text{condição de estacionaridade}) \tag{A.2}$$

$$h_j(x^*) = 0, j = 1, \dots, p \quad (\text{condição de factibilidade}) \tag{A.3}$$

Consideremos agora o seguinte problema de programação não-linear com restrições de desigualdade

$$\begin{aligned} &\text{Minimizar} && f(x) \\ &\text{sujeito a} && g_i(x) \geq 0, \quad i = 1, \dots, m \end{aligned} \tag{A.4}$$

a função Lagrangeana associada ao problema (A.4) é

$$\mathcal{L}(x, \mu) = f(x) - \sum_{i=1}^m \mu_i g_i(x)$$

onde μ_i 's são os multiplicadores de Lagrange.

Suponhamos que x^* e μ^* sejam soluções ótimas do problema (A.4). As condições necessárias de primeira ordem para esse problema são,

$$\nabla_x \mathcal{L}(x^*, \mu^*) = 0 \quad (\text{condição de estacionaridade}) \tag{A.5}$$

$$g_i(x^*) \geq 0, i = 1, \dots, m \quad (\text{condição de factibilidade}) \tag{A.6}$$

$$\mu_i^* \geq 0, i = 1, \dots, m \quad (\text{condição de factibilidade}) \tag{A.7}$$

$$\mu_i^* g_i(x^*) = 0, i = 1, \dots, m \quad (\text{condição de complementariedade}) \tag{A.8}$$

As condições (A.2), (A.3) e (A.5)-(A.8) são conhecidas também como condições de *Karush-Kuhn-Tucker* ou apenas por condições de KKT.

Apêndice B

Método de Newton

Método de Newton

A direção de Newton é dada por

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

O princípio desse método é minimizar uma função f através de uma aproximação local por uma função quadrática,

$$f(x(\lambda)) \approx f_q(x(\lambda)) = f(x^k) + \nabla f(x^k)^T (x(\lambda) - x^k) + \frac{1}{2} (x(\lambda) - x^k)^T \nabla^2 f(x^k) (x(\lambda) - x^k)$$

sendo $x(\lambda) = x^k + \lambda d$.

A iteração do método de Newton é definida como

$$x^{k+1} = x^k - \lambda \nabla^2 f(x^k)^{-1} \nabla f(x^k).$$

Algoritmo B.0.1 *Algoritmo do método de Newton*

Dado $x^0 \in \mathbb{R}^n$ arbitrário

$k = 0$

Enquanto $\nabla f(x^k) \neq 0$

1. Resolver $\nabla^2 f(x^k) d_k = -\nabla f(x^k)$

2. Busca linear para encontrar λ_k

3. Fazer $x^{k+1} = x^k + \lambda_k d_k$

4. $k = k + 1$

Nem sempre será possível determinar a direção de descida d_k no passo 1 do Algoritmo B.0.1, pois a matriz pode não ser definida positiva. Caso isso ocorra, uma estratégia seria dar continuidade com a iteração do método do gradiente até que a hessiana seja definida positiva.

Para o Algoritmo B.0.1 tem-se o seguinte resultado:

Teorema B.0.1 *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^3$ e $x^* \in \mathbb{R}^n$ um minimizador local de f , tal que a matriz $\nabla^2 f(x^*)$ é definida positiva. Então, existe $\varepsilon > 0$ tal que se $x^0 \in B(x^*, \varepsilon)$ e $\lambda_k = 1$ para todo $k \in \mathbb{N}$, a sequência $\{x^k\}$ gerada pelo Algoritmo B.0.1 verifica:*

(i) $\nabla^2 f(x^k)$ é definida positiva para todo $k \in \mathbb{N}$;

(ii) $\lim_{k \rightarrow \infty} x^k = x^*$;

(iii) Existe $c > 0$ tal que $\|x^{k+1} - x^*\| \leq c\|x^k - x^*\|^2$ para todo $k \in \mathbb{N}$.

Prova: Ver [Lue86]. $\square$

Esse é um resultado de convergência local que diz que se escolher um x^0 suficientemente próximo de x^* garante-se que,

(i) os sistemas lineares do passo 1 sempre serão possíveis de resolver, d_k está bem-definido para todo $k \in \mathbb{N}$;

(ii) a sequência converge a x^* ;

(iii) a ordem de convergência é pelo menos quadrática.

Apêndice C

Gráficos

Neste apêndice apresentamos as seqüências de soluções primais geradas pelos métodos da Lagrangeana aumentada e barreira modificada. Cada figura representa a seqüência de um dos problemas apresentados no Capítulo 5 com as seguintes abordagens:

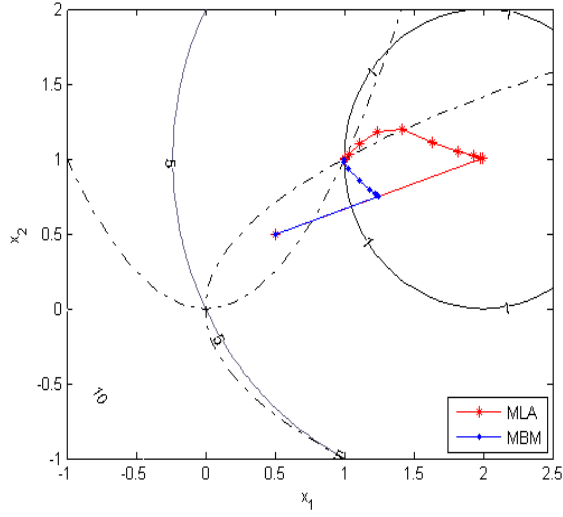
1ª abordagem: variando c e soluções iniciais x^0 factíveis;

2ª abordagem: variando c e soluções iniciais x^0 infactíveis;

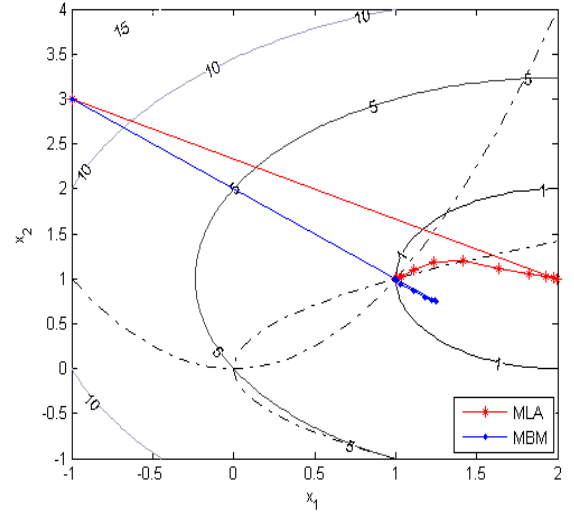
3ª abordagem: valor fixo de c e soluções iniciais x^0 factíveis,

4ª abordagem: valor fixo c e soluções iniciais x^0 infactíveis.

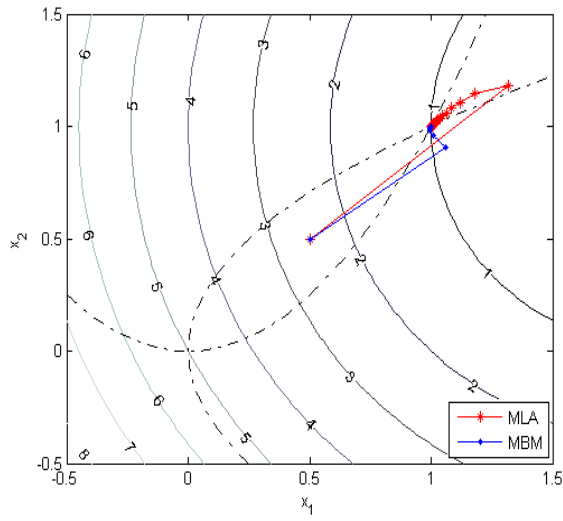
onde c para a primeira e segunda abordagens vale $1e - 3$ com $\beta = 3$ e para a terceira e quarta abordagens $c = 1$.



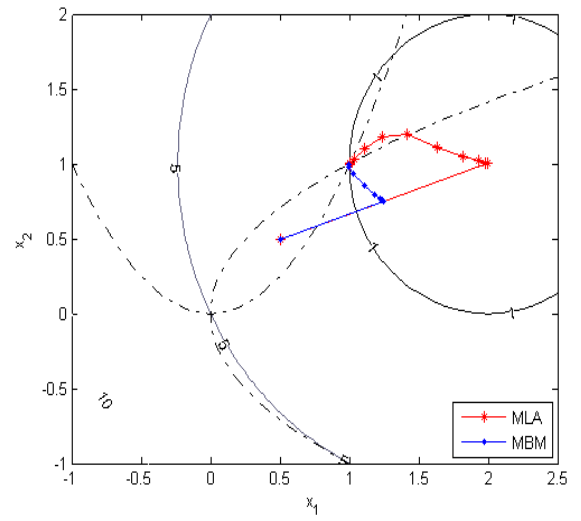
(a) Primeira abordagem



(b) Segunda abordagem

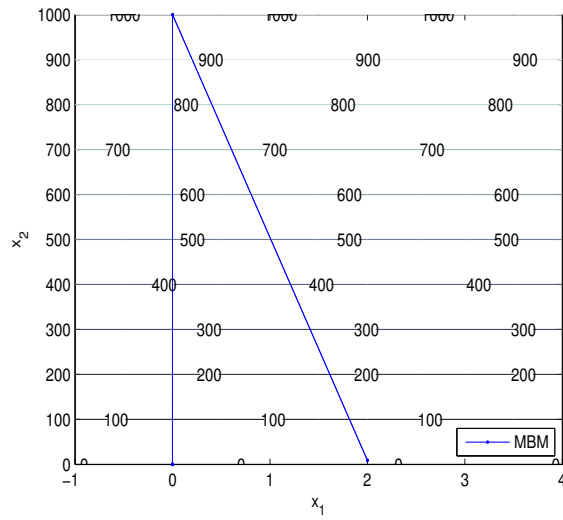


(c) Terceira abordagem

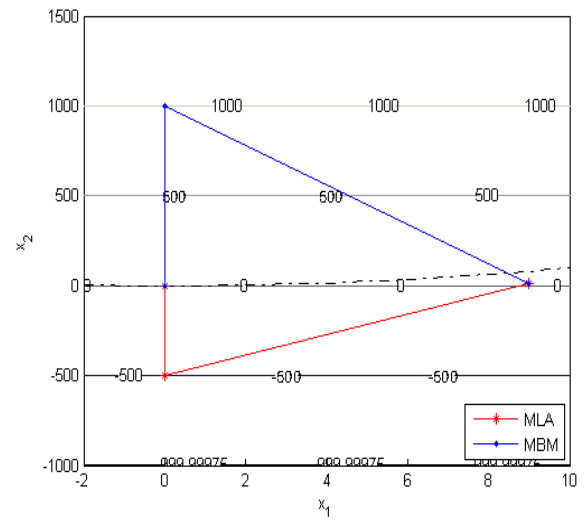


(d) Quarta abordagem

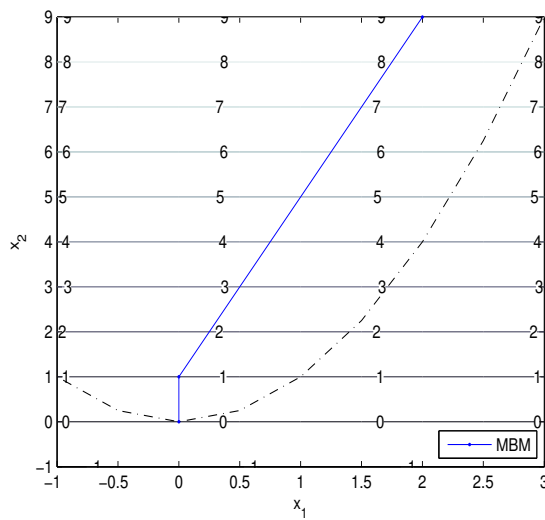
Figura C.1: Sequências de soluções primais geradas pelos métodos para o problema 1.



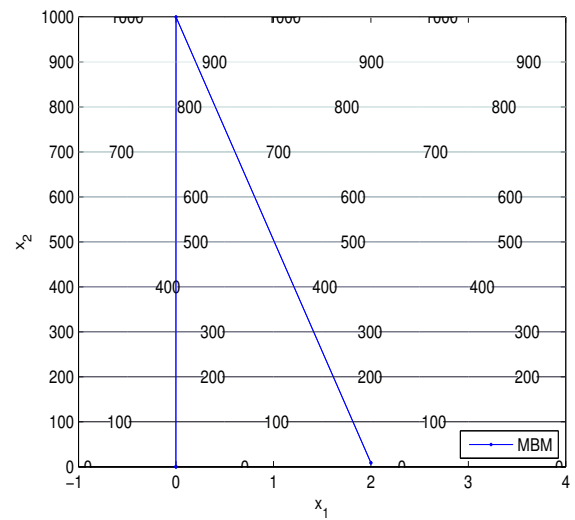
(a) Primeira abordagem



(b) Segunda abordagem

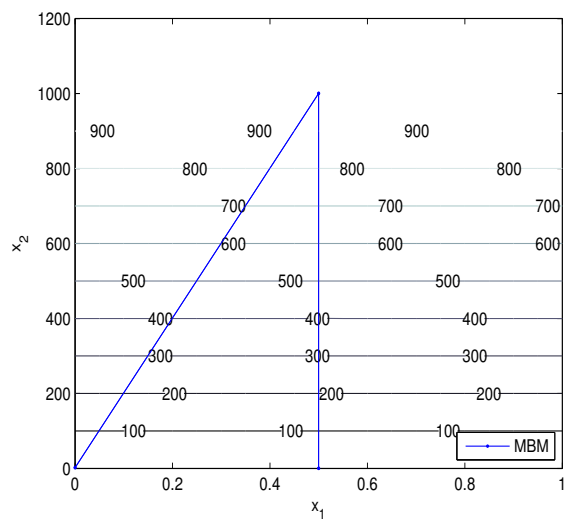


(c) Terceira abordagem

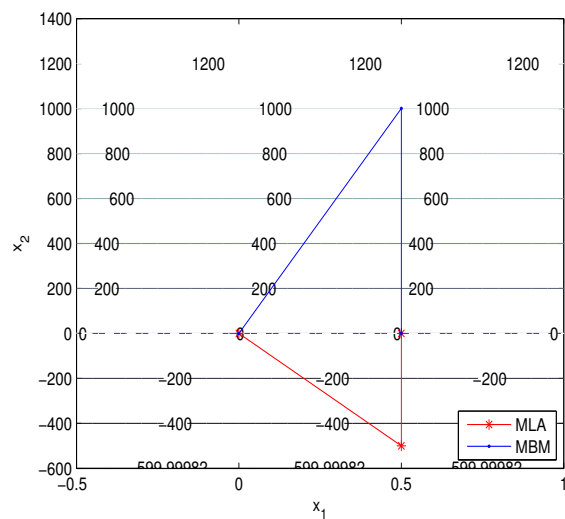


(d) Quarta abordagem

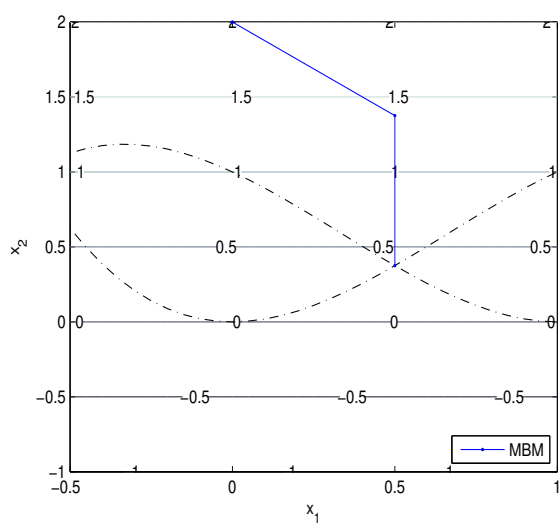
Figura C.2: Sequências de soluções primais geradas pelos métodos para o problema 2.



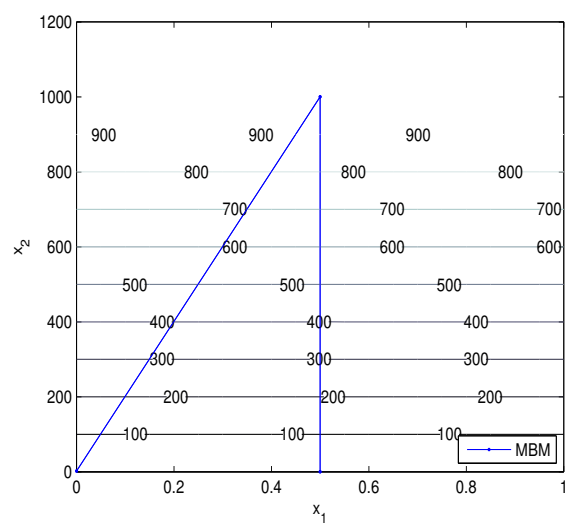
(a) Primeira abordagem



(b) Segunda abordagem

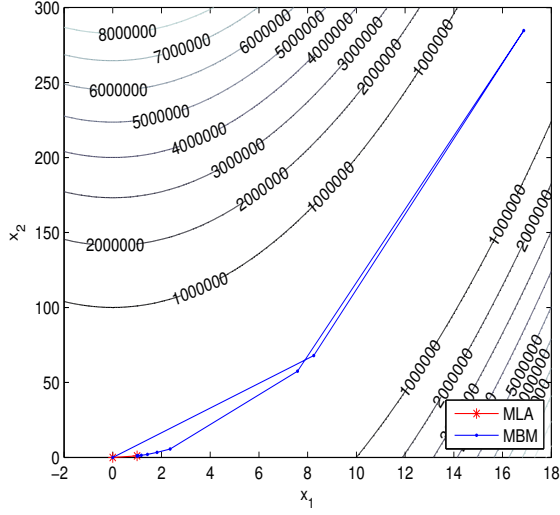


(c) Terceira abordagem

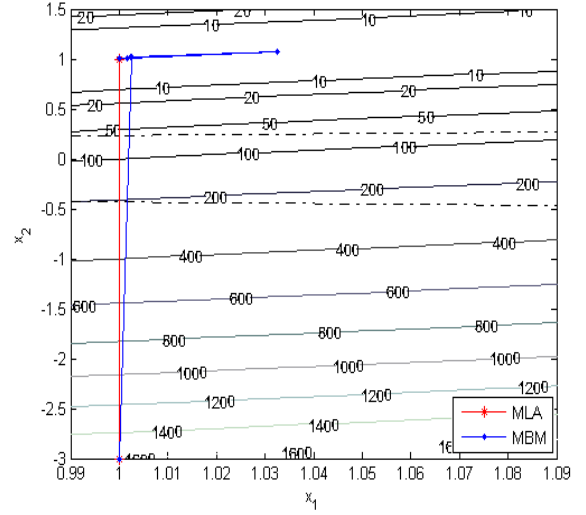


(d) Quarta abordagem

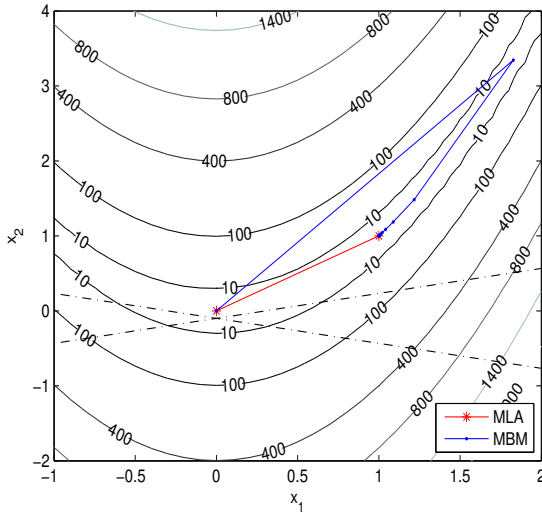
Figura C.3: Sequências de soluções primais geradas pelos métodos para o problema 3.



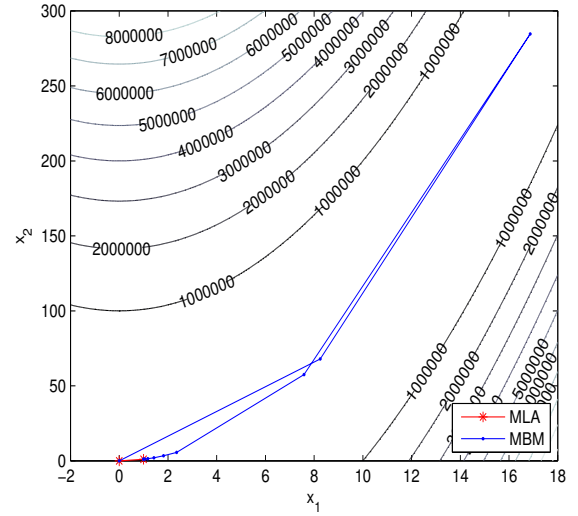
(a) Primeira abordagem



(b) Segunda abordagem

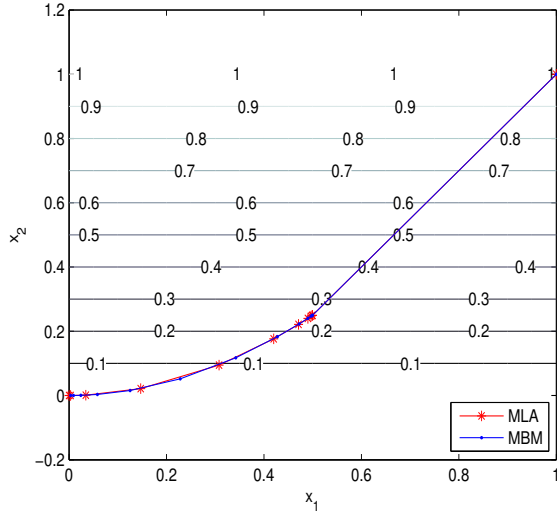


(c) Terceira abordagem

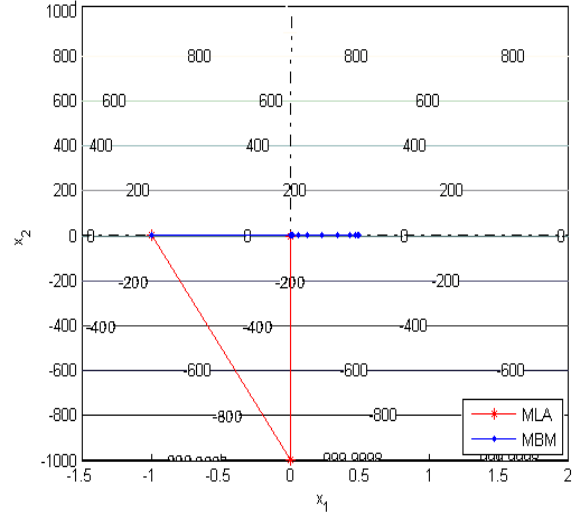


(d) Quarta abordagem

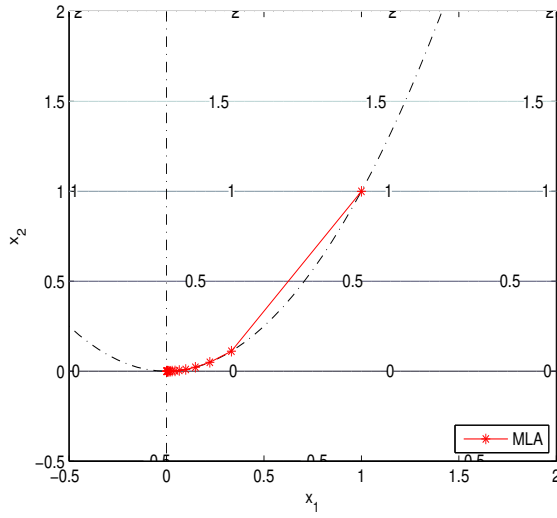
Figura C.4: Sequências de soluções primais geradas pelos métodos para o problema 4.



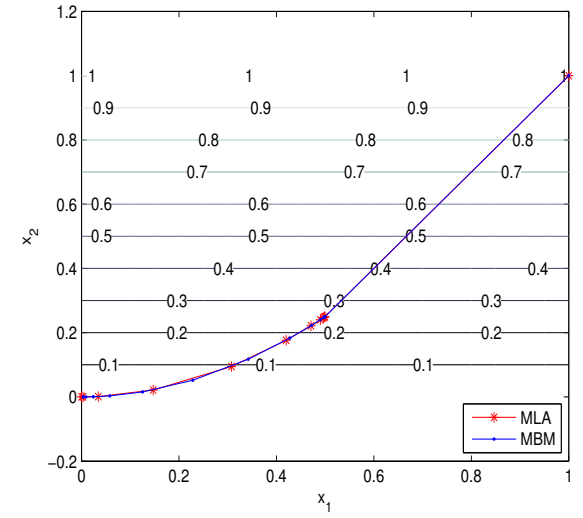
(a) Primeira abordagem



(b) Segunda abordagem



(c) Terceira abordagem



(d) Quarta abordagem

Figura C.5: Função auxiliar $\mathcal{F}_{\mathcal{P}}(x)$ para o problema (2.3) com diferentes parâmetros de penalidade.