

ISMAEL RODRIGO BLEYER

**UM ALGORITMO DE PONTOS INTERIORES PARA
O SUB-PROBLEMA TANGENCIAL EM MÉTODOS
DE FILTRO PARA PROGRAMAÇÃO NÃO LINEAR**

FLORIANÓPOLIS
2007

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS FÍSICAS E MATEMÁTICAS
DEPARTAMENTO DE MATEMÁTICA

UM ALGORITMO DE PONTOS INTERIORES PARA O SUB-PROBLEMA TANGENCIAL EM MÉTODOS DE FILTRO PARA PROGRAMAÇÃO NÃO LINEAR

Trabalho de Conclusão de Curso
apresentado ao Curso de Matemática
Habilitação Bacharelado
Departamento de Matemática
Centro de Ciências Físicas e Matemáticas
Universidade Federal de Santa Catarina

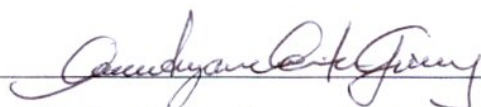
Ismael Rodrigo Bleyer¹

Orientador
Dr. Clóvis Caesar Gonzaga

Florianópolis, fevereiro de 2007.

¹Este trabalho teve apoio financeiro do PIBIC/CNPq/UFSC

Esta monografia foi julgada adequada como **TRABALHO DE CONCLUSÃO DE CURSO** no curso de Matemática - Habilitação Bacharelado e aprovada em sua forma final pela Banca Examinadora designada pela Portaria nº. 08/CCM/07.



Prof^ª. Carmen Suzane Comitre Gimenez

Professora responsável pela disciplina

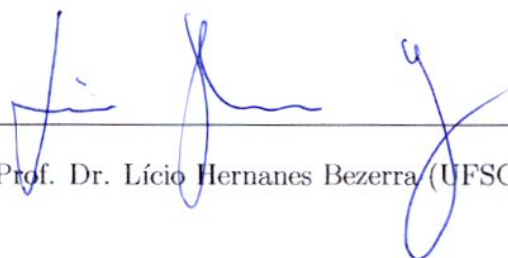
Banca Examinadora:



Prof. Dr. Clóvis Caesar Gonzaga (UFSC-Orientador)



Prof. Dr. Juliano de Bem Francisco (UFSC)



Prof. Dr. Lício Hernanes Bezerra (UFSC)

Sumário

Lista de Figuras	vi
Lista de Algoritmos	vii
Introdução	1
1 Considerações Gerais	3
1.1 Condições de otimalidade	3
1.2 Problema	4
1.3 Condições de KKT	5
1.4 Região de confiança para o problema irrestrito	6
1.5 Região de confiança para o problema restrito linearmente	9
2 Programação Quadrática Seqüencial e Filtros	10
2.1 Função Lagrangeano	11
2.2 Passo normal	12
2.3 Passo tangente	13
2.3.1 A região de confiança	15
2.3.2 Passo de região de confiança	15
2.4 Algoritmo de filtros	16
3 Método primal-dual de pontos interiores	19
3.1 Solução do problema	19

3.1.1	Uma iteração	21
3.2	Algoritmo de passo mais longo	23
4	Resultados principais	27
4.1	O passo normal	28
4.2	O passo tangente	28
4.2.1	Região de Confiança	29
4.2.2	O conjunto Ω	30
4.3	Transformação do Problema	32
5	Implementação	35
5.1	Pré-requisitos	35
5.2	Restauração	36
5.3	Otimização	36
5.4	Exemplo em \mathbb{R}^2	37
	Conclusão	39
	Referências Bibliográficas	42

Lista de Figuras

1.1	Região de confiança	7
1.2	Comparação de normas	8
2.1	Uma iteração de PQS	14
2.2	Construção do par (f^k, h^k)	17
2.3	Região proibida pelo filtro F	18
3.1	Passo mais longo	24
3.2	Comparação das vizinhanças da trajetória central	25
4.1	Construção da Região de confiança para o sub-problema tangencial	29
5.1	Solução aproximada do problema (5.2)	38
5.2	Região proibida pelo filtro	38

Lista de Algoritmos

1	Região de Confiança	8
2	Construção da caixa	29
3	Método tradicional	31
4	Método novo	31

Introdução

O problema geral de otimização restrita em \mathbb{R}^n é

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

com $\Omega \subset \mathbb{R}^n$ um conjunto fechado e $f : \Omega \rightarrow \mathbb{R}$ uma função contínua.

No problema de otimização diferenciável tomamos f continuamente diferenciável em um aberto que contém Ω . No estudo de algoritmos que utilizam aproximações quadráticas, todas as funções envolvidas são de classe C^2 , ou seja, duplamente continuamente diferenciáveis. Normalmente suporemos que todas as funções são definidas em todo o espaço.

Este trabalho é parte de uma linha de pesquisa iniciada com o artigo [12] sobre métodos de filtro para programação não linear (PNL), e que neste momento está sendo desenvolvido na tese de mestrado da Ana Maria Basei e na tese de Doutorado de Rafael Casali, usando ainda o resultado de Roger Behling [1].

O problema de PNL estudado tem o formato

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && c(x) = 0 \\ &&& x \geq 0 \end{aligned} \tag{2}$$

em que $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $c : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são funções de classe C^2 com $m < n$. Esse formato é bastante geral, uma vez que restrições de desigualdades podem ser reduzidas a igualdades com a introdução de variáveis de folga.

Nas seções iniciais do trabalho explicaremos como um algoritmo de Programação Quadrática Seqüencial com filtro gera em cada iteração um sub-problema de programação quadrática, o “problema tangencial”, com o seguinte formato:

$$\begin{aligned}
 &\text{minimizar} && v^T x + \frac{1}{2} x^T H x \\
 &\text{sujeito a} && Ax = b \\
 &&& l \leq x \leq u \\
 &&& x \in \Omega
 \end{aligned} \tag{3}$$

em que $v \in \mathbb{R}^n$, $H \in \mathbb{R}^{n \times n}$ é simétrica semi-definida positiva, $l < u$ e $A \in \mathbb{R}^{m \times n}$ é uma matriz de posto completo igual a m . O conjunto Ω é determinado por restrições impostas a cada iteração pelo filtro e pela restrição $x \geq 0$, e será comentado adiante. O ponto inicial dado $z \in \mathbb{R}^n$ é o centro da caixa $l \leq x \leq u$.

Este sub-problema, que é um problema geral de programação quadrática com restrições lineares, não deve ser resolvido exatamente: busca-se somente uma solução que satisfaça um critério de aceitação que será bastante detalhado adiante.

O método que se propõe nesta pesquisa conjunta consiste no seguinte: a partir do ponto z (centro da caixa), segue-se a trajetória central associada ao problema de programação quadrática até obter-se uma “boa” precisão ou até que se viole a restrição $x \in \Omega$.

O desenvolvimento detalhado da teoria estará na tese de mestrado da Ana Maria, e aqui expomos o mínimo necessário para caracterizar o problema e mostrar resultados de implementação feita em Matlab.

Capítulo 1

Considerações Gerais

Neste capítulo descreveremos o problema de programação não linear com restrições de igualdade e o algoritmo de regiões de confiança para problemas irrestritos e restritos linearmente.

1.1 Condições de otimalidade

Em otimização irrestrita, o conjunto viável Ω é o espaço todo, assim o problema (1) torna-se

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && x \in \mathbb{R}^n \end{aligned} \tag{1.1}$$

Nosso objetivo é encontrar um ponto x^* tal que $f(x^*) \leq f(x)$ para todo $x \in \Omega$, ou seja, um *minimizador global*. Visto que tal tarefa é difícil, vamos aceitar um minimizador local. Formalmente, um ponto x^* é *minimizador local* se existe uma vizinhança $V \subset \Omega$ de x^* tal que $f(x^*) \leq f(x)$ para todo $x \in V$.

Outra maneira de caracterizar se um ponto x^* é solução do problema irrestrito é dada pelas condições enunciadas pelos teoremas abaixo:

Teorema 1.1 (Condições Necessárias de Otimalidade de 1ª ordem) *Se $x^* \in \mathbb{R}^n$ é um minimizador local de f e $f \in C^1$ em uma vizinhança aberta de x^* , então*

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

Teorema 1.2 (Condições Necessárias de Otimalidade de 2ª ordem) *Se $x^* \in \mathbb{R}^n$ é um minimizador local de f e $\nabla^2 f$ é contínua em uma vizinhança aberta de x^* , então $\nabla f(x^*) = 0$ e $\nabla^2 f(x^*)$ é semi-definida positiva.*

Teorema 1.3 (Condições Suficientes de Otimalidade de 2ª ordem) *Suponha que $x^* \in \mathbb{R}^n$, $\nabla^2 f(x^*)$ é contínua em uma vizinhança aberta de x^* , $\nabla f(x^*) = 0$ e $\nabla^2 f(x^*)$ é definida positiva. Então x^* é um minimizador local.*

As demonstrações desses teoremas seguem na referência [20].

1.2 Problema

Dentre as variantes do problema geral descrito em (1), vamos nos dedicar inicialmente aos problemas de otimização com restrições de igualdade. Aqui definimos nosso problema padrão para esta seção:

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && c(x) = 0 \end{aligned} \tag{1.2}$$

com $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe C^2 , chamada de *função objetivo*, e $c : \mathbb{R}^n \rightarrow \mathbb{R}^m$, com $m < n$, tal que

$$x \longmapsto c(x) = [c_1(x), c_2(x), \dots, c_m(x)]^T$$

com $c_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$ funções de classe C^2 .

Vamos, ainda, introduzir uma notação para $x \in \mathbb{R}^n$,

$$A(x)^T = \begin{bmatrix} | & | & & | \\ \nabla c_1(x) & \nabla c_2(x) & \dots & \nabla c_m(x) \\ | & | & & | \end{bmatrix}.$$

$A(x)$ é a matriz das derivadas do problema (1.2), chamada de matriz *Jacobiana*. Durante todo nosso trabalho vamos supor que vale a condição de qualificação de independência linear (LICQ), ou seja, em todos os pontos de interesse o conjunto dos gradientes das restrições é linearmente independente. Portanto, temos que $A(x)^T$ tem colunas linearmente independentes, ou seja, $\text{posto}(A(x)^T) = m$.

1.3 Condições de KKT

Procurar uma solução para um problema com restrições, como (1.2), necessita de condições especiais de otimalidade, como segue:

Teorema 1.4 *Suponhamos f e c funções como descritas no problema (1.2) e ainda que a LICQ seja válida, se x é solução de (1.2), então existe um único $\lambda \in \mathbb{R}^n$ tal que*

$$\begin{cases} \nabla f(x) + A(x)^T \lambda = 0 \\ c(x) = 0 \end{cases} \quad (1.3)$$

As condições de otimalidade de 1ª ordem para o problema (1.2) dadas no teorema acima também são conhecidas como condições de Karush-Kunn-Tucker (KKT).

No par $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m$ que satisfaz (1.3), os pontos x são chamados *pontos estacionários*, enquanto λ são conhecidos como *multiplicadores de Lagrange*.

No sentido geométrico, o Teorema 1.4 nos diz que: o gradiente da função objetivo, $\nabla f(x)$, pode ser escrito como uma combinação linear dos gradientes das restrições. Isto equivale a dizer que $\nabla f(x)$ é ortogonal ao núcleo de $A(x)$. Para maiores detalhes dessa seção ver [13].

1.4 Região de confiança para o problema irrestrito

Nesta seção descreveremos um procedimento para encontrar uma solução para um problema irrestrito (1.1).

Em métodos de região de confiança, geramos uma seqüência de pontos $\{x^k\}_{k \in \mathbb{N}}$ convergindo para um minimizador local de f . Em cada iteração, dado um ponto x_k , construímos um modelo da função objetivo f em torno do ponto corrente e minimizamos esse modelo numa certa região. O tamanho da região é controlado iterativamente de modo que se possa confiar no modelo da função. Caso o passo não seja aceitável, reduzimos o tamanho da região e encontramos um novo minimizador.

Suponhamos que o nosso modelo seja quadrático, dado pelos três primeiros termos da expansão da série de Taylor de f em torno do ponto dado $x^k \in \mathbb{R}^n$, ou seja,

$$m_k(x) = f(x^k) + \nabla f(x^k)^T(x - x^k) + \frac{1}{2}(x - x^k)^T \nabla^2 f(x^k)(x - x^k) \quad (1.4)$$

Em cada iteração, procuramos a solução do sub-problema

$$\begin{aligned} &\text{minimizar} && m_k(x) \\ &\text{sujeito a} && \|x - x^k\| \leq \Delta_k \end{aligned} \quad (1.5)$$

no qual $\Delta_k > 0$ é chamado raio da região de confiança.

Definimos a redução *prevista* pelo modelo, no passo $d = x - x^k$, como

$$pred = m_k(x^k) - m_k(x) \quad (1.6)$$

e a redução *efetiva* como

$$ared = f(x^k) - f(x). \quad (1.7)$$

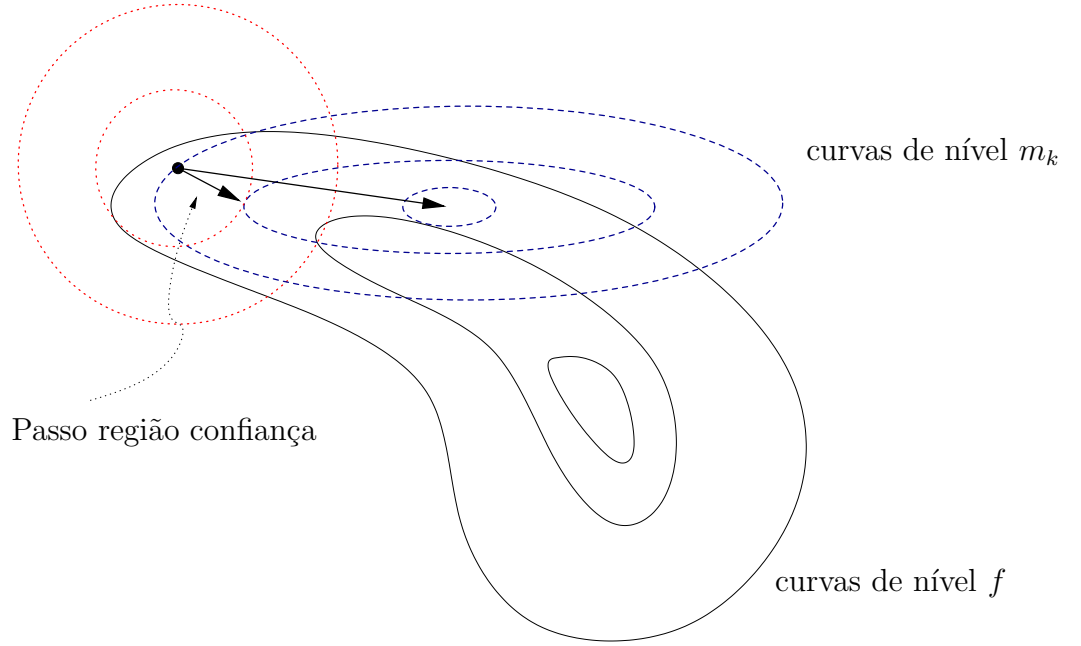


Figura 1.1: Região de confiança

Se

$$ared \geq \alpha \text{ pred} \quad (1.8)$$

com $\alpha \in (0, 1)$ fixo, podemos confiar no modelo na região definida e o passo d é aceito.

A região definida por $\|d\| \leq \Delta_k$ será uma bola caso usarmos a norma euclidiana (Norma $\|\cdot\|_2$) e neste caso encontramos vários métodos na literatura que resolvem tal problema, por exemplo o método de Moré e Sorensen e os métodos aproximados *Dog-Leg* e *Steihaug* (ver [20]). Para a norma infinito (Norma $\|\cdot\|_\infty$), ver os trabalhos de Martínez [17] e do grupo de pesquisa da UNICAMP, que desenvolveram o método BOX-QUACAN [16] baseado nesta metodologia.

Em nosso trabalho vamos resolver tal problema usando a norma infinito. Agora, nossa região torna-se uma caixa, o que nos trará vantagem pelo tamanho da região em relação à bola.

O algoritmo 1 utiliza os seguintes parâmetros:

- $\alpha \in (0, 1)$, usualmente $\alpha = 0.25$ (ou $\alpha = 0.01$ em implementações práticas). Se $ared/pred < \alpha$, o modelo não é confiável, não se aceita o passo e reduz-se a

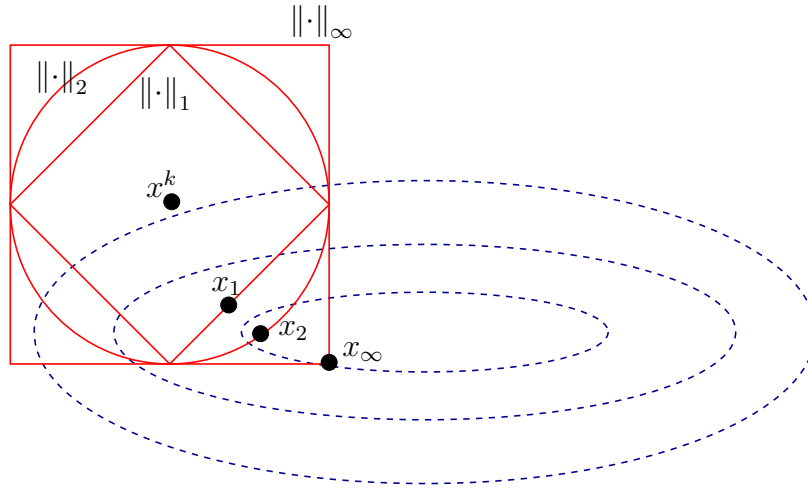


Figura 1.2: Comparação de normas

Algoritmo 1 Região de Confiança

Dados: $x^0 \in \mathbb{R}^n$, $\Delta_0 > 0$, $0 < \alpha < \gamma < 1$, $\beta_1 \in (0, 1)$, $\beta_2 > 1$

- 1: **para** $k = 0, 1, 2, \dots$ **faça**
 - 2: construir modelo $m_k(\cdot)$ (1.4)
 - 3: obter x solução do problema de região de confiança (1.5)
 - 4: calcular $ared = f(x^k) - f(x)$ e $pred = m_k(x^k) - m_k(x)$
 - 5: $r = ared/pred$
 - 6: **se** $r \geq \alpha$ **então** ▷ aceitar o passo
 - 7: $x^{k+1} = x^k + d$
 - 8: **se** $r > \gamma$ **então**
 - 9: $\Delta_{k+1} = \beta_2 \Delta_k$ ▷ aumentar a região de confiança
 - 10: **senão**
 - 11: $\Delta_{k+1} = \Delta_k$
 - 12: **fim se**
 - 13: **senão**
 - 14: $x^{k+1} = x^k$
 - 15: $\Delta_{k+1} = \beta_1 \Delta_k$ ▷ reduzir a região de confiança
 - 16: **fim se**
 - 17: **fim para**
 - 18: até obter convergência
-

região de confiança.

- $\beta_1 \in (0, 1)$, usualmente $\beta_1 = 0.25$: fator de redução da região de confiança.
- $\gamma \in (0, 1)$, usualmente $\gamma = 0.75$: $ared/pred > \gamma$ indica não somente que o passo deve ser aceito, como que o modelo está muito preciso e pode-se aumentar o raio da região de confiança para a iteração seguinte.
- $\beta_2 > 1$, usualmente $\beta_2 = 2$: fator de aumento da região de confiança.

1.5 Região de confiança para o problema restrito linearmente

O problema de programação não linear com restrições de igualdade é

$$\begin{aligned} & \text{minimizar} && f(x) \\ & \text{sujeito a} && Ax = b \end{aligned} \tag{1.9}$$

em que $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $A \in \mathbb{R}^{m \times n}$ é uma matriz de posto completo com $m < n$.

O problema de região de confiança a partir de x^k tal que $Ax^k = b$ agora passa a ser

$$\begin{aligned} & \text{minimizar} && m_k(x) \\ & \text{sujeito a} && \|x - x^k\| \leq \Delta_k \\ & && A(x - x^k) = 0 \end{aligned} \tag{1.10}$$

Este problema é mais difícil do que (1.5). Se a norma utilizada é $\|\cdot\|_2$, os métodos de solução são bem conhecidos e são semelhantes aos do caso irrestrito (ver [20]). Com a norma $\|\cdot\|_\infty$, o problema é bem mais difícil: trata-se de um problema geral de programação quadrática. Nossa pesquisa trata deste problema, com restrições adicionais e será detalhada adiante.

Capítulo 2

Programação Quadrática

Seqüencial e Filtros

Neste capítulo vamos expor, de forma bem reduzida, algoritmos de Programação Quadrática Seqüencial (PQS) baseados em restauração inexata e filtro, com a finalidade de construir o problema de otimização tangencial exposto na introdução.

Para manter o tratamento simples, consideremos neste capítulo somente restrições de igualdade $c(x) = 0$. As restrições de não negatividade $x \geq 0$ serão incorporadas mais adiante.

O método de PQS é um dos métodos mais eficazes para otimização não linear com restrições.

A idéia desse método consiste em substituir, a cada iteração, a função objetivo por um modelo quadrático e as restrições de igualdade por equações lineares, fazendo uma aproximação linear para as mesmas.

Em tal método temos dois objetivos conflitantes a considerar: otimalidade e viabilidade, dando preferência ao último, uma vez que a solução deverá ser viável. Tais objetivos serão separados em passos distintos, o passo normal (passo de restauração) e o passo tangente. Essa metodologia foi proposta por Martínez e Pilota [18], e é conhecida como método de restauração inexata. Cada iteração do método

proposto é formada por um passo normal seguido por um passo tangente (passo de otimalidade), dessa maneira construímos uma sequência de pontos $\{x^k\}_{k \in \mathbb{N}}$ nos quais, em cada iteração, nos aproximamos do conjunto viável e em seguida reduzimos o valor da função objetivo. Tal sequência deverá convergir para um ponto viável minimizador local de f . Note-se que é impossível exigir que ambos critérios sejam reduzidos em cada iteração, já que são conflitantes.

Antes de descrevermos cada passo, precisamos introduzir a função Lagrangeano.

2.1 Função Lagrangeano

Dado um problema como em (1.2) a função Lagrangeano associado ao problema fixado é dada por

$$(x, \lambda) \in \mathbb{R}^{n+m} \mapsto \mathcal{L}(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i c_i(x) = f(x) + \lambda^T c(x) \quad (2.1)$$

em que $x \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}^m$ é o vetor dos multiplicadores de Lagrange.

Vale a pena observar que

$$\nabla_x \mathcal{L}(x, \lambda) = \nabla f(x) + \sum_{i=1}^m \lambda_i \nabla c_i(x) = \nabla f(x) + A(x)^T \lambda \quad (2.2)$$

$$\nabla_\lambda \mathcal{L}(x, \lambda) = c(x) \quad (2.3)$$

Utilizando as igualdades encontradas acima, podemos reescrever as condições de KKT (1.3) para o problema com restrições de igualdade como

$$\begin{cases} \nabla_x \mathcal{L}(x, \lambda) = 0 \\ \nabla_\lambda \mathcal{L}(x, \lambda) = 0, \end{cases} \quad (2.4)$$

ou ainda,

$$\begin{cases} \nabla_x \mathcal{L}(x, \lambda) = 0 \\ c(x) = 0 \end{cases} \quad (2.5)$$

como é mais encontrado na literatura.

2.2 Passo normal

No passo normal, também chamado passo de restauração, desejamos em cada iteração melhorar a viabilidade do problema, isto é, dado um ponto $x^k \in \mathbb{R}^n$ tal que $c(x^k) \neq 0$ queremos encontrar um novo ponto $z \in \mathbb{R}^n$ da forma $z = x^k + d$, em que $d \in \mathbb{R}^n$ possua a propriedade de melhorar a viabilidade do ponto corrente.

Para avaliar se o novo ponto encontrado é melhor ou não, definimos uma medida de inviabilidade, a saber

$$x \in \mathbb{R}^n \mapsto h(x) = \|c(x)\| \quad (2.6)$$

com $\|\cdot\|$ uma norma qualquer em \mathbb{R}^n , por exemplo a norma euclidiana.

Dessa maneira, o ponto será aceito se

$$h(z) \ll h(x^k). \quad (2.7)$$

Não vamos nos ater muito a este passo, para encontrar tal ponto usamos algum método de programação não linear que minimize $h(\cdot)$, aceitando o passo se houver um decréscimo de no mínimo 50%. Os métodos mais utilizados baseiam-se em passos de Newton para o problema $c(x) = 0$ (sistema de equações não linear).

2.3 Passo tangente

No passo tangente, ou passo de otimalidade, dado um ponto $z \in \mathbb{R}^n$ do passo anterior, queremos obter um novo ponto $w \in \mathbb{R}^n$ da forma

$$w = z + d, \quad d \in \mathbb{R}^n \quad (2.8)$$

que reduza o valor da função objetivo, isto é,

$$f(w) < f(z), \quad (2.9)$$

mas sem piorar muito a viabilidade. Para garantir isso, usamos direções¹ no espaço tangente às restrições em z , ou seja,

$$d \in \ker(A(z)).$$

Poderíamos tentar resolver o problema seguinte:

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && A(z)(x - z) = 0. \end{aligned} \quad (2.10)$$

No entanto, esse problema ignora a curvatura das restrições. Resultados muito melhores são obtidos se tentarmos minimizar a função Lagrangeano, utilizando multiplicadores de Lagrange bem escolhidos. Para um tratamento completo ver [20].

Vamos portanto tentar reduzir $\mathcal{L}(\cdot, \lambda)$ a partir de z . Estudaremos um passo de região de confiança para esse problema.

Se $\bar{\lambda}$, multiplicador ótimo, é conhecido, conseguimos construir o modelo quadrático do Lagrangeano em z

$$m_{\mathcal{L}}(z + d, \bar{\lambda}) = \mathcal{L}(z, \bar{\lambda}) + \nabla \mathcal{L}(z, \bar{\lambda})^T d + \frac{1}{2} d^T L(z, \bar{\lambda}) d \quad (2.11)$$

¹Aqui estamos novamente usando a notação d para direção, não fazer confusão com a direção do passo normal.

em que $L(z, \bar{\lambda})$ denota a hessiana do Lagrangeno e $\nabla \mathcal{L}(z, \bar{\lambda})$ o vetor gradiente, dado pela derivada em relação a primeira variável.

Caso $\bar{\lambda}$ seja desconhecido, podemos calcular um estimador para os multiplicadores de Lagrange utilizando a primeira condição de KKT (1.3). A solução do problema de mínimos quadráticos para esta igualdade é dada por

$$\tilde{\lambda} = -(A(x)A(x)^T)^{-1}A(x)\nabla f(x). \quad (2.12)$$

Conhecido λ , multiplicador ótimo ou não, o melhor procedimento nesta etapa é minimizar $m_{\mathcal{L}}(\cdot, \lambda)$ ao longo do núcleo de $A(z)$, que resulta em resolver um novo problema:

$$\begin{aligned} \text{minimizar} \quad & m_{\mathcal{L}}(z + d, \lambda) \\ \text{sujeito a} \quad & A(z)d = 0 \end{aligned} \quad (2.13)$$

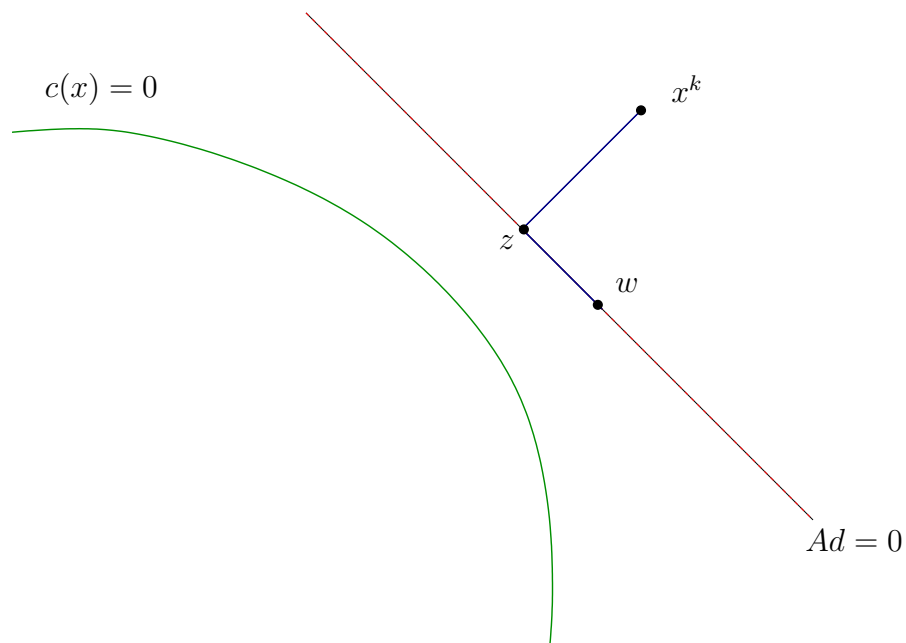


Figura 2.1: Uma iteração de PQS

Convém fazer umas observações interessantes:

- A função $\mathcal{L}(x, \lambda)$ incorpora informação sobre a curvatura das restrições;

- Em uma solução ótima \bar{x} com multiplicadores $\bar{\lambda}$, a hessiana do Lagrangeano $L(\bar{x}, \bar{\lambda})$ é semi definida positiva no espaço do núcleo da matriz Jacobiana $A(\bar{x})$, isto é,

$$d^T L(\bar{x}, \bar{\lambda})d \geq 0, \forall d \in \ker(A(\bar{x})).$$

Se for satisfeita a condição de otimalidade de segunda ordem (Teorema 1.2), $L(\bar{x}, \bar{\lambda})$ é definida positiva, e esta condição vale em uma vizinhança de $(\bar{x}, \bar{\lambda})$ em \mathbb{R}^{n+m} .

2.3.1 A região de confiança

Nosso problema ainda não está completo, uma vez que estamos utilizando um modelo quadrático precisamos verificar se a aproximação dada por ele é confiável ou não, isto é, devemos construir uma região em torno do ponto intermediário z e comparar a redução prevista pelo modelo com a redução da função a ser minimizada. Para isto basta adicionar uma restrição de região de confiança no problema (2.13) e trabalhar com as idéias expostas na seção 1.4. Agora nosso problema de região de confiança é

$$\begin{aligned} \text{minimizar} \quad & m_{\mathcal{L}}(z + d, \lambda) \\ \text{sujeito a} \quad & A(z)d = 0 \\ & \|d\| \leq \Delta_k \end{aligned} \tag{2.14}$$

2.3.2 Passo de região de confiança

Um passo completo de região de confiança pode ser feito como no algoritmo 1, reduzindo Δ_k se necessário até que a redução efetiva do Lagrangeano seja aceitável. No capítulo 4 estudaremos o passo completo.

Suponhamos que o ponto $w = z + d$ solução de (2.14) satisfaz o critério de redução efetiva. Resta-nos saber se o ponto candidato $w = z + d$ solução do problema (2.13) deve ser aceito ou não, ou melhor, se w melhora a otimalidade sem piorar

muito a viabilidade. Caso afirmativo definimos $x^{k+1} = w$ como o próximo ponto para a seqüência gerada pelo algoritmo de PQS e damos continuidade no procedimento até obter convergência.

2.4 Algoritmo de filtros

Continuando com a mesma notação da seção anterior, iniciamos com um ponto x^k e encontramos um ponto intermediário z do passo normal. Resolvido² o problema (2.14) devemos tomar a decisão de aceitar $w = z + d$ como novo ponto para a seqüência $\{x^k\}_{k \in \mathbb{N}}$, ou melhor, se tal ponto melhora a otimalidade sem piorar muito a viabilidade. Usaremos o método de *filtros* para tal decisão.

Definição: O ponto $x \in \mathbb{R}^n$ domina o ponto $y \in \mathbb{R}^n$ se e só se $f(x) \leq f(y)$ e $h(x) \leq h(y)$.

Convém lembrar que h é a medida de inviabilidade definida em (2.6).

Idealmente, gostaríamos de obter w tal que w domina x^k , mas isso é geralmente impossível porque os critérios são conflitantes. Mas não se pode admitir que x^k domine w , pois ambos critérios piorariam. Mais do que isso, não se pode aceitar que w seja dominado por nenhum iterado x^i , $i = 1, \dots, k$.

Não vamos descrever detalhes do método de filtros, veja [12].

Dado um iterando x^k , definimos o seguinte par:

$$(f^k, h^k) = (f(x^k) - \delta, h(x^k) - \delta) \text{ , com } \delta = 0.2h(x^k) \quad (2.15)$$

A região $\mathcal{F}_k = \{x \in \mathbb{R}^n \mid f(x) > f^k \text{ e } h(x) > h^k\}$ será proibida em iterações futuras. Tomadas várias iterações do algoritmo (nem todos os pontos $\{x^k\}_{k=1,2,\dots}$ são utilizados no filtro - ver referência [12]), construímos o filtro: o filtro é um conjunto de pares ordenados $\{(f^i, h^i)\}_{i=1,\dots,p}$, cada par é construído usando a equação (2.15), que guarda informações sobre as iterações passadas definindo uma região proibida como

²Essa suposição ainda é forte, pois ainda não temos ferramentas suficientes para isso.

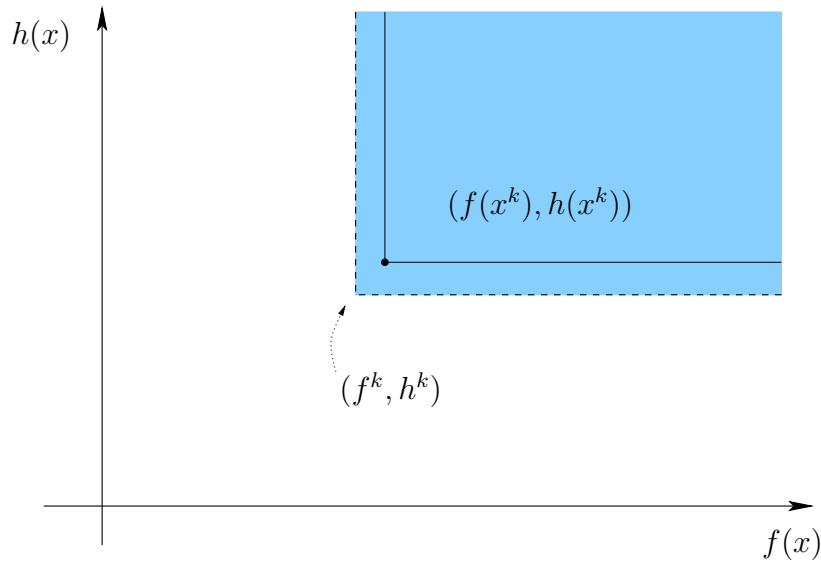


Figura 2.2: Construção do par (f^k, h^k)

na figura 2.3.

Dado um filtro, define-se a região proibida em \mathbb{R}^n como

$$\mathcal{F} = \{x \in \mathbb{R}^n \mid f(x) > f^i \text{ e } h(x) > h^i \text{ para algum } i = 1, \dots, p\}. \quad (2.16)$$

Não se pode descrever essa região do \mathbb{R}^n , mas é muito fácil determinar se um ponto $w \in \mathbb{R}^n$ dado pertence a \mathcal{F} : basta fazer p comparações em \mathbb{R}^2 .

No algoritmo de PQS, exigimos o seguinte:

- Dado um ponto x^k , adicionamos ao filtro o ponto correspondente como acima, aumentando o filtro.
- Exigimos que o passo de restauração gere um ponto z não proibido (o que é fácil).

No método de filtro exposto em [12], faz-se o passo de região de confiança tangencial, obtendo w . Se w está na região proibida, é necessário reduzir Δ_k e refazer o passo tangencial, até obter w na região permitida.

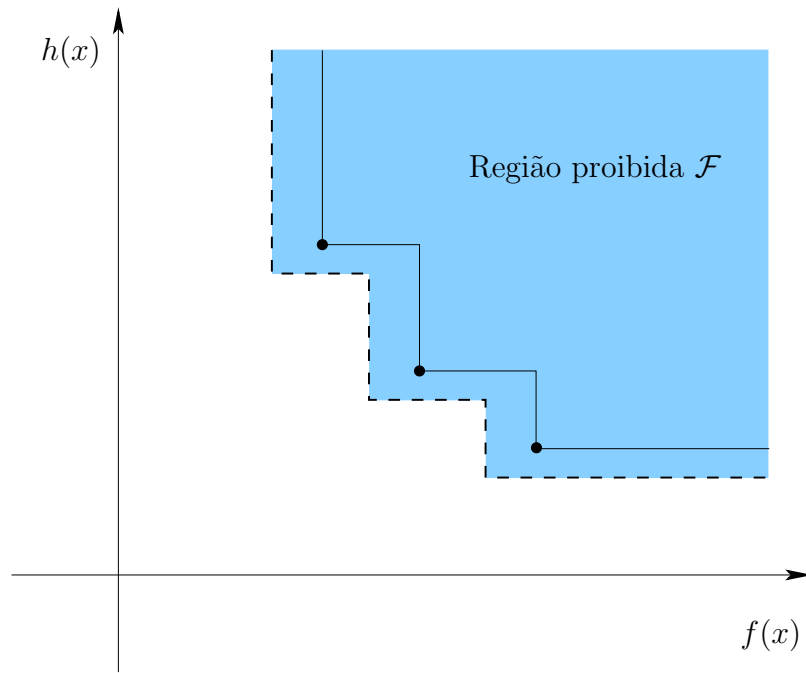


Figura 2.3: Região proibida pelo filtro F

Neste trabalho usaremos outra estratégia, sem refazer o passo de região de confiança: a restrição $x \notin \mathcal{F}$ será incorporada à definição de Ω no enunciado do passo tangencial descrito na introdução (3).

Capítulo 3

Método primal-dual de pontos interiores

Este capítulo é independente dos outros. Ele descreve sucintamente um algoritmo de trajetória central primal-dual para o problema geral de programação quadrática convexo. Baseia-se nas referências [11] e [23].

O formato padrão para o problema a ser resolvido neste capítulo é o problema geral de programação quadrática convexa com restrições lineares:

$$\begin{aligned} &\text{minimizar} && v^T x + \frac{1}{2}x^T H x \\ &\text{sujeito a} && Ax = b \\ &&& x \geq 0 \end{aligned} \tag{3.1}$$

com $x, v \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ e $H \in \mathbb{R}^{n \times n}$ uma matriz simétrica semi-definida positiva. Um problema misto com restrições de igualdade e desigualdades.

3.1 Solução do problema

A fim de obter a solução do problema (3.1), vamos examinar as condições de KKT (maiores detalhes ver [23]) para um problema com restrições mistas:

$$v + Hx + A^T y - s = 0 \quad (3.2)$$

$$Ax = b \quad (3.3)$$

$$x \cdot s = 0 \quad (3.4)$$

$$(x, s) \geq 0 \quad (3.5)$$

em que y e s denotam os multiplicadores de Lagrange para a restrição de igualdade e desigualdade, respectivamente. A notação $x \cdot s$ denota o produto de x por s componente a componente.

Para resolver o problema (3.1) devemos encontrar (x, y, s) que resolvem as condições de KKT dadas acima. Poderíamos tentar encontrar uma solução através do método de Newton, definindo uma função vetorial afim de encontrar o “zero” dessa função, contudo, isso é uma tarefa muito difícil, pois (3.4) é um problema combinatorio. Para contornar essa dificuldade relaxamos a condição de complementaridade (3.4) fazendo para $\tau > 0$,

$$x_i s_i = \tau, \quad i = 1, 2, \dots, n \quad (3.6)$$

com n a dimensão do vetor x . O método *Primal-dual* resolve as condições de KKT relaxadas. Com isto, elimina-se a restrição difícil $x \cdot s = 0$ e também as desigualdades que passam a $(x, s) > 0$, definindo um conjunto aberto. Sabe-se que com nossas hipóteses, esse problema sempre tem solução única (ver [23]). Se iniciarmos em um ponto próximo da solução ótima do problema relaxado (discutiremos “proximidade” abaixo), sabe-se que o método de Newton é muito eficiente.

Variando o parâmetro $\tau > 0$, os pontos (x_τ, y_τ, s_τ) , solução do sistema de KKT acima com (3.6), formam um curva denominada *trajetória central*

$$\mathcal{C} = \{(x_\tau, y_\tau, s_\tau) \mid \tau > 0\}. \quad (3.7)$$

Quando τ tende a zero, a solução do sistema de KKT relaxada se aproxima de uma solução do sistema de KKT exato, satisfazendo (3.4).

Métodos de trajetória central resolvem aproximadamente o problema relaxado para uma seqüência decrescente de valores de τ tendendo a zero, seguindo a trajetória central.

Vamos descrever um algoritmo de trajetória central, que foi implementado neste trabalho. O algoritmo parte de uma solução viável (x^0, s^0, y^0) , que será discutida adiante. Aqui, supomos que (x^0, s^0, y^0) está disponível.

3.1.1 Uma iteração

Uma iteração de um algoritmo de trajetória parte de uma solução primal-dual viável (x^k, s^k, y^k) satisfazendo (3.2) e (3.3) com $(x^k, s^k) > 0$. Calcula-se uma medida de dualidade $\mu = \frac{(x^k)^T s^k}{n}$, introduzindo um parâmetro de centralização $\sigma \in [0, 1]$ e faz-se um passo de Newton para resolver o sistema

$$\begin{aligned} v + Hx + A^T y - s &= 0 \\ Ax &= b \\ x \cdot s &= \sigma \mu \end{aligned}$$

Calculando o passo de Newton, obtemos

$$(x^{k+1}, y^{k+1}, s^{k+1}) = (x^k, y^k, s^k) + (\Delta x, \Delta y, \Delta s) \quad (3.8)$$

resolvendo

$$\begin{bmatrix} H & A^T & -I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -XSe + \sigma\mu e \end{bmatrix}, \quad (3.9)$$

usando S e X para denotar as matrizes cujas diagonais são dadas pelos vetores s^k e x^k , respectivamente. O vetor formado por uns é denotado por e com mesma dimensão de x^k . A matriz do lado esquerdo de (3.9) é a matriz Jacobiana $J(x, y, s)$ para a função vetorial

$$F(x, y, s) = \begin{bmatrix} v + Hx + A^T y - s \\ Ax - b \\ XSe - \sigma\mu e \end{bmatrix} \quad (3.10)$$

definida pelas condições de KKT relaxadas. Aqui $X = \text{diag}(x)$ e $S = \text{diag}(s)$.¹

Em resumo, estamos interessados em resolver

$$J(x^k, y^k, s^k) \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = -F(x^k, y^k, s^k) \quad (3.11)$$

e atualizar o ponto fazendo

$$(x^{k+1}, y^{k+1}, s^{k+1}) = (x^k, y^k, s^k) + (\Delta x, \Delta y, \Delta s). \quad (3.12)$$

Neste instante termina uma iteração do método. Recalculamos o valor de μ e iniciamos uma nova iteração, até obter convergência.

O problema central é a escolha de σ em cada iteração. A próxima seção mostra um método para a escolha desse parâmetro, de modo a garantir que todas as iterações são bem definidas.

Vamos usar o seguinte teorema muito importante (aqui em uma versão

¹Estamos usando a mesma linguagem do Matlab.

simplificada):

Teorema 3.1 *Se o ponto (x, y, s) satisfaz $\left\| \frac{x \cdot s}{\sigma \mu} - e \right\| \leq 0.5$, então o ponto (x^+, y^+, s^+) resultante de um passo de Newton satisfaz $\left\| \frac{x^+ \cdot s^+}{\sigma \mu} - e \right\| \leq 0.25$.*

Usaremos esse resultado sem demonstrá-lo. A demonstração é encontrada no artigo de Gonzaga (ver [10]).

Proximidade: Definimos para $(x, s) > 0$, $\mu > 0$ a seguinte medida de proximidade, usando o teorema acima:

$$\delta(x, s, \mu) = \left\| \frac{x \cdot s}{\sigma \mu} - e \right\|.$$

3.2 Algoritmo de passo mais longo

Existe uma vasta literatura sobre o passo primal-dual. A questão central é a escolha de σ : σ pequeno força um passo longo, mas o passo de Newton pode não ser aceitável; σ próximo de 1 faz uma “centralização”, aproximando o ponto da trajetória central sem avançar.

Sabemos que o passo de Newton é eficiente se todos os pontos estiverem em uma região “próxima” da trajetória central. O método exposto adiante define uma medida de proximidade e procura o passo de Newton mais longo (escolhendo σ) que permanece próximo da trajetória. Para detalhes e teoria, ver [11].

Quando fazemos $\sigma = 1$, a solução do sistema (3.9) define uma direção de *centralização*; quando $\sigma = 0$ temos uma direção *afim-escala*. Assim, definimos

$$\sigma = 0 : (x^a, y^a, s^a) = (\Delta x^a, \Delta y^a, \Delta s^a) + (x^k, y^k, s^k) \quad (3.13)$$

e

$$\sigma = 1 : (x^c, y^c, s^c) = (\Delta x^c, \Delta y^c, \Delta s^c) + (x^k, y^k, s^k). \quad (3.14)$$

Em geral, (x^a, y^a, s^a) é inviável, e pelo teorema 3.1, se $\delta(x, s, \mu) \leq 0.5$ então $\delta(x^c, s^c, \mu) \leq 0.25$.

Com esses novos pontos, vamos caminhar ao longo de uma semi-reta que liga x^a e x^c , a saber

$$r = \sigma x^c + (1 - \sigma)x^a \quad (3.15)$$

com $\sigma \in [0, 1]$, sem ultrapassarmos uma vizinhança da trajetória central \mathcal{C} , entretanto, desejamos andar o máximo possível nessa região.

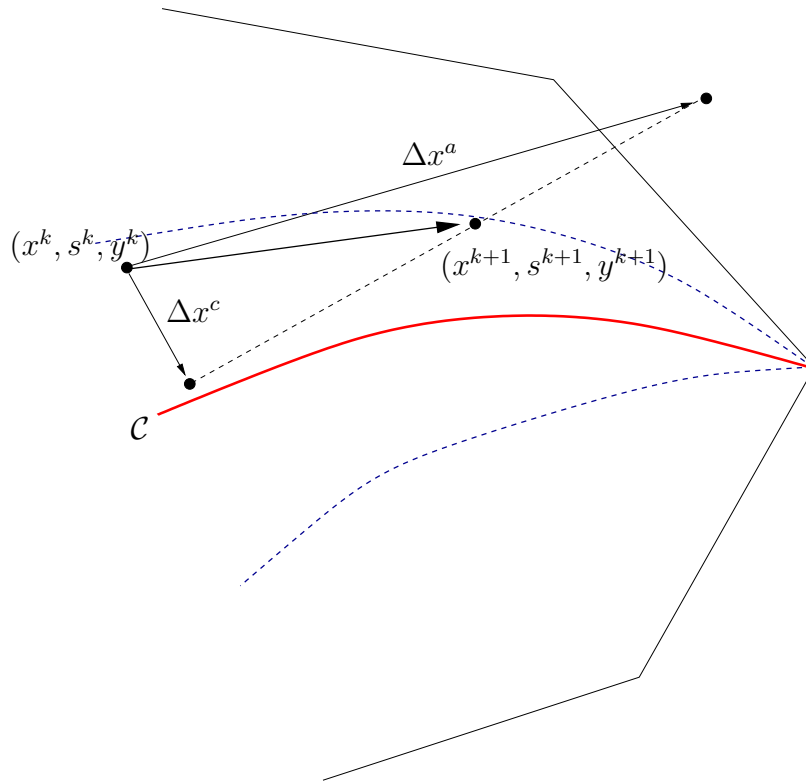


Figura 3.1: Passo mais longo

Definamos o conjunto viável \mathcal{V} e o conjunto estritamente viável \mathcal{V}° da seguinte forma

$$\mathcal{V} = \{(x, s) \mid Ax = b, x \geq 0, s \geq 0\} \quad (3.16)$$

$$\mathcal{V}^o = \{(x, s) \mid Ax = b, x > 0, s > 0\} \quad (3.17)$$

As duas vizinhanças de \mathcal{C} mais interessantes encontradas na literatura são

$$N_2(\theta) = \left\{ (x, s) \in \mathcal{V}^o \mid \left\| \frac{x \cdot s}{\mu} - e \right\|_2 \leq \theta \right\} \quad (3.18)$$

para algum $\theta \in (0, 1)$ e

$$N_{-\infty}(\gamma) = \{(x, s) \in \mathcal{V}^o \mid x \cdot s \geq \gamma e\} \quad (3.19)$$

para algum $\gamma \in (0, 1)$.

Valores típicos para os parâmetros são $\theta = 0.5$ e $\gamma = 10^{-3}$. É possível mostrar que $N_2 \subset N_{-\infty}$. Em geral, $N_{-\infty}$ é muito maior que N_2 , mas a teoria de convergência para $N_{-\infty}$ é muito mais complicada.

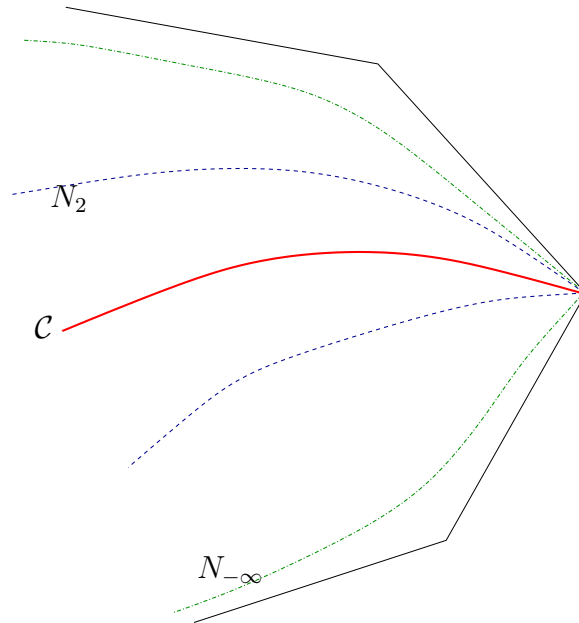


Figura 3.2: Comparação das vizinhanças da trajetória central

Vistas estas novas definições, lembramos que nosso objetivo é encontrar o

parâmetro σ em (3.15) tal que andamos o máximo possível dentro de uma vizinhança.

Vamos trabalhar inicialmente com a vizinhança N_2 : resolveremos a equação

$$\left\| \frac{(\sigma x^c + (1 - \sigma)x^a) \cdot (\sigma s^c + (1 - \sigma)s^a)}{\sigma \mu} - e \right\|_2 = \theta. \quad (3.20)$$

Ao resolver a equação acima, que é uma equação de quarto grau, escolhemos $\sigma \in [0, 1)$ real: caso tenha mais de uma raiz em $(0, 1)$ tomamos a menor. Se não existir raiz real, o passo de Newton completo com $\sigma = 0$ é aceito e o problema é fácil.

Os novos iterandos serão dados por

$$x = \sigma x^c + (1 - \sigma)x^a \quad (3.21)$$

$$s = \sigma s^c + (1 - \sigma)s^a \quad (3.22)$$

$$y = \sigma y^c + (1 - \sigma)y^a \quad (3.23)$$

$$\mu = \sigma \mu. \quad (3.24)$$

Feitas tais atualizações reiniciamos o processo e assim geramos uma sequência de pontos que estão próximos da trajetória central até obter a convergência requerida, convergindo para um ponto solução do problema (3.1) com μ menor que uma tolerância dada, isto é, muito próximo de zero e assim (3.6) se aproxima de (3.4).

Capítulo 4

Resultados principais

Até o presente momento reunimos informações, descritas de uma maneira geral para o problema indicado em cada seção, necessárias para dar continuidade ao nosso trabalho. Neste capítulo, o principal do trabalho, vamos conectar todas as informações anteriores fazendo alguma modificação, se necessário, para adequarmos ao nosso problema.

Estamos interessados em resolver o problema (1.2) exposto no capítulo 1, por sua importância vamos reescrevê-lo aqui,

$$\begin{aligned} & \text{minimizar} && f(x) \\ & \text{sujeito a} && c(x) = 0 \\ & && x \geq 0 \end{aligned} \tag{4.1}$$

adicionando a restrição de positividade, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe C^2 e $c : \mathbb{R}^n \rightarrow \mathbb{R}^m$, com $m < n$, tal que

$$x \mapsto c(x) = [c_1(x), c_2(x), \dots, c_m(x)]^T$$

com $c_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$ funções de classe C^2 .

Resolveremos esse problema usando o método descrito no capítulo 2, método de restauração inexata. Em resumo, cada iteração é caracterizada pela substituição

da função por seu modelo quadrático e das restrições de igualdade por sua linearização, resolvendo os dois passos separadamente. No primeiro, passo normal, encontramos um ponto intermediário, em seguida construímos uma região de confiança nesse ponto e encontramos um novo ponto, minimizando o modelo quadrático do Lagrangeano no núcleo da matriz Jacobiana, o que caracteriza o passo tangente.

4.1 O passo normal

Manteremos sempre a restrição $x \geq 0$ satisfeita. A viabilidade é medida como definida em (2.6), por $h(x) = \|c(x)\|$, mas no passo normal devemos reduzir $h(\cdot)$ com a restrição $x \geq 0$. Esse passo foi estudado por Francisco [7] e não detalharemos aqui. Vamos estudar o passo tangente.

4.2 O passo tangente

O passo tangente parte de um ponto $z \in \mathbb{R}^n$ satisfazendo

$$z > 0 \text{ e } z \notin \mathcal{F}$$

em que \mathcal{F} é a região proibida pelo filtro na iteração k . Ao ponto z associa-se um estimador de multiplicador de Lagrange para as restrições de igualdade $\lambda \in \mathbb{R}^m$.

O modelo quadrático do Lagrangeano em torno de z é construído como no capítulo 2, e tem o formato

$$m_{\mathcal{L}}(x, \lambda) = v^T x + \frac{1}{2} x^T H x, \tag{4.2}$$

em que H é a hessiana do Lagrangeano $\mathcal{L}(\cdot, \lambda)$ em torno de z .

4.2.1 Região de Confiança

Dado um raio Δ para a região de confiança, a minimização de $m_{\mathcal{L}}$ deve ser restrita à caixa $\{x \in \mathbb{R}^n \mid \|x - z\|_{\infty} \leq \Delta\}$. Mas existem também as restrições $x \geq 0$. Na tese da Ana Maria faz-se a seguinte construção:

Seja Pv a projeção de v sobre $N(A(z))$ (que coincide com o gradiente projetado de $f(\cdot)$).

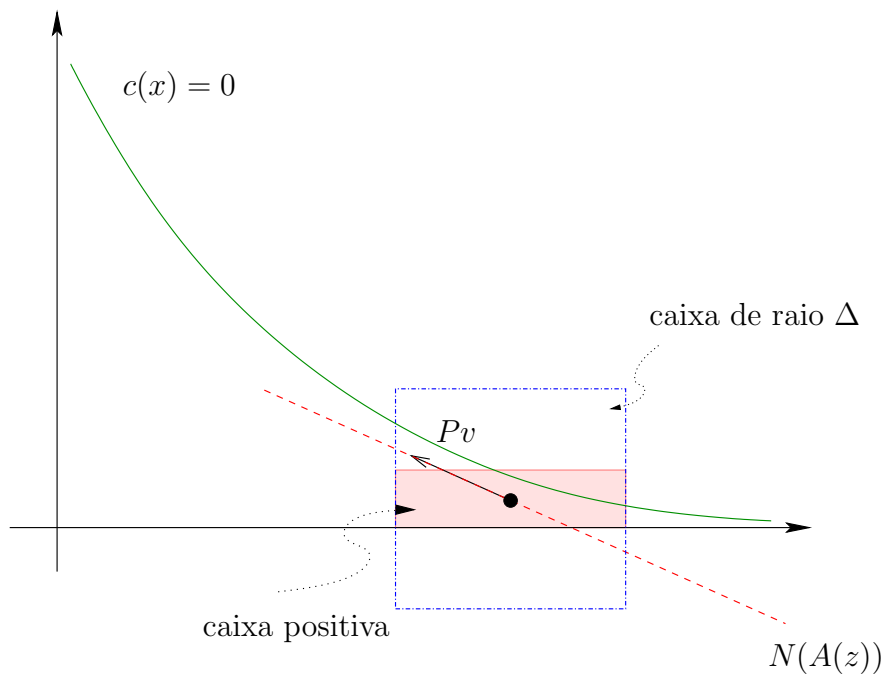


Figura 4.1: Construção da Região de confiança para o sub-problema tangencial

O procedimento segue no algoritmo 2.

Algoritmo 2 Construção da caixa

Dados: $z \in \mathbb{R}^n$, $Pv \in \mathbb{R}^n$, $\Delta > 0$

- 1: **para** $k = 0, 1, 2, \dots, n$ **faça**
 - 2: **se** $Pv_i \geq 0$ **então**
 - 3: $l_i = 0$
 - 4: $u_i = 2z_i$
 - 5: **senão**
 - 6: $l_i = z_i - \Delta$
 - 7: $u_i = z_i + \Delta$
 - 8: **fim se**
 - 9: **fim para**
-

A região de confiança será definida por

$$l \leq x \leq u.$$

Na construção da região desprezamos as restrições $x_i \geq 0$ para as coordenadas i tais que $Pv_i < 0$.

4.2.2 O conjunto Ω

O conjunto Ω é definido pelas restrições de filtro e por $x \geq 0$. Esta última restrição é necessária porque pode ter sido ignorada na construção da região de confiança para algumas componentes.

$$\Omega = \{x \in \mathbb{R}^n \mid x \notin \mathcal{F} \text{ e } x \geq 0\}. \quad (4.3)$$

O problema tangencial é definido por

$$\begin{aligned} \text{minimizar} \quad & v^T x + \frac{1}{2} x^T H x \\ \text{sujeito a} \quad & A(z)(x - z) = 0 \\ & l \leq x \leq u \\ & x \in \Omega. \end{aligned} \quad (4.4)$$

Definindo $A = A(z)$ e $b = Az$, a restrição de igualdade é $Ax = b$ e assim chegamos ao formato apresentado na introdução, como em (3).

Vamos organizar a resolução desse problema no algoritmo 4, desenvolvido em detalhes na tese da Ana Maria. Primeiramente vamos mostrar como o método tradicional é feito, no algoritmo 3.

Para ambos vamos definir o seguinte problema:

$$\begin{aligned}
& \text{minimizar} && v^T x + \frac{1}{2} x^T H x \\
& \text{sujeito a} && Ax = b \\
& && l \leq x \leq u
\end{aligned} \tag{4.5}$$

Algoritmo 3 Método tradicional

Dados: $z \in \mathbb{R}^n$, $\Delta > 0$, $\alpha \in (0, 1)$, Ω , μ_0

- 1: **para** $k = 0, 1, 2, \dots$ **faça**
 - 2: calcular l e u pelo algoritmo 2
 - 3: encontrar w solução de (4.5)
 - 4: calcular $ared = \mathcal{L}(z, \lambda) - \mathcal{L}(w, \lambda)$
 - 5: $pred = m_{\mathcal{L}}(z) - m_{\mathcal{L}}(w)$
 - 6: **se** $ared/pred \geq \alpha$ e $w \in \Omega$ **então**
 - 7: aceitar o passo, com $x^{k+1} = w$
 - 8: **senão**
 - 9: $\Delta = \Delta/2$
 - 10: **fim se**
 - 11: **fim para**
 - 12: até aceitar o passo
-

Algoritmo 4 Método novo

Dados: $z \in \mathbb{R}^n$, $\Delta > 0$, $\alpha \in (0, 1)$, Ω , μ_0

- 1: calcular l e u pelo algoritmo 2
 - 2: siga a trajetória central do problema (4.5) para encontrar um ponto (aproximadamente) central w_μ tal que:
 - 3: $w_\mu \in \Omega$
 - 4: $ared/pred \geq \alpha$
 - 5: μ é pequeno
 - 6: para a iteração seguinte, faça $\Delta = \min\{\Delta, 2\|w - z\|_\infty\}$
-

Observação

- A trajetória central para o problema primal (4.5) é a parte primal da trajetória primal-dual correspondente, que será formalizada na próxima seção.
- O valor “pequeno” de μ é obtido usando o algoritmo de trajetória central, reduzindo μ enquanto $w_\mu \in \Omega$ e o critério de redução efetiva (confiança no modelo quadrático).
- O ponto primal-dual inicial para a trajetória resulta da tese de Roger Behling

[1], e corresponde a um passo de região de confiança com região elipsoidal em torno de z .

4.3 Transformação do Problema

Convém observar que no passo tangente de PQS em (2.14), estamos pensando na norma infinito, ou melhor, tal região é definida por uma caixa

$$l \leq x \leq u \tag{4.6}$$

para l e u vetores com dimensões apropriadas que representam um limitante inferior e superior, respectivamente.

Neste passo o problema a ser resolvido, de forma geral, é

$$\begin{aligned} &\text{minimizar} && v^T x + \frac{1}{2} x^T H x \\ &\text{sujeito a} && Ax = b \\ &&& l \leq x \leq u \end{aligned} \tag{4.7}$$

com v e H o vetor gradiente e a matriz hessiana, respectivamente, da função indicada aplicados no ponto corrente.

Chamamos atenção quanto à notação, aqui usaremos a variável x no lugar da variável d do problema (2.14). Ainda, sem fazer confusão, estamos usando a mesma notação do capítulo 3 para o modelo quadrático, ao invés de usar (2.11).

Adicionando variáveis de folga no problema (4.7) podemos transformá-lo no problema padrão (3.1) e usar a técnica ali desenvolvida.

Subtraindo l em (4.6) resulta

$$0 \leq x - l \leq u - l$$

que por sua vez, definindo $x^{(1)} := x - l$ e $b^{(2)} := u - l$, temos $x^{(1)} \leq b^{(2)}$.

Adicionando uma variável de folga $x^{(2)}$, temos a igualdade

$$x^{(1)} + x^{(2)} = b^{(2)} \quad (4.8)$$

e da expressão $Ax = b$, temos $A(x^{(1)} + l) = b$, ou ainda, $Ax^{(1)} = b - Al$, fazendo $b^{(1)} := b - Al$, segue

$$Ax^{(1)} = b^{(1)} \quad (4.9)$$

Combinando (4.8) e (4.9), segue

$$\begin{bmatrix} A & 0 \\ I & I \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix} = \begin{bmatrix} b^{(1)} \\ b^{(2)} \end{bmatrix} \quad (4.10)$$

Por outro lado, $v^T x + \frac{1}{2}x^T H x$ deve ser reescrito como

$$\begin{bmatrix} \tilde{v}^T & 0 \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} (x^{(1)})^T & (x^{(2)})^T \end{bmatrix} \begin{bmatrix} H & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix} \quad (4.11)$$

em que 0 denota a matriz nula e I a matriz identidade, com as dimensões devidamente atribuídas e ainda $\tilde{v} = (v + Hl)$.

Fazendo $\bar{H} = \begin{bmatrix} H & 0 \\ 0 & 0 \end{bmatrix}$, $\bar{A} = \begin{bmatrix} A & 0 \\ I & I \end{bmatrix}$, $\bar{x} = \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix}$, $\bar{v} = \begin{bmatrix} \tilde{v} \\ 0 \end{bmatrix}$ e

$$\bar{b} = \begin{bmatrix} b^{(1)} \\ b^{(2)} \end{bmatrix} \text{ obtemos}$$

$$\begin{aligned} &\text{minimizar} && \bar{v}^T \bar{x} + \frac{1}{2} \bar{x}^T \bar{H} \bar{x} \\ &\text{sujeito a} && \bar{A} \bar{x} = \bar{b} \\ &&& \bar{x} \geq 0 \end{aligned} \quad (4.12)$$

com $\bar{v}, \bar{x} \in \mathbb{R}^{2n}$, $\bar{H} \in \mathbb{R}^{2n \times 2n}$, $\bar{A} \in \mathbb{R}^{(m+n) \times 2n}$ e $\bar{b} \in \mathbb{R}^{m+n}$, conforme (3.1).

Nosso problema de minimização no espaço tangente (4.7) foi transformado no novo problema (4.12), com um custo de aumentar a dimensão do problema original. Agora podemos resolvê-lo combinando o procedimento desenvolvido no capítulo 3 com os devidos cuidados da região de confiança, seção 1.4.

Capítulo 5

Implementação

Neste capítulo descreveremos alguns algoritmos implementados em Matlab a fim de colocar em prática a teoria até aqui estudada. Nosso principal objetivo era resolver o sub-problema tangencial que aparece no segundo passo de algoritmo de filtros para PQS. A idéia desse algoritmo, que leva o título de nosso trabalho, segue no algoritmo conceitual 4.

Aqui foi implementado o algoritmo de filtro para PQS com restauração inexata, o que consiste em gerar um passo de restauração e o passo tangencial.

5.1 Pré-requisitos

Para iniciar nosso trabalho prático estudamos a função Lagrangeno associada ao problema com restrições de igualdade e construímos a função com o cálculo do gradiente e da hessiana para o ponto dado.

A função de inviabilidade aqui foi definida com uma modificação, a saber

$$h(x) = \frac{1}{2} \|c(x)\|^2, \quad (5.1)$$

é uma função de classe C^2 . Novamente, construímos o gradiente e a hessiana para um ponto dado.

Antes de iniciar o algoritmo principal, construímos uma entrada para o filtro e uma rotina para testar se o ponto está na região proibida ou não.

5.2 Restauração

Esse passo não era o nosso objetivo, como dito antes, e programamos a rotina a fim de gerar um ponto intermediário z para então resolver o passo tangencial. Como em cada iteração exigimos $x^k > 0$, também não aceitamos o ponto z se violar a restrição dada por Ω .

Iniciamos com o método de Newton, sabemos que ele funciona bem quando está próximo da solução. Assim, se o passo dado pelo método Newton for viável, aceitamos; caso contrário, damos um passo de Cauchy com mudança de escala, idéia dada pelo método de Dikin, teoria estudada em programação linear (maiores detalhes ver [7]). Ainda nesse passo tivemos que usar Armijo para obter o decréscimo suficiente quando usamos a direção de Cauchy (para os métodos Armijo e Cauchy ver [20]).

5.3 Otimização

A rotina para este passo foi a mais extensa, reunindo as idéias do método de pontos interiores a fim de caminhar numa região próxima da trajetória central e de região de confiança, uma vez que estamos trabalhando com o modelo quadrático do Lagrangeano.

Iniciamos nossa rotina construindo a caixa dada pelo algoritmo 2, transformamos o problema como visto na seção 4.3 aumentando a dimensão do problema. Partimos de um ponto com componentes positivas e próximo do centro analítico da caixa, esse foi dado pelo algoritmo da tese de Behling [1] e programado aqui.

Em cada iteração desse passo encontramos a direção de centralização e a direção afim-escala para combiná-las de maneira a caminhar o máximo possível dentro da vizinhança, que no nosso caso foi N_2 . A iteração é aceita caso o modelo quadrático

do Lagrangeano seja confiável, o ponto gerado não pertença a região proibida pelo filtro \mathcal{F} e seja não negativo. Caso uma delas falhe, paramos e ficamos com o ponto anterior que já respeitava as três condições. Em particular se o modelo quadrático é ruim, reduzimos o tamanho do raio Δ para a próxima iteração. Seguindo a idéia de região de confiança, caso atingimos o final da trajetória central e o modelo é muito confiável, aumentamos o tamanho da nossa região.

5.4 Exemplo em \mathbb{R}^2

Vamos testar nosso algoritmo para resolver um exemplo em \mathbb{R}^2 , o que facilita nossa visualização.

Exemplo: Uma função linear sujeita a uma bola de raio 1. Iniciamos com um ponto no octante positivo $x_0 = (1, 2)$, tolerância $\varepsilon = 10^{-6}$ e $\Delta = 0.5$. Em 4 iterações encontramos como solução aproximada $(1.0007, 0.0096)$, ver as figuras 5.1 e 5.2

$$\begin{aligned} &\text{minimizar } y \\ &\text{sujeito a } x^2 + y^2 - 1 = 0 \\ &\quad (x, y) \geq 0 \end{aligned} \tag{5.2}$$

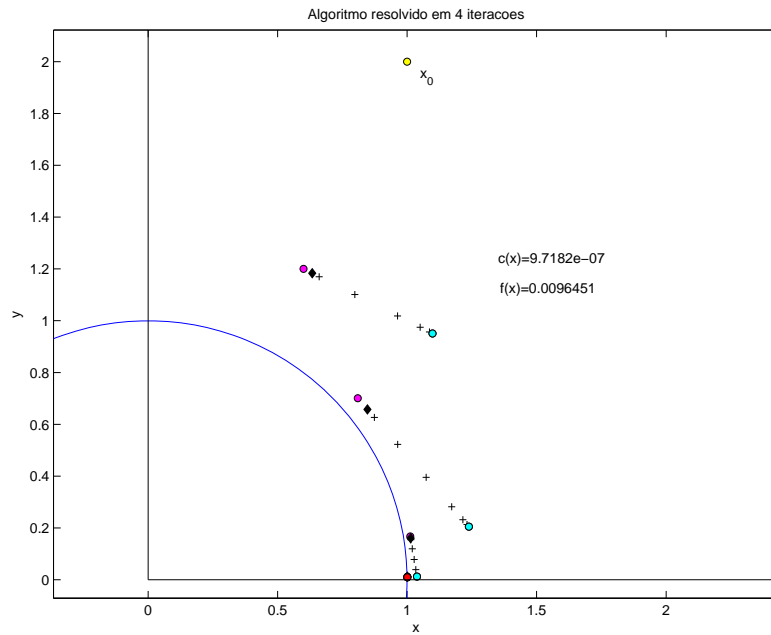


Figura 5.1: Solução aproximada do problema (5.2)

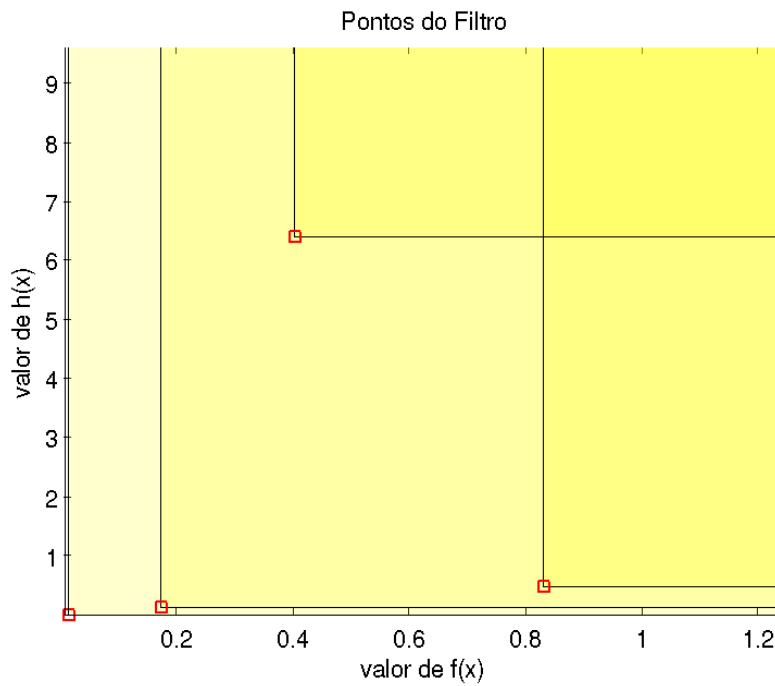


Figura 5.2: Região proibida pelo filtro

Conclusão

No presente trabalho procuramos relembrar alguns conceitos e revisar métodos necessários para a solução do nosso problema, tais como região de confiança e trajetória primal-dual, com o intuito de usá-los para a elaboração de um novo algoritmo para a solução do sub-problema tangencial, problema proveniente do segundo passo do algoritmo de PQS com métodos de filtro, o que completou nosso objetivo para este trabalho.

O mais empolgante, foi acompanhar o desenvolvimento da teoria elaborada em conjunto com o professor Clóvis Gonzaga e a aluna Ana Maria Basei, o que possibilitou a troca de idéias e informações importantes para essa linha de pesquisa, unindo ferramentas estudadas em programação linear e não linear.

Por fim, conseguimos fixar ainda mais as idéias e conceitos ao programar os métodos citados no capítulo 5, aferindo com a teoria explanada nos capítulos iniciais. Em alguns exemplos, embora simples, percebemos o bom funcionamento e convergência para tal método.

Sem dúvida nenhuma este trabalho abre uma porta na área de otimização, o que me motiva a seguir esse caminho dentro da matemática aplicada, dando continuidade para esse estudo que está longe de se exaurir aqui.

Referências Bibliográficas

- [1] BEHLING, Roger. *Minimização de quadráticas em caixas sobre variedades afins, um subproblema de PQS*. Dissertação de mestrado. Universidade Federal de Santa Catarina, Santa Catarina, Brazil, 2006.
- [2] BELLAVIAS, S.; MACCONI, M; MORINI, B. *An affine scaling trust-region method approach to bound-constrained nonlinear systems*. Applied Numerical Mathematics, vol 44: 257-280, 2003.
- [3] CASALI, Rafael. *Resolução de problemas de programação matemática com restrição de equilíbrio aplicado a logística e transporte*. Technical report, Universidade Federal de Santa Catarina, 2006.
- [4] COLEMAN, T. F., LI, Y. *An interior trust region approach for nonlinear minimization subject to bounds*. SIAM Journal on Optimization, vol 6: 418-445, 1996.
- [5] DIRKSE, S., FERRIS, M. and MEERAUS, A. *Mathematical programs with equilibrium constraints: Automatic reformulation and solution via constrained optimization*. Technical Report NA-02/11, Oxford University Computing Laboratory, jul 2002.
- [6] FERRIS, M. , WATHEN, A. and ARMAND, P. *Limited memory solution of complementarity problems arising in video games*. Technical report, Oxford University, 2006.

- [7] FRANCISCO, Juliano B. *Viabilidade em Programação Não-linear: Restauração e Aplicações*. Tese de doutorado. Unicamp, São Paulo: 2005.
- [8] FLETCHER, R. and LEYFFER, S. *Nonlinear programming without a penalty function*. Technical Report NA/171, Dundee University, Dept. of Mathematics, 1997. Revised 2000.
- [9] GILBERT, J. Charles, GONZAGA, Clovis C. and KARAS, Elizabeth. *Examples of ill-behaved central paths in convex optimization*. Mathematical Programming, 103(1): 63-94, May 2005.
- [10] GONZAGA, Clovis C. *Path following methods for linear programming*. SIAM Review, 34(2); 167-227, 1992.
- [11] GONZAGA, Clovis C. *The largest step path following algorithm for monotone linear complementarity problems*. Mathematical Programming, 76: 309-332, 1997.
- [12] GONZAGA, Clovis C. , KARAS, Elizabeth and VANTI, M. *A globally convergent filter method for nonlinear programming*. SIAM Journal on Optimization, 14(3): 646-669, 2003.
- [13] IZMAILOV, A. and SOLODOV, M. *Otimização - volume 1: Condições de Otimidade, Elementos de Análise Convexa e de Dualidade*. Rio de Janeiro: IMPA, 2005.
- [14] KARAS, Elizabeth W. *Exemplos de trajetória central mal comportada em otimização convexa e um algoritmo de filtro para programação não linear*. Tese de doutorado. Universidade Federal de Santa Catarina, Santa Catarina, Brazil, 2002.
- [15] LEYFFER, Sven. *Complementarity constraints as nonlinear equations: Theory and numerical experience*. Technical report, Argonne National Laboratory, Mathematics and Computer Science Division, 2003.

- [16] MARTÍNEZ, J. M. *Box-quacan and the implementation of augmented lagrangian algorithms for minimization with inequality constraints*. Computational & Applied Mathematics, 19: 31-56, 2000.
- [17] MARTÍNEZ, J. M.; SANTOS, S. A. *Métodos Computacionais de Otimização*. IMPA: 1995.
- [18] MARTÍNEZ, J. M.; PILOTA, E. A. *A Inexact restoration algorithm for constrained optimization*. Journal of Optimization Theory and Applications, 111 : 39-58, 2001.
- [19] MOREÉ, J.J.; SORENSEN, D.C. *Computing a trust region step*. SIAM Journal Scientific Computing, 4: 553-572, 1983.
- [20] NOCEDAL, J. and WRIGHT, S. J. *Numerical Optimization*. Springer Series in Operations Research. Springer-Verlag, 1999.
- [21] SCHEEL, Holger and SCHOLTES, Stefan. *Mathematical programs with complementarity constraints: stationarity, optimality and sensitivity*. Mathematics to Operations Research, 25(1), February 2000.
- [22] VANTI, Márcia R. V. *Melhoria da segurança dinâmica por otimização e algoritmos de programação não linear*. Tese de doutorado. Universidade Federal de Santa Catarina, Santa Catarina, Brazil, 2003.
- [23] WRIGHT, Stephen J. *Primal-Dual Interior-Point Methods*. Philadelphia, SIAM: 1997.