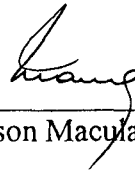


# CONTRIBUIÇÃO AO PROBLEMA DE VIABILIDADE CONVEXA

Yrma Alejandrina Raymundo Huaroto

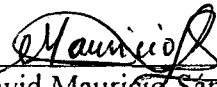
TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS À OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS (D. Sc.) EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Aprovada por:



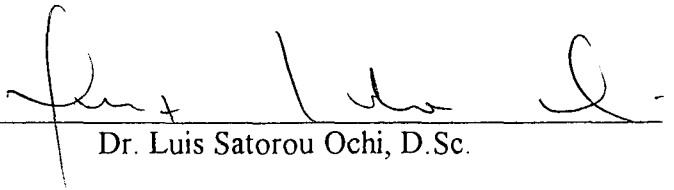
---

Prof. Nelson Maculan Filho, D.Sc.



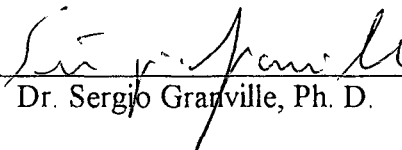
---

Prof. David Mauricio Sánchez, D.Sc.



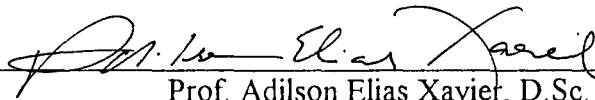
---

Dr. Luis Satorou Ochi, D.Sc.



---

Dr. Sergio Granville, Ph. D.



---

Prof. Adilson Elias Xavier, D.Sc.

RIO DE JANEIRO, RJ - BRASIL  
JUNHO DE 1997

RAYMUNDO HUAROTO, YRMA  
ALEJANDRINA

Contribuição ao Problema de  
Viabilidade Convexa [Rio de Janeiro] 1997

VIII,113 p. 29,7 cm (COPPE/UFRJ,  
D.Sc.,Engenharia de Sistemas, 1997)

Tese - Universidade Federal de Rio de  
Janeiro, COPPE

1. Viabilidade Convexa-Planos de Corte

I.COPPE/UFRJ      II. Título(série)

## **Dedicatória**

À memória dos meus pais.  
A minha querida família no Perú.  
Em especial a meus irmãos e sobrinhos.

## **Agradecimentos:**

Ao Dr. Nelson Maculan Filho, meu orientador, por seu valioso estímulo e apoio na realização do presente trabalho.

Ao Dr. Clovis C. Gonzaga por seus valiosos ensinamentos que contribuíram na realização desta pesquisa.

A meu amigo Dr. David Mauricio que me deu muito estímulo para continuar na elaboração da tese e pelas úteis sugestões.

A todos meus amigos que de um modo ou outro mostraram seu estímulo e apoio. Em especial a meu amigo Freddy Poma que me prestou seu valioso apoio na digitação da tese.

Finalmente agradeço o apoio financeiro do CNPq recebido durante vários anos de estudo no Programa de Engenharia de Sistemas e Computação.

Finalmente agradeço a Universidade Estadual do Norte Fluminense (UENF) pelo apoio financeiro prestado para a conclusão da tese.

# SUMÁRIO

## CAPÍTULO 1- INTRODUÇÃO

1.1 MÉTODO DE PONTOS INTERIORES .....	01
1.2 MÉTODO DE PLANOS DE CORTE .....	02
1.3 O PROBLEMA DE VIABILIDADE CONVEXA .....	04
1.4 ALCANCE DA TESE .....	05

## CAPÍTULO 2 - REDUNDÂNCIA EM UM POLÍTOPO

2.1 INTRODUÇÃO .....	07
2.2 POLIEDROS .....	08
2.3 REPRESENTAÇÃO DE UM POLIEDRO .....	12
2.4 DESCRIÇÕES IRREDUNDANTES DE UM POLIEDRO .....	13
2.5 MÉTODO PRÁTICO PARA IDENTIFICAR ALGUMAS RESTRIÇÕES REDUNDANTES .....	17

## CAPÍTULO 3-O PROBLEMA DE VIABILIDADE CONVEXA

3.1. INTRODUÇÃO .....	19
3.2. NOTAS HISTÓRICAS .....	20
3.3. PRELIMINARES.....	25
3.3.1. FUNÇÃO BARREIRA LOGARÍTMICA .....	25
3.3.2. ALGORITMO PADRÃO DE PLANOS DE CORTE CENTRO ANALÍTICO .....	26
3.3.3. PROPRIEDADES DO ELIPSOIDE HESSIANO .....	27
3.3.4. VALORES DA FUNÇÃO BARREIRA LOGARÍTMICA PERTO DO CENTRO ANALÍTICO .....	34
3.3.5. MEDIDA DE PROXIMIDADE AO CENTRO ANALÍTICO . .....	44

## CAPÍTULO 4- O MÉTODO CENTRO ANALÍTICO COM ELIMINAÇÃO DE CORTES

4.1. INTRODUÇÃO .....	47
4.2. ALGORITMOS DE PLANOS DE CORTE-CENTRO ANALÍTICO COM ELIMINAÇÃO DE CORTES .....	47
4.3. ALGORITMO DE ATKINSON-VAIDYA .....	48
4.4. O MÉTODO DE ELIPSOIDES .....	51

**CAPÍTULO 5- UM ALGORITMO DE PLANO DE CORTE-CENTRO  
ANALÍTICO COM ELIMINAÇÃO DE CORTES  
REDUNDANTES**

5.1.	INTRODUÇÃO .....	55
5.2.	UM CRITERIO DE IDENTIFICAÇÃO DE RESTRIÇÕES REDUNDANTES NUM POLÍTOPO . .....	55
5.3.	O ALGORITMO (A.E.R) .....	58
5.4.	DECRECIMENTO DO VOLUME DOS ELIPSOIDES USADOS NO ALGORITMO(R,R).....	61
5.5.	POLINOMIALIDADE DO ALGORITMO (AER) .....	67

**CAPÍTULO 6- UMA APLICAÇÃO AO PROBLEMA DE PROGRAMAÇÃO  
CONVEXA.**

6.1	INTRODUÇÃO .....	70
6.1.1	ALGORITMO DE GOFFIN-LUO-YE PARA O PROBLEMA DE VIABILIDADE CONVEXA .....	70
6.1.2	ALGORITMO DE KIWIEL PARA O PROBLEMA DE PROGRAMAÇÃO CONVEXA.....	73
6.2.0	O PROBLEMA CONVEXO.....	76
6.3	ALGUMAS DEFINIÇÕES.....	76
6.4.	ALGORITMO (PV $\epsilon$ ) PARA PROGRAMAÇÃO CONVEXA RESTRITA.....	78
6.5	REDUÇÃO DO MÁX-POTENCIAL.....	82
6.6	COMPLEXIDADE DO ALGORITMO (PV $\epsilon$ ).....	92
6.7	UM ALGORITMO PARA O PROBLEMA DE PROGRAMAÇÃO CONVEXA IRRESTRITA.....	106
<b>CAPÍTULO 7 - CONCLUSÕES</b>		
7.1	CONCLUSÕES.....	108
<b>BIBLIOGRAFIA.....</b>		<b>109</b>

# CAPÍTULO 1

## INTRODUÇÃO

Neste capítulo, descrevemos e motivamos o objetivo da tese. Inicialmente é feita uma breve descrição da história do método de pontos interiores. Em seguida descrevemos brevemente o método de planos de corte e também falamos sobre o problema de viabilidade convexa. Finalmente damos a conhecer o conteúdo e a contribuição da tese.

### 1.1 MÉTODO DE PONTOS INTERIORES (MPI)

A pesquisa sobre o método de pontos interiores para programação convexa surge ao redor dos anos 70. A primeira onda começou em 1979 com a publicação de Khachian sobre o método de elipsóides [25]. Interessante por ser o primeiro método polinomial para programação linear (PL), bem como e pelas muitas conseqüências teóricas, entretanto muito lento na prática. Khachian usou algumas técnicas de programação não linear (P.N.L) desenvolvido por Shor [35] e desenvolvido por Yudin-Nemirovsky [44]. Ao contrário do método Simplex o método de elipsóides não usa a estrutura combinatorial de PL.

Uma segunda onda de interesse foi iniciada em 1984, com a publicação do algoritmo polinomial de Karmakar [26], para PL. Karmakar conseguiu resolver problemas de grande escala para PL, com uma maior rapidez do que o método simplex. Atualmente torna-se claro que o trabalho de Karmakar tem dado abertura a um novo campo de pesquisa, o campo dos métodos de pontos interiores. Inspirados na publicação de Karmakar muitos outros métodos polinomiais foram propostos, alguns contendo novas idéias, outros contendo somente poucas modificações.

O algoritmo de Karmakar requer de  $O((m^{1.5}n^2+m^2n)L)$  operações aritméticas, onde “m” é o número de hiperplanos definindo o sistema linear, “n” a dimensão do espaço e “L” o comprimento da entrada. Após a publicação de

Karmakar, o método original projetivo de redução de potencial de Karmakar, foi estudado e modificado por muitos pesquisadores que simplificaram a análise, melhoraram o tempo total de complexidade, diminuíram as suposições iniciais, etc.

Outros autores tem estendido idéias de pontos interiores para programação não linear. Também, o método de pontos interiores conduz ao renascimento de alguns métodos antigos como por exemplo:

- O método da função barreira logarítmica de Frisch [14];
- O método de centros de Huard (Center method)[24]; e
- O método afim escala de Dikin [9].

Após a publicação de Karmakar, os pesquisadores tornaram-se interessados no clássico método barreira logarítmica e no método dos centros. Ambos métodos são chamados de métodos seguindo trajetória, pois eles seguem a trajetória central do problema.

Gonzaga [16] foi o primeiro a mostrar a polinomialidade de ordem  $O(n^3L)$  para uma versão especial do método barreira logarítmica para PL. Se bem que ele mostrou que a complexidade total é  $O(n^3L)$ , na prática o método é muito lento.

O número de artigos sobre MPI é enorme (aproximadamente de 2,000) que podem ser classificadas categorizar nas seguintes quatro categorias (em ordem cronológicos):

- Métodos seguindo trajetória
- Métodos afim escala
- Métodos projetivo de redução potencial
- Métodos afim de redução potencial

As diferenças entre estas categorias é algumas vezes vago. Todos estes métodos contam com algumas noções e conceitos comuns. Também, diferenciam-se pelo fato de ser em algoritmos primais, duais, ou primais-duais.

## 1.2 MÉTODOS DE PLANOS DE CORTE

A idéia de planos de corte parece ser a idéia básica mais natural da maioria dos métodos modernos de otimização não diferenciável. Dado um problema por exemplo:



$$\begin{cases} \text{Minimizar } f(x) \\ x \in S \subset \mathbb{R}^n \end{cases} \quad (1.1)$$

onde  $f$  é uma função convexa e  $S$  é um conjunto convexo aberto, pode-se escolher um ponto  $x^i$  no conjunto  $S$  como um corrente ponto teste, calcular um subgradiente  $g$  da função  $f$  no ponto  $x^i$ , e então, diminuir a região viável do problema inicial fazendo:

$$S := S \cap \{ x \in \mathbb{R}^n : g^T (x^i - x) \geq 0 \}$$

Assim, cada estratégia de encontrar o próximo ponto teste no corrente conjunto viável do problema (1.1) define um método de minimização não diferenciável. Um dos primeiros métodos deste tipo, “o método de centro de gravidade”, foi desenvolvido por A. Levin [30] e D. Newman [32]. Estes autores propuseram o uso do centro de gravidade do corrente conjunto viável como um ponto teste. Naturalmente esta estratégia pode ter somente um valor teórico, pois o problema de encontrar o centro de gravidade é extremamente difícil mesmo para regiões simples.

O próximo método de plano de corte, “O método de elipsóides” desenvolvido por D.Yudin-A.Nemirovsky [44] e por Shor [37] tornaram-se uma sensação na programação matemática. Era provado que o método de elipsóides tem um raio de convergência linear, bem como este raio depende somente do número de variáveis. Até o momento a lista de métodos de planos de corte inclui também os seguintes métodos:

- O método de elipsóides inscritos de Tarasov-Khachian-Erlich [41];
- O método de centros volumétricos de Vaidya [42];
- O método de centros analíticos de G. Sonnevend [38].

Estreitamente relacionado ao esquema dos planos de corte, tem-se os chamados métodos de feixe, que foram desenvolvidos por C. Lemarechal [31].

O primeiro, e ainda provavelmente o mais popular, algoritmo de planos de corte para programação, convexa era derivadas por Kelly [28]. Desde então, vários métodos de plano de corte e variantes foram desenvolvidos. Uma das vantagens destes métodos de plano de corte, é que os subproblemas a serem resolvidos são

problemas de programação linear (PL), e estes PL problemas diferenciam-se ligeiramente do prévio subproblema. Portanto, as eficientes técnicas para PL são aplicáveis. Além disso, nenhuma busca linear é necessária, e perto da linearidade das funções não lineares é bem preservada.

Os inconvenientes são que o tamanho dos PL subproblemas cresce (pode-se por hábil estratégia eliminar restrições). Convexidade é uma exigência essencial para se conseguir hiperplanos suportes válidos. Para conseguir uma boa solução, a geração de um grande número de planos de corte pode ser necessário. Um dos mais sérios inconvenientes é que nenhum ponto viável é gerado pelos subproblemas. Apenas os pontos de acumulação da seqüência gerada de otimal soluções das relaxações são viáveis e otimal para o problema de programação convexa.

Os assim chamados “Central Cutting Plane Methods” de Elzinga-Moore [11] e Goffin-Vial [20] são consideradas completamente eficientes, ver por exemplo Goffin-Haurie-Vial [19] e Bahn-Goffin-Merle [2]. Eles calculam um certo “centro” da relaxação PL. Se o centro é viável, eles adicionam um corte objetivo, se o centro é inviável, então um novo hiperplano separador é gerado. Portanto estes métodos geram pontos viáveis durante o algoritmo.

Os algoritmos de planos de corte centro analítico usam como “centro” o chamado “centro analítico” da relaxação PL. O centro analítico é calculado por uma variante do algoritmo de Karmakar. A diferença das aproximações clássicas, estes algoritmos tendem a transmitir estimável informação aos subproblemas das quais é um centro.

### **1.3 O PROBLEMA DE VIABILIDADE CONVEXA (PVC)**

Dado um conjunto  $D \subset \mathfrak{R}^n$ ,  $D$  convexo, com interior não vazio, o problema de viabilidade convexa consiste em encontrar um ponto em  $D$ .

O problema de viabilidade convexa é um problema bastante geral de programação convexa, pois um problema de programação convexa pode ser considerado como um problema (PVC), como veremos no Capítulo 6.

A maioria dos algoritmos de viabilidade convexa envolvem um oráculo para  $D$ , que aceita como entrada um ponto  $\bar{y} \in \mathfrak{R}^n$  e produz uma saída

dependendo sobre se  $\bar{y} \in D$ . Se  $\bar{y} \in D$  o algoritmo pára, senão o oráculo retorna um hiperplano separador (corte)

$$H = \{ x \in \mathfrak{R}^n : \mathbf{a}^T x \leq \mathbf{a}^T \bar{y} \} \supset D.$$

Assumindo que  $D$  esteja contido numa região inicial limitada  $P$ , gera-se em cada iteração do algoritmo uma nova região  $P := P \cap H$  tal que o novo  $P$  contém  $D$ . Um novo ponto teste  $\bar{y}$  é encontrado na nova região  $P$ , e assim o algoritmo prosegue. Naturalmente a confiança é que as regiões  $P$  convirjam ao conjunto viável  $D$ . Se  $P$  é um polítopo, então o ponto teste  $\bar{y}$  é que geralmente o “centro analítico” do polítopo  $P$ . Os métodos mais populares para o problema de viabilidade convexa são os métodos de planos de corte-centro analítico, que na prática têm demonstrado desempenho superior aos outros métodos, como é demonstrado em [2],[3], [19],[21],[39]. A idéia de usar o centro analítico como ponto teste foi iniciada por SONNEVEND [29]. A razão de usar o centro analítico no método de planos de corte é evidente, pois é facil de calcular e atualizar, e tem desempenhado um rol importante em muitos algoritmos de pontos interiores para PL e PNL.

#### 1.4 ALCANCE DA TESE

A questão mais importante desta tese é estudar o problema de viabilidade convexa (PVC). Sendo que nosso principal propósito, é conseguir eliminar cortes redundantes surgidos nos algoritmos de viabilidade convexa. Em nossa pesquisa consegue-se criar um algoritmo para o (PVC) que identifica algumas restrições redundantes dos polítopos gerados durante o algoritmo. O critério é baseado no método de elipsóides. Em cada iteração construímos um elipsóide  $E(B',b')$  de volume menor que o volume do elipsóide  $E(B,b)$  considerado na iteração anterior, tal que  $E(B',b')$  contenha ao corrente polítopo:

$$P = \{ x \in \mathfrak{R}^n : \mathbf{a}_i^T x \leq b_i, i = 1, \dots, m \}$$

então, a restrição  $\mathbf{a}_i^T x \leq b_i$  é considerado redundante na definição do polítopo  $P$ , se o hiperplano  $\mathbf{a}_i^T x \leq b_i$  não corta ao elipsóide  $E(B',b')$  ou simplesmente é tangente ao elipsóide.

Usando o fato que o volume dos elipsóides construídos é decrescente, mostramos a complexidade polinomialidade de nosso algoritmo de viabilidade convexa com eliminação de cortes redundantes, que apresentamos no capítulo 5 da tese.

No capítulo 2, damos um resumo da teoria de poliedros, para dar uma base teórica ao nosso critério de identificação de restrições redundantes num polítopo.

No capítulo 3, apresentamos e comentamos amplamente sobre o problema de viabilidade convexa (PVC), e damos um resumo histórico dos algoritmos existentes em torno deste problema. Também neste mesmo capítulo estudamos algumas propriedades úteis da função barreira logarítmica o que servirão para definir medidas de proximidade ao centro analítico.

No capítulo 4, comentamos sobre os algoritmos de planos de corte centro analítico, que usam eliminação de cortes. Entre eles comentamos com maior detalhe o algoritmo de Atkinson-Vaidya[1] que usa eliminação de cortes para provar a polinomialidade do algoritmo. Também neste mesmo capítulo, apresentamos o algoritmo do método de elipsóides, que serviu de base para construir nosso critério de eliminação de cortes redundantes.

No capítulo 5, apresentamos a nossa contribuição ao problema de viabilidade Convexa, criando um algoritmo para o (PVC) com um critério de eliminação de cortes redundantes, na qual usamos como pontos testes os centros analíticos dos polítopos gerados e os centros dos elipsóides construídos durante o algoritmo.

No capítulo 6, apresentamos uma outra contribuição ao problema de programação convexa restrita (P.C.R), criando um algoritmo de plano corte-centro analítico, com cortes profundos para encontrar uma  $\epsilon$ -solução aproximada para o problema (P.C.R), baseado fundamentalmente no algoritmo de Goffin-Luo-Ye [21], e no algoritmo de Kiwiel [27]. Também, damos uma estimativa da complexidade do algoritmo. Apresentamos também um algoritmo para programação convexa irrestrita, onde os pontos teste são gerados por um oráculo de viabilidade, e cuja complexidade total depende do algoritmo para o problema de viabilidade convexa usada para gerar os pontos testes.

# CAPÍTULO 2

## REDUNDÂNCIA EM UM POLÍTOPO

### 2.1 INTRODUÇÃO

Neste capítulo, apresentamos primeiro algumas definições e resultados sobre a teoria de poliedros, para estabelecer condições que determinam quando uma restrição é redundante na definição de um polítopo.

Informalmente entendemos por uma restrição redundante, uma restrição que quando é eliminada do sistema não muda a região viável. Em vez do termo ‘redundante’, alguns autores usam outros termos: ‘trivial’, ‘supérfluo’, ‘irrelevante’, ‘não essencial’ etc. Da natureza de redundância, fica claro que restrições redundantes podem ser omitidas do problema.

Restrições redundantes aparecem com freqüência em problemas práticos, quando se trata de dar uma formulação matemática do modelo. Os modeladores podem querer evitar a possibilidade de ter muitas restrições, e ter o estritamente necessário no modelo; pois o modelo com muitas restrições às vezes torna o modelo grande e complicado.

Na solução de um problema de programação matemática, algumas técnicas requerem restrições extras as quais podem causar redundância. Estas técnicas abrangem todos os métodos de planos de corte para programação linear (descomposição de Dantzig-Wolfe e forma dual de Dantzig and Wolfe (1960)), inteira (Gomory [1958]), inteira mista (Benders [1962]) e programação não linear convexa (Kelly [1963]) e todos os métodos ‘branch and bound’. (ver Garfinkel and Nemhauser [1972]). Em programação paramétrica ( Gal [1979]) as restrições redundantes podem se tornar em não redundantes e vice-versa ( Gal [1975]).

São muitos os efeitos desvantajosos causados pela inclusão de restrições redundantes num sistema linear de restrições. Uma das desvantagens é que restrições redundantes requerem armazenar espaço, o qual pode ser crítica se o problema cresce em tamanho.

Identificar e remover restrições redundantes, pode causar algumas simplificações, como por exemplo reduzir o esforço computacional da resolução do problema. Do o ponto de vista teórico a presença de restrições redundantes pode ser favorável mante-las.

Dado um sistema de desigualdades lineares, o reconhecimento de restrições redundantes no sistema é uma tarefa árdua. Existem algoritmos exatos que servem para determinar quando uma restrição é redundante ou não redundante, e desde o ponto de vista teórico está resolvido, mas na prática não é possível reconhecer todas as restrições redundantes de um sistema linear qualquer. Existem métodos práticos que reconhecem pelo menos algumas das restrições redundantes do sistema.

No capítulo 5 apresentamos um critério que criamos para reconhecer algumas restrições redundantes na definição de um polítopo gerado dentro de um algoritmo de planos de corte; na qual usamos elipsóides contendo o polítopo.

Em seguida tratamos sobre a teoria de poliedros. As definições e as demonstrações estão contidas em [15].

## 2.2 POLIEDROS

Notação: para  $x = (x_1, \dots, x_n)$  e  $y = (y_1, \dots, y_n)$  em  $\mathcal{R}^n$ ,  $x \leq y$  se, e somente se,  $x_i \leq y_i$  para cada  $1 \leq i \leq n$

### Conceitos básicos

Um subconjunto  $P \subseteq \mathcal{R}^n$  é chamado um poliedro. Se  $P = \{x \in \mathcal{R}^n : Ax \leq b\}$  para alguma matriz  $A \in \mathcal{R}^{m \times n}$  e algum vetor  $b \in \mathcal{R}^m$ . Se  $P$  é um poliedro limitado (ou seja existem  $\ell, u \in \mathcal{R}^n$  tal que  $\ell \leq x \leq u$  para cada  $x \in P$ ), então  $P$  é chamado um polítopo.

Se  $\alpha \in \mathcal{R}^n$ ,  $\alpha \neq \bar{0}$  e  $\alpha \in \mathcal{R}$ , então:

O poliedro  $\{x \in \mathcal{R}^n : \alpha^T x \leq \alpha\}$ , é chamado um semi-espço.

O conjunto de pontos  $H = \{x \in \mathcal{R}^n : \alpha^T x = \alpha\}$ , é chamado um hiperplano em  $\mathcal{R}^n$ .

O hiperplano  $H$  divide o espaço em dois semi-espacos

$$\{x \in \mathfrak{R}^n : a^T x \leq \alpha\}(\text{fechado}) \quad \text{e} \quad \{x \in \mathfrak{R}^n : -a^T x < -\alpha\}(\text{aberto})$$

Na definição de um poliedro  $P = \{x \in \mathfrak{R}^n : Ax \leq b\}$ , cada uma das linhas  $a_i^T$  da matriz  $A$  conjuntamente com a correspondente componente  $b_i$  do vetor  $\bar{b}$  define um semiespaço  $\{x \in \mathfrak{R}^n : a_i^T x \leq b_i\}$ . Então, a definição de poliedro

pode ser reformulado, como: Um poliedro é a interseção de um número finito de semi-espacos. É trivial mostrar que um poliedro é um conjunto convexo, isto é, se  $P$  é um poliedro então, para  $x^1, x^2 \in P$ , tem-se que

$$\lambda x^1 + (1-\lambda) x^2 \in P \text{ para qualquer } 0 \leq \lambda \leq 1.$$

### Dimensão de um poliedro:

Definição 2.1 Um número finito de vetores (pontos)  $x^1, x^2, \dots, x^k$  de  $\mathfrak{R}^n$  é dito afim-independentes se a única solução de

$$\sum_{i=1}^k \alpha_i x_i = 0 \quad \text{e} \quad \sum_{i=1}^k \alpha_i = 0$$

é  $\alpha_i = 0$  para  $i = 1, \dots, k$ .

Definição 2.2 Um conjunto  $S$ ,  $S \subset \mathfrak{R}^n$  é de dimensão  $m$ , que denotamos por  $\dim(S) = m$ , se o maior número de vetores afim-independentes em  $S$  é  $m + 1$ . Se  $\dim(S) = n$  então é dito que  $S$  é de dimensão plena.

Seja  $P$  o poliedro  $P = \{x \in \mathfrak{R}^n : Ax \leq b\}$  onde  $A \in \mathfrak{R}^{m \times n}$ ,  $b \in \mathfrak{R}^m$ , um ponto  $x \in P$  é chamado um ponto interior de  $P$  se  $Ax < b$ . É mostrado que  $P$  é um poliedro de dimensão plena se, e somente se,  $P$  tem um ponto interior.

### Faces de um poliedro:

Dado um poliedro  $P = \{x \in \mathfrak{R}^n : a_i^T x \leq b_i, i = 1, \dots, m\}$  a questão que discutiremos é identificar quais das desigualdades  $a_i^T x \leq b_i$  são necessárias na descrição de  $P$  e quais podem ser eliminadas. Para isso precisamos de algumas definições que damos a seguir.

Seja  $S \subseteq \mathfrak{R}^n$ ,  $a \in \mathfrak{R}^n$ ,  $\alpha \in \mathfrak{R}$ . Uma inequação  $a^T x \leq \alpha$  é válida em relação a  $S$ , se  $S \subseteq \{x \in \mathfrak{R}^n : a^T x \leq \alpha\}$ . Um hiperplano  $H = \{x \in \mathfrak{R}^n : a^T x = \alpha\}$  é um hiperplano suporte de  $S$ , se  $a^T x \leq \alpha$  é válida em relação a  $S$  e  $S \cap H \neq \emptyset$ .

Portanto, um hiperplano  $H = \{x \in \mathfrak{R}^n : a^T x = \alpha\}$ , é um hiperplano suporte de  $S$ , se, e somente se, o problema

$$\begin{cases} \text{maximize } a^T x \\ x \in S \end{cases} \quad (2.1)$$

tem uma solução ótima  $\bar{x} \in S$  com  $a^T \bar{x} = \alpha$ .

Seja  $P \subseteq \mathfrak{R}^n$  um poliedro. Um conjunto  $F \subseteq P$  é uma face de  $P$ , se existe uma inequação  $a^T x \leq \alpha$  válida em relação a  $P$  tal que

$$F = P \cap \{x \in \mathfrak{R}^n : a^T x = \alpha\},$$

e neste caso dizemos que  $(a, \alpha)$  é uma representação da face  $F$ . Dizemos que  $F$  é uma face própria de  $P$  se  $F \neq P$ , e que  $F$  é uma face não trivial, se  $\emptyset \neq F \neq P$ .

Se  $a^T x \leq \alpha$  é válida em relação a  $P$ , então dizemos que:

$F = P \cap \{x \in \mathfrak{R}^n : a^T x = \alpha\}$ , é uma face induzida ou definida por  $a^T x \leq \alpha$ .

Quando  $F$  é uma face não vazia de um poliedro  $P$  dizemos que a face  $F$  suporta  $P$ . Logo pela caracterização de hiperplanos suportes dada em (2.1) a face  $F$  do poliedro  $P$  com representação  $(a, \alpha)$  suporta  $P$  se, e somente se,

$$\max \{ a^T x : x \in P \} = \alpha. \text{ (ver Fig. 2.1)}$$



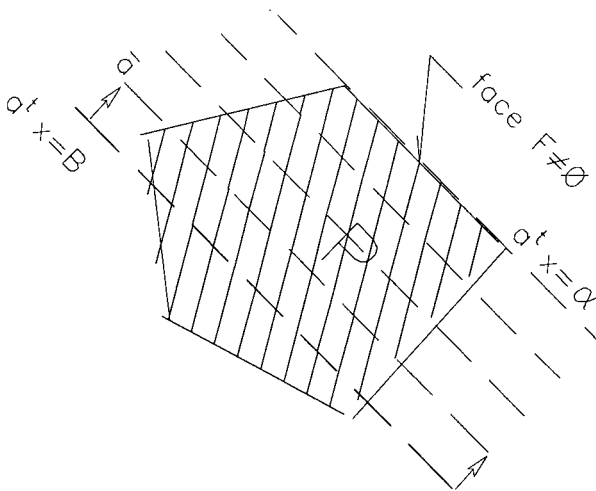


Figura 2.1

Mais ainda,  $F$  é uma face de um poliedro  $P$  se, e somente se existe  $\alpha \in \mathcal{R}^n$  tal que  $F = \operatorname{argmax} \{ \alpha^T x : x \in P \}$ . Portanto, um primeiro passo para descartar desigualdades ‘supérfluas’ na definição de um polítopo

$$P = \{x \in \mathcal{R}^n : Ax \leq b\} = \{x \in \mathcal{R}^n : \alpha_i^T x \leq b_i, i = 1, \dots, m\}$$

é descartar desigualdades  $\alpha_i^T x \leq b_i$  tal que o hiperplano

$$H = \{x \in \mathcal{R}^n : \alpha_i^T x = b_i\}$$

não seja um hiperplano suporte do polítopo  $P$  ou seja quando:

$$F_i = \{x \in \mathcal{R}^n : \alpha_i^T x = b_i\} \cap P = \emptyset \quad (F_i \text{ face vazia}) \quad (\text{ver Fig. 2.2})$$

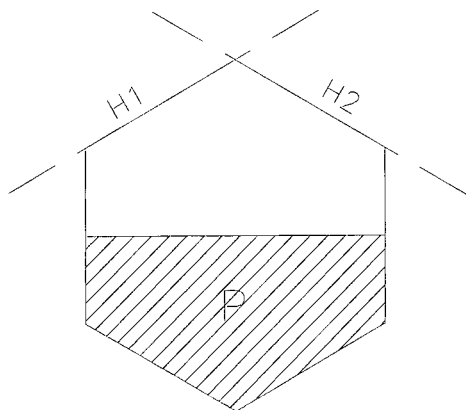


Figura 2.2

$H_1, H_2$  não são suportes

Algumas faces de um poliedro receberão nomes especiais, dependendo de sua dimensão. Lembremos que a dimensão de um conjunto  $F \subset \mathfrak{R}^n$  é igual a  $\dim(F) = d$  se o maior número de vetores afim-independentes em  $F$  é  $d+1$ . É mostrado que se  $F$  é uma face própria de um poliedro  $P$ , então

$$\dim(F) \leq \dim(P)-1$$

Faces de dimensão zero são chamadas vértices de um poliedro.

### Facetas:

As faces próprias maximais de um poliedro são chamadas facetas. Isto é, uma face  $F$  não trivial de um poliedro  $P$  é uma faceta de  $P$ , se  $F$  não está contida em nenhuma face própria de  $P$ .

É mostrado que, uma face  $F$  não trivial de um poliedro  $P$  é uma faceta de  $P$  se, somente se  $\dim(F) = \dim(P) - 1$ .

## **2.3 REPRESENTAÇÃO DE UM POLIEDRO:**

Um poliedro  $P \subset \mathfrak{R}^n$  foi definido como uma interseção de um número finito de semi-espacos. Uma representação (ou uma descrição) de um poliedro  $P$  é um sistema de desigualdades  $Ax \leq b$  onde  $A \in \mathfrak{R}^{m \times n}$ ,  $b \in \mathfrak{R}^m$  tal que  $P = \{x \in \mathfrak{R}^n : Ax \leq b\}$ . Neste caso denotamos  $P = P(A,b)$ .

É comum fazer-se confusão entre um poliedro e uma determinada representação dela. Isto é inofensivo se os conceitos estão claros.

Dada uma determinada descrição de um poliedro  $P = P(A,b)$  através de um sistema  $Ax \leq b$  de inequações e uma face  $F$  deste poliedro, é interessante determinar um conjunto de inequações deste sistema que definem  $F$ .

Definimos os seguintes conjuntos:

$\text{ind}(A)$ , o conjunto dos índices das linhas da matriz  $A$ . Se  $F \subseteq P$   $\text{ig}(F) = \{i \in \text{ind}(A) : \mathcal{O}_i^T x = b_i; \text{ para todo } x \in F\}$ , onde  $\mathcal{O}_i^T$  é a  $i$ -essima linha da matriz  $A$ ,  $\text{ig}(F)$  é o conjunto igualdade de  $F$ .

É certo que se  $P \subset \mathfrak{R}^n$  é um poliedro e  $\text{ig}(P) = \emptyset$ , então  $P$  tem dimensão plena isto é  $\dim(P) = n$ . Para  $I \subseteq \text{ind}(A)$ , seja  $f_a(I) = \{x \in P : A_I x = b_I\}$  onde  $A_I$  é a submatriz de  $A$  com índices das linhas em  $I$  e  $b_I$  o subvetor de  $b$  com índices das componentes em  $I$ . Veremos no próximo teorema que  $f_a(I)$  constitui uma face de  $P$ , chamada a

face induzida por  $I$ . Se  $I = \{i\} \subset \text{ind}(A)$ ,  $F = \{x \in P: a_i^T x = b_i\}$  é uma face induzida pela inequação  $a_i^T x \leq b_i$ . O próximo teorema caracteriza as faces de um poliedro.

### Teorema 2.1

Sejam  $P = P(A,b)$  um poliedro e  $F \neq \emptyset$  um subconjunto de  $P$ . As seguintes afirmações são equivalentes:

- (a)  $F$  é uma face de  $P$
- (b) existe  $I \subseteq \text{ind}(A)$  tal que  $F = f_a(I)$
- (c)  $F = f_a(\text{ig}(F))$ .

Este teorema nos dá uma forma de obter faces de um poliedro  $P = P(A,b)$ . Para isto, basta escolhermos algumas inequações do sistema  $Ax \leq b$  e exigirmos igualdade nessas inequações. Esta caracterização nos dá um limitante superior para o número de faces de um poliedro, pois pelo teorema 2.1 o número máximo de faces não vazias é dado pelo número de subconjuntos de  $\text{ind}(A)$ , assim o poliedro  $P = P(A,b)$  com  $A \in \mathcal{R}^{m \times n}$  e  $b \in \mathcal{R}^m$  tem no máximo  $(2^m + 1)$  faces.

Por simplicidade, se uma inequação induz uma faceta então chamamos tal inequação de faceta. A seguir na próxima subseção mostraremos que as facetas constituem um sistema “minimal” de inequações para representar um poliedro.

## **2.4 DESCRIÇÕES IRREDUNDANTES DE UM POLIEDRO:**

Seja  $Ax \leq b$  um sistema de inequações. Por simplicidade denotamos o conjunto  $\text{ind}(A)$  por  $M$ . Seja  $I \subseteq M$  e  $D = M - I$ .

O sistema  $A_I x \leq b_I$  é redundante em relação a  $Ax \leq b$  se  $P(A,b) = P(A_D, b_D)$  onde  $A_D$  é uma submatriz de  $A$  com índices das linhas em  $D$  e  $b_D$  é um subvetor de  $b$  com índices das componentes em  $D$ . Se o sistema  $Ax \leq b$  contém algum subsistema

$A_I x \leq b_I$  redundante, então dizemos que o sistema  $Ax \leq b$  é redundante; caso contrário o sistema é dito irredundante.

Observar que a redundância ou irredundância não é propriedade de poliedros, mas sim de sistemas de inequações.

Um poliedro pode ter várias descrições irredundantes e várias descrições redundantes. Também, não é verdade que dado um sistema redundante ao removermos as inequações redundantes do sistema se obtenha um sistema irredundante.

Veremos a continuação que um sistema constituído pelas facetas de um poliedro é uma descrição irredundante do poliedro.

### Teorema 2.2

Seja  $P = P(A, b)$  um poliedro e  $\mathcal{F}$  um conjunto de facetas de  $P$  então

(a) Se  $F$  é uma faceta de  $P$ , então para cada  $i \in \text{ig}(F) \setminus \text{ig}(P)$  ( $\text{ig}(P) \subseteq \text{ig}(F)$ ) tem-se que

$$F = f_a(\{i\}) = \{x \in P : a_i^T x = b_i\}$$

(b)  $F_1, F_2 \in \mathcal{F}, F_1 \neq F_2$  então  $\text{ig}(F_1) \cap \text{ig}(F_2) = \text{ig}(P)$

(c)  $|\mathcal{F}| \leq |M| - |\text{ig}(P)|$

(d) existe  $I \subseteq M$  tal que  $I \subseteq M \setminus \text{ig}(P)$ ,  $|I| = |\mathcal{F}|$  e  $F \in \mathcal{F}$  se e somente se existe um único  $i \in I$  tal que  $F = f_a(\{i\})$ .

Um conjunto  $I \subseteq M$  com as propriedades descritas em (d) é chamado conjunto dos índices-faceta. Pela parte (a) do teorema temos que as facetas do polítopo  $P$  são obtidas impondo-se igualdade em apenas uma inequação  $a_i^T x \leq b_i$  do sistema

$Ax \leq b$ .

A seguir, caracterizamos descrições irredundantes de poliedros

### Teorema 2.3

Sejam  $P = P(A, b)$  um poliedro,  $P \neq \emptyset$ ,  $I \subseteq M \setminus \text{ig}(P)$ ,  $J \subseteq \text{ig}(P)$  e suponhamos que  $P = \{x \in \mathbb{R}^n : A_J x = b_J, A_I x = b_I\}$ . Esta descrição de  $P$  é irredundante se, e somente se

(a)  $I$  é um conjunto de índices faceta de  $P$ ;

(b)  $A_j$  é uma matriz  $p \times m$  de posto-linha completo (isto é, as linhas de  $A_j$  são linearmente independentes).

### Corolário 2.1

Seja  $P = P(A, b) \subseteq \mathfrak{R}^n$  um poliedro de dimensão plena. Então, para todo  $I \subseteq M$ ,  $A_I x \leq b_I$  é uma descrição irredundante de  $P$  se, e somente se,  $I$  é o conjunto de índices-faceta.

Dos resultados expostos acima, vemos que dado um polítopo  $P$  com a representação dada pelo sistema redundante  $Ax \leq b$  é possível (pelo menos na teoria) obter uma descrição irredundante de  $P$  da forma dada no teorema 2.3. Cabe notar que para cada faceta  $F$  de  $P$  uma das desigualdades representando  $F$  é necessário na descrição de  $P$ , e que cada inequação  $a_i^T x \leq b_i$  que induz uma face  $F$  de  $P$  com  $\dim(F) < \dim(P) - 1$  é irrelevante na descrição de  $P$ .

A tarefa de encontrar uma descrição irredundante de um polítopo é um problema difícil, pois eliminar inequações redundantes numa dada descrição de um polítopo é um problema ainda não resolvido com complexidade polinomial.

Existem somente algumas heurísticas para eliminar algumas inequações redundantes. A tarefa principal para eliminar inequações redundantes na descrição de um polítopo, é primeiro ser capaz de criar um algoritmo que reconheça se uma inequação induz ou não uma face do polítopo. Sabemos que dado um polítopo

$P = P(A, b) = \{x \in \mathfrak{R}^n : a_i^T x \leq b_i, i = 1, \dots, m\}$ , com  $A \in \mathfrak{R}^{m \times n}$ ,  $b \in \mathfrak{R}^m$  para cada  $i \in M = \text{ind}(A) = \{1, 2, \dots, m\}$  a inequação  $a_i^T x \leq b_i$  induz uma face  $F = \{x \in P : a_i^T x = b_i\}$  não vazia do polítopo  $P$  se, e somente se

$$\max \{a_i^T x : x \in P\} = b_i, \text{ mais ainda } F = \arg \max \{a_i^T x : x \in P\}.$$

portanto,  $a_i^T x \leq b_i$  não induz nenhuma face  $F \neq \emptyset$ ,  $F \neq P$  se, e somente se

$$\max \{ a_i^T x : x \in P \} = \alpha_i < b_i \quad (2.2) \quad (\text{ver Fig. 2.3})$$

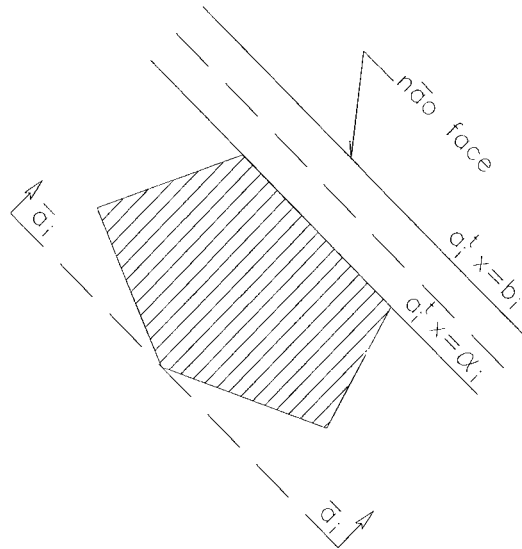


Figura 2.3

Vemos que se o número de desigualdades na descrição do polítopo é muito grande, testar se cada um induz uma face, mediante um problema de programação linear dado por (2.2) vai levar a um número grande de problemas de programação linear. Logo, esse teste pode servir somente para polítopos com pequeno número de desigualdades na sua descrição.

Em conclusão, dado um polítopo

$$P = \{x \in \mathbb{R}^n : Ax \leq b\} = \{x \in \mathbb{R}^n : a_i^T x \leq b_i, i = 1, \dots, m\},$$

podemos caracterizar quando uma restrição é redundante na definição do polítopo P, usando o seguinte lema.

### Lema 2.1

Uma restrição  $a_{\bar{j}}^T x \leq b_{\bar{j}}$  é redundante na definição do polítopo P se, e somente se,

$$\max \{ a_{\bar{j}}^T x : a_i^T x \leq b_i, i = 1, \dots, m, i \neq \bar{j} \} < b_{\bar{j}}$$

## 2.5 MÉTODO PRÁTICO PARA IDENTIFICAR ALGUMAS RESTRIÇÕES REDUNDANTES DE UM POLÍTOPO

Dado um polítopo  $P = \{x \in \mathfrak{R}^n : \alpha_i^T x \leq b_i, i = 1, \dots, m\}$ , uma maneira de conseguir reconhecer algumas restrições redundantes do polítopo é conseguir contruir um elipsóide que contenha o polítopo  $P$ , então os hiperplanos  $\alpha_i^T x = b_i$ , que não interseptam o elipsóide ou são tangentes ao elipsóide serão considerados redundantes para  $P$ .

Seja  $B$  uma matriz definida positiva, o conjunto

$$E(B,b,r) = \{x \in \mathfrak{R}^n : (x-b)^T B (x-b) \leq r^2\}$$

é chamado uma elipse de centro  $b$  e raio  $r > 0$

Lema 2.2 Seja  $a \in \mathfrak{R}^n$ , um vetor fixo. Então,

$$\max_{x \in E(B,b,r)} a^T x = a^T b + r \sqrt{a^T B^{-1} a}$$

$$\min_{x \in E(B,b,r)} a^T x = a^T b - r \sqrt{a^T B^{-1} a}$$

Corolário 2.2 O hiperplano  $H = \{x \in \mathfrak{R}^n : \alpha_i^T x = b_i\}$  interseptam o elipsóide  $E(B,b,r)$ , se e somente se

$$\alpha_i^T b - r \sqrt{\alpha_i^T B^{-1} \alpha_i} \leq b_i \leq \alpha_i^T b + r \sqrt{\alpha_i^T B^{-1} \alpha_i}$$

isto é se, somente se  $\left| \frac{\alpha_i^T b - b_i}{\alpha_i^T B^{-1} \alpha_i} \right| \leq r$

Logo, podemos estabelecer o seguinte algoritmo para reconhecer algumas restrições redundantes no polítopo  $P$



## Algoritmo

(0) Entrada:

$E = E(B, b, r)$  elipsóide contendo o polítopo

$$P = \{x \in \mathcal{R}^n : \alpha_i^T x \leq b_i, i = 1, \dots, m\}$$

$i = 1.$

(1) teste de redundancia

$$\text{Se } \left| \frac{\alpha_i^T b - b_i}{\alpha_i^T B^{-1} \alpha_i} \right| \geq r$$

então o hiperplano  $\alpha_i^T x = b_i$  é redundante na definição de  $P$

(2) teste de parada

Se  $i = m+1$  parar.

(3) Fazer

$i := i+1,$

ir ao passo (1).

# CAPÍTULO 3

## O PROBLEMA DE VIABILIDADE CONVEXA

### 3.1 INTRODUÇÃO

Dado um conjunto  $S \subset \mathfrak{R}^n$ ,  $S$  convexo,  $S$  limitado o problema de viabilidade convexa consiste em encontrar um ponto  $\bar{x} \in \mathfrak{R}^n$  tal que  $\bar{x} \in S$ .

O problema de viabilidade convexa (PVC) é um problema bastante geral de programação convexa, pois por exemplo o problema de minimizar uma função convexa  $f$  sobre um conjunto convexo  $\Omega \subset \mathfrak{R}^n$ , pode-se reduzir ao problema de encontrar um ponto  $\bar{x}$  no conjunto convexo  $\{x \in \Omega: f(x) - f(x^*) < \varepsilon\}$  para algum  $\varepsilon > 0$ , onde  $x^*$  é uma solução ótima do problema convexo.

A maioria dos algoritmos existentes para resolver o problema de viabilidade convexa envolvem um oráculo para  $S$ , que é um programa que aceita como entrada um determinado ponto  $\bar{z} \in \mathfrak{R}^n$  (ponto teste) e produz uma saída que depende da condição de  $\bar{z}$  pertencer ou não ao conjunto  $S$ . Se  $\bar{z} \in S$  o oráculo retorna ‘sim’ e o algoritmo pára, pois neste caso é resolvido o problema de viabilidade. De outro modo, se  $\bar{z} \notin S$  o oráculo retorna um ‘não’ acompanhado de um hiperplano separador de  $\bar{z}$  e  $S$ , isto é, retorna um vetor  $c$  tal que  $S \subset \{x \in \mathfrak{R}^n: c^T x \geq c^T \bar{z}\}$ , tal plano separador existe pois  $S$  é um conjunto convexo e  $\bar{z} \notin S$ .

Um algoritmo padrão de planos de corte para o problema de viabilidade convexa é como segue. Supondo que é conhecido que  $S$  esteja contido em alguma região prédefinida limitada  $P$ , e que se  $S$  é não vazia contenha alguma bola de determinado tamanho, isto é que  $S$  não tenha interior vazio, em cada iteração o algoritmo gera um ponto teste  $\bar{z} \in P$  e chama ao oráculo com  $\bar{z}$  como entrada, se  $\bar{z} \notin S$ , o oráculo retorna um vetor  $c$  tal que para cada  $x \in S$  se verifique que  $c^T x \geq c^T \bar{z}$ , isto é retorna um hiperplano separador de  $\bar{z}$  com  $S$  que geralmente é usado para cortar para outro lado uma seção de  $P$  que não contém pontos de  $S$ . Então, escolhendo  $\beta$  adequado tal que  $\beta \leq c^T \bar{z}$ , a região  $P$  é atualizada por  $P^* := P \cap$

$\{x \in \mathcal{R}^n: c^T x \geq \beta\}$ , é evidente que  $S \subset P^a$ . Então um novo ponto teste é escolhido dentro do novo bloco  $P^a$  para começar uma nova iteração, e assim o algoritmo prossegue, gerando novos blocos que contêm  $S$ . Naturalmente confia-se que os blocos ‘converjam’ ao conjunto  $S$ , o raio de convergência depende muito sobre o modo de selecionar os pontos testes que alimentam ao oráculo.

Em vários algoritmos de planos de corte para o problema geral de viabilidade convexa é mostrado que quando o algoritmo prossegue o volume dos blocos  $P^a$  gerados decrescem suficientemente num determinado raio, então, se  $S$  é não vazia, o algoritmo para com um ponto em  $S$  antes que o volume de  $P$  fique abaixo do volume da bola de raio  $\varepsilon$  contida em  $S$  ou o algoritmo para no primeiro momento que o volume de  $P$  decai abaixo do volume da bola de raio  $\varepsilon$  contida em  $S$ , mostrando que  $S$  é vazia. No transcurso do algoritmo a região  $P$  pode tornar-se complicada e a escolha do ponto teste pode ser cara; assim se  $P$  torna-se muito complicada é melhor substituir a região  $P$  por uma região mais simples que contenha  $P$ , esta substituição negocia volume por eficiência computacional, e o algoritmo ainda converge.

## 3.2 NOTAS HISTÓRICAS

A seguir apresentaremos uma breve resenha histórica da evolução dos métodos para o Problema de Viabilidade Convexa (PVC).

Um dos primeiros algoritmos muito conhecido e representativo que cai dentro do esquema padrão de planos de corte descrito acima é o algoritmo chamado de método de elipsóides. Neste algoritmo a região  $P$  é um elipsóide  $E$  que contém  $S$ , e o ponto teste  $\bar{z}$  é o centro do elipsóide. Um plano de corte corta ao elipsóide  $E$  em duas partes, uma delas contém o conjunto  $S$ , e então é construindo outro elipsóide  $E^a$  de volume menor do que  $E$  tal que  $E^a$  contém a parte de  $E$  que contém  $S$ , e assim o algoritmo prossegue até encontrar um ponto em  $S$  ou declarar que  $S$  é vazia.

Os primeiros trabalhos que contribuíram na gestação do método das elipsóides, foram os trabalhos de Shor[36] e de Nemirovsky [45], Shor descreveu

um novo algoritmo gradiente projetado com dilatação de espaço para programação convexa não diferenciável e Nemirovsky observou que os algoritmos de Shor providenciavam uma alternativa ao problema discutido por Levin [30] e que de alguma maneira dava um esboço ao método de elipsóides. O primeiro estabelecimento explícito do método de elipsóides, que é conhecido atualmente é devido a Shor[37] para programação linear não convexa. Este método tornou-se bastante conhecido e causou sensação com a publicação de Khachiyan[25] quem adaptou o método de elipsóides para provar o tempo polinomial de solução de problemas de programação linear (PL). O algoritmo Khachiyan foi o primeiro algoritmo polinomial para PL. Contrariamente ao método simplex, o método de elipsóides não usa a estrutura combinatorial de PL. Se bem que o algoritmo de khachiyan tem o mérito de ser o primeiro algoritmo polinomial para PL, na pratica é de pouco valor ao menos até o presente tempo, e têm pouco uso, principalmente devido a sua lenta convergência, além de outras desvantagens computacionais. Não obstante, acredita-se que o método dos elipsóides tem uma fundamental potência teórica por suas muitas conseqüências teóricas, pois é uma ferramenta elegante para provar o tempo polinomial de solução de muitos problemas de otimização geométrica e combinatória ([17][4]). Existem conexões entre o método de elipsóides e os algoritmos quasi-Newton para programação não linear e com o algoritmo de Karmarkar para programação linear[18], [46]. Contrariamente ao seu comportamento em PL, parece que a versão central do método de elipsóides é robusta para problemas de programação não linear em relação a sua eficiência, compete com outros algoritmos propostos [12],[10],[13]. Em 1988 Tarasov-Khachiyan-Erlich [41] usaram um método de planos de corte chamado o método de Elipsóides Inscritos para resolver um problema geral de PNL, na qual o ponto de teste é uma aproximação ao centro do elipsóide de máximo volume inscrito num polítopo que contém o conjunto de localização (que contém as soluções de PNL).

Outro método que cabe mencionar para o problema de viabilidade convexa que pertence ao esquema padrão de planos de corte, é o Método de Centros Volumétricos de P.M.Vaidya [42] que usa o centro volumétrico de um polítopo como ponto teste, O centro volumétrico de um polítopo, é um ponto que minimiza

o logaritmo do determinante da matriz hessiana da função barreira logarítmica sobre o polítopo. Vaidya estabeleceu resultados de convergência e complexidade polinomial por eliminação de planos “não influentes”, mostrou que o algoritmo converge em  $O(nl)$  iterações. Até antes de Vaidya o método de elipsóides era o método mais conhecido para o problema de viabilidade que converge dentro de  $O(n^2 L)$  iterações.

Por outro lado, vários pesquisadores desenvolveram métodos de planos de corte, usando como ponto teste o centro analítico (ou aproximado) de um polítopo, que vem a ser o ponto que minimiza a função barreira logarítmica sobre o polítopo. A primeira idéia de usar cortes através do centro analítico foi dada por Sonnevend [38], ele sugeriu fazer desse modo, mais não deu um análise algorítmico.

A razão de usar o centro analítico no método de planos de corte é evidente, pois é fácil de calcular, e atualizar, e tem desempenhado um rol importante em muitos algoritmos de pontos interiores para PL e PNL. Alguns resultados experimentais têm mostrado que eleger o centro analítico como ponto teste é muito eficiente, as publicações de Goffin, Bahn, Merle e Vidal são testemunha de tal eficiência [2], [19], [3], todos estes trabalhos reportaram um bom comportamento de seus algoritmos de planos de corte pelo centro analítico para problema de programação convexa, mais nenhum deles apresentou uma análise de convergência e complexidade.

Foi, Ye [47] o primeiro que deu um resultado de complexidade para um algoritmo de plano de corte pelo centro analítico, Ye desenvolveu um algoritmo de decomposição e redução de potencial seguindo o esquema de geração de colunas baseado no centro analítico para resolver o problema de viabilidade linear, isto é, quando o conjunto convexo  $S$  é um polítopo. Ele mostrou que o algoritmo necessita somente  $O(qL)$  iterações para gerar um ponto no polítopo  $S$ , sendo “ $q$ ” o número de cortes adicionados ao sistema (sem eliminação de cortes) e  $L$  é o tamanho dos dados de entrada. A desvantagem é que embora que  $q$  seja geralmente pequeno para resolver muitos problemas práticos, pode crescer em um número exponencial, daí que a complexidade obtida por YE é pior que a complexidade do

método de elipsóides para alguns problemas PL com grande número de desigualdades.

No ano 1992 Atkinson e Vaidya [1] apresentaram um algoritmo de plano corte-centro analítico para resolver o problema geral de viabilidade convexa. Eles estabeleceram resultados de convergências e complexidade polinomial, usando um critério de eliminação de planos de corte considerados “não importantes”. A diferença do resultado de Ye a complexidade obtida por Atkinson e Vaidya é independente do número de planos de corte gerados sendo na ordem  $O(nL^2)$  iterações. Constituindo assim o primeiro resultado de complexidade polinomial convincente de um algoritmo de plano de corte centro analítico para o problema de viabilidade convexa.

Por outro lado Y. Nesterov[33], estudou dois esquemas de métodos de planos de corte para minimização de uma função convexa, baseados sobre modificações da função barreira logarítmica. O primeiro método é uma variante conceitual do tradicional método de centros analíticos, e usa uma bola como localizador inicial de uma solução. Enquanto que o segundo método é uma variante conceptual do método proximal e do método centro analítico. Chamado método proximal barreira centro analítico. Nesterov dá para ambos esquemas estimativas do raio de convergência. Considerando uma modificação implementável do segundo esquema, na qual remove o cálculo exato do minimizador de função barreira e introduzindo uma técnica padrão de atualizar aproximações da inversa da hessiana da função barreira logarítmica, (técnica bem conhecida que melhora estimativas de eficiência de algoritmos polinomiais de ponto interior em PL e programação quadrática). Ele reduz a total complexidade aritmética do método. Em seus resultados usa fortemente resultados da teoria de funções chamadas auto concordantes (Self-Concordant). Mas Nesterov não apresenta uma extensão de seus algoritmos para o problema de viabilidade convexa.

Continuando com a pesquisa sobre o problema de viabilidade convexa, no ano 1994 Goffin, Luo e Ye [21] fizeram uma análise mais profunda da convergência e complexidade de um algoritmo plano de corte centro analítico para resolver o problema geral de viabilidade convexa definidos por um oráculo de

separação. Usando-os centros analíticos aproximados dos polítopos gerados no algoritmo, eles mostraram que o algoritmo converge em um tempo finito, sem usar eliminação de cortes e provaram que o algoritmo é um pleno esquema da aproximação polinomial o que significa que para  $\varepsilon > 0$  fixo (a exatidão da terminação) o método é polinomial na dimensão do espaço  $n$ , exatamente na ordem de  $O^*(n^2/\varepsilon^2)$  iterações onde a notação  $O^*$  significa que os termos de menor ordem são ignorados.

Baseados fundamentalmente no algoritmo Goffin-Luo-Ye[21] apareceram outras variantes. Assim no mesmo ano 1994, Y.Ye [48] analisou a complexidade de uma variante do algoritmo de Goffin-Luo-Ye, quando são adicionados em cada iteração múltiplas cortes pelo mesmo centro analítico exato. Ye mostrou que o algoritmo é ainda um algoritmo de plena aproximação polinomial com um limitante de ordem  $O^*(\eta^2 n^2/\varepsilon^2)$  sobre o número total de iterações necessitadas para encontrar uma  $\varepsilon$ -solução viável, sendo  $\eta$  o número máximo de cortes que pode ser adicionado em cada iteração e  $n$  a dimensão do espaço. Por outro lado, usando aproximados “pesados” centros analíticos Q.Luo [43] mostrou que o método de cortes múltiplas de Ye [48] tem complexidade  $O^*(\eta n^2/\varepsilon^2)$  melhorando a complexidade obtida por Ye. Novamente J.L.Goffin [22] melhora os resultados obtidos em [21] usando o primal-dual direção de atualização, e cortes profundos pelo centro analítico aproximado. Paralelamente, apareceram as publicações de K.Kiwiel [27] baseado fundamentalmente no algoritmo de Goffin-Luo-Ye [21], Kiwiel estuda métodos de plano de corte para minimização de uma função convexa sujeita a restrições de caixa. Em cada iteração, cortes subgradiente definem um polítopo que localiza o mínimo. A função objetivo e seus sub-gradientes são avaliados no centro analítico deste polítopo para produzir um ou dois tipos de cortes que melhoram o conjunto localizador. Kiwiel dá também estimativas de complexidades de seu método.

Atualmente continuam aparecendo outras variantes do método de plano de corte centro analítico para problemas de programação convexa, e outras para o problema geral de viabilidade convexa que ainda não estão publicadas.

Baseados em várias das publicações existentes para o (PVC) nosso objetivo na presente tese é abordar o problema geral de viabilidade convexa, usando fundamentalmente o método de planos de corte centro e o método de elipsóides usando a eliminação de cortes redundantes com a intenção de melhorar a eficiência computacional dos métodos de planos de corte, mantendo a mesma complexidade polinomial do método, o que apresentamos no capítulo 5.



### 3.3 PRELIMINARES

#### 3.3.1 FUNÇÃO BARREIRA LOGARÍTMICA

Seja

$$- \quad \mathbf{A} \in \mathcal{R}^{m \times n}, \mathbf{b} \in \mathcal{R}^m \text{ e } P = \{x \in \mathcal{R}^n : \mathbf{A}x \geq \mathbf{b}\},$$

Se

$$- \quad \mathcal{A}_i^T \text{ são as linhas de } \mathbf{A} \text{ e } \mathbf{b} = (b_1, \dots, b_m),$$

então

$$P = \{x \in \mathcal{R}^n : \mathcal{A}_i^T x \geq b_i, i = 1, 2, \dots, m\}.$$

Suponhamos que  $P$  seja um polítopo em  $\mathcal{R}^n$  com interior

$$- \quad \text{int}P = \{x \in \mathcal{R}^n : \mathcal{A}_i^T x > b_i, i = 1, 2, \dots, m\} \text{ não vazio.}$$

Então definimos a **função barreira logarítmica** sobre  $P$  por

$$- \quad F(x) = - \sum_{i=1}^m \ln(\mathcal{A}_i^T x - b_i), \text{ para cada } x \in \text{int}P.$$

A função barreira logarítmica é estritamente convexa, portanto tem um único minimizador  $w$  sobre  $\text{int}P$ . O ponto  $w$  é chamado o **centro analítico** do polítopo  $P$ .

É importante saber que o centro analítico é um conceito analítico e não um conceito geométrico. Como consequência, o centro analítico depende também sobre a maneira de descrever o polítopo  $P$ , daí que quando se elimina restrições redundantes no polítopo  $P$ , o centro analítico pode variar, mesmo que o polítopo  $P$  continue sendo o mesmo conjunto.

Para cada  $x \in \text{int}P$  o vetor gradiente  $\nabla F(x)$  e matriz Hessiana  $\nabla^2 F(x)$  são dadas por

$$- \quad \nabla F(x) = - \sum_{i=1}^m \frac{\mathcal{A}_i^T}{\mathcal{A}_i^T x - b_i}$$

$$- \quad \nabla^2 F(x) = \sum_{i=1}^m \frac{\mathcal{A}_i \mathcal{A}_i^T}{(\mathcal{A}_i^T x - b_i)^2}$$

Se definimos as folgas  $s_i = \mathcal{A}_i^T x - b_i > 0$ , para  $i = 1, \dots, m$  e consideramos o vetor  $\mathbf{e} = (1, \dots, 1)^T \in \mathcal{R}^m$  (com todas as componentes iguais a 1), então pode-se reescrever

$$- \quad \nabla F(x) = - \mathbf{A}^T \mathbf{S}^{-1} \mathbf{e}$$

$$- \quad \nabla^2 F(x) = \mathbf{A}^T \mathbf{S}^{-2} \mathbf{A}, \text{ onde } \mathbf{S}^{-1} \text{ e } \mathbf{S}^{-2} \text{ são as matrizes diagonais definidas por}$$

$$- \quad \mathbf{S}^{-1} = \text{diag.}(1/s_i)$$

$$S^{-2} = \text{diag.}(1/s_i^2)$$

A matriz Hessiana  $\nabla^2 F(x)$  é uma matriz simétrica definida positiva para cada  $x \in \text{int.}P$ , daí que a função barreira logarítmica  $F(x)$  é estritamente convexa..

### 3.3.2 ALGORITMO PADRÃO DE PLANOS DE CORTE-CENTRO ANALÍTICO (APC)

A seguir apresentamos o esquema padrão da maioria dos algoritmos de planos de corte que usam o centro analítico (ou aproximados) como ponto teste, em particular nosso algoritmo (AER) que apresentamos no capítulo 5 cai dentro deste esquema.

#### ALGORITMO PADRÃO (APC)

##### ENTRADA

s um conjunto convexo em  $\mathcal{R}^n$  e P um polítopo em  $\mathcal{R}^n$  tal que  $S \subseteq P$

##### ITERAÇÃO GENÉRICA

1. Calcular (aproximadamente) o centro analítico  $\omega$  do polítopo P
2. Chamar ao oráculo de S com  $\omega$  como entrada:
  - (a) Se  $\omega \in S$  parar.
  - (b) Se  $\omega \notin S$ , o oráculo retorna um vetor  $\bar{c}$  tal que  $\{x \in \mathcal{R}^n : c^T x \geq c^T \omega\} \supseteq S$
3. Adicionar o corte  $\bar{c}^T x \geq \beta$  onde  $\bar{c}^T \omega \geq \beta$   
Fazer  $P := P \cap \{x : \bar{c}^T x \geq \beta\}$
4. ir para passo 1.

Dentro deste padrão geral, encaixam a maioria dos algoritmos de planos de corte existentes. A diferença entre estes algoritmos radica principalmente nos testes de parada do algoritmo, e também na maneira de atualizar o centro analítico aproximado.

As demonstrações de complexidade para os algoritmos de planos de corte-centro analítico tem duas componentes : uma orientada a mostrar a complexidade do método de centros analíticos e uma segunda orientada a mostrar que é possível calcular em  $O(1)$  passos de Newton um centro analítico aproximado de  $P^{k+1}$  dado um centro analítico aproximado de  $P^k$  sendo ambos componentes discutidas amplamente na maioria dos algoritmos existentes.

### 3.3.3 PROPRIEDADES DO ELIPSÓIDE HESSIANO

Na continuação apresentaremos vários resultados relacionados com a matriz Hessiana  $\nabla^2 F(x)$  da função barreira logarítmica e os elipsóides definidos pela matriz  $\nabla^2 F(x)$ . As demonstrações estão contidas em [1].

Para cada matriz definida positiva  $A$ , o conjunto

$$E(A, a, \nu) = \{ x \in \mathbb{R}^n : (x-a)^T A (x-a) \leq \nu^2 \}, \nu > 0$$

é chamado de elipsóide de centro  $a$  e raio  $\nu > 0$ .

Quando  $A = \nabla^2 F(z)$ , o elipsóide  $E(\nabla^2 F(z), z, \nu)$  é chamado elipsóide Hessiano.

#### LEMA 3.1

Para cada  $z \in \text{int.}P$ , tem-se que  $E(\nabla^2 F(z), z, 1) \subset P$ .

#### Prova

Sabemos que

$$E(\nabla^2 F(z), z, 1) = \{ x \in \mathbb{R}^n : (x-z)^T \nabla^2 F(z) (x-z) \leq 1 \},$$

$$\text{onde } \nabla^2 F(z) = \sum_{i=1}^m \frac{a_i a_i^T}{(a_i^T z - b_i)^2}$$

Se  $x \in E(\nabla^2 F(x), z, 1)$ ,

$$\text{então } (x-z)^T \sum_{i=1}^m \frac{a_i a_i^T}{(a_i^T z - b_i)^2} (x-z) \leq 1$$

$$\text{o que implica que } \sum_{i=1}^m \frac{(a_i^T (x-z))^2}{(a_i^T z - b_i)^2} \leq 1.$$

Logo para cada  $i = 1, \dots, m$  tem-se que  $\frac{(\alpha_i^T(x-z))^2}{(\alpha_i^T z - b_i)^2} \leq 1$

e assim  $|\alpha_i^T(x-z)| \leq |\alpha_i^T z - b_i|$

de onde  $|\alpha_i^T(x-z)| \leq \alpha_i^T z - b_i$

portanto,  $\alpha_i^T z - b_i \geq 0$ , ou seja  $x \in P$ .

-

### LEMA 3.2

Seja  $B$  uma  $n \times m$  matriz simétrica, definida positiva, e seja

$E(B, z, \nu) = \{x \in \mathfrak{R}^n : (x-z)^T B(x-z) \leq \nu^2\}$ , e  $c \in \mathfrak{R}^n$  um vetor fixo. Então

$$\max_{x \in E(B, z, \nu)} c^T(x-z) = \nu \sqrt{c^T B^{-1} c}$$

Prova:

Seja  $x^*$  o ponto que maximiza a função linear sobre o elipsóide  $E(B, z, r)$ . Então das condições de Karush-Kuhn-Tucker segue-se que

$$B(x^*-z) = \lambda c, \text{ para algum } \lambda \in \mathfrak{R}$$

Logo  $x^* = z + \lambda B^{-1} c$ , e como  $x^*$  pertence a fronteira do elipsóide, então

$$(x^*-z)^T B(x^*-z) = \nu^2. \text{ Portanto}$$

$$(\lambda B^{-1} c)^T B(\lambda B^{-1} c) = \nu^2$$

ou seja  $\lambda^2 (c^T B^{-1} c) = \nu^2$  o que dá  $\lambda = \frac{\nu}{\sqrt{c^T B^{-1} c}}$

$$\text{então } x^* = z + \frac{\nu}{\sqrt{c^T B^{-1} c}} B^{-1} c$$

Assim

$$c^T (x^*-z) = \frac{\nu}{\sqrt{c^T B^{-1} c}} c^T B^{-1} c = \nu \sqrt{c^T B^{-1} c}$$

ou seja

$$\max_{x \in E(B, z, \nu)} c^T(x-z) = \nu \sqrt{c^T B^{-1} c}.$$

-

## FUNÇÃO $\Psi(x)$

Definimos a função  $\Psi(x)$  por

$$\Psi(x) = \nabla F(x)^T (\nabla^2 F(x))^{-1} \nabla F(x) \text{ para } x \in \text{int}P. \text{ Logo } \Psi(x) > 0.$$

Esta função é útil, frequentemente usada na discussão das propriedades de função barreira logaritmica, como veremos depois.

Consideremos a expansão de Taylor de segundo grau da Função  $F(x)$  entorno do ponto  $z$

$$F(x) - F(z) = \nabla F(z)^T (x-z) + \frac{1}{2} (x-z)^T \nabla^2 F(z)(x-z) + e_\nu$$

onde  $e_\nu$  é o erro de aproximação.

É ilustrativo observar que em um passo na direção de Newton desde o ponto  $z$ , temos

$$z' = z - (\nabla^2 F(z))^{-1} \nabla F(z)$$

então

$$F(z') - F(z) = -\Psi(z) + \frac{1}{2} \Psi(z) + e_\nu = -\frac{1}{2} \Psi(z) + e_\nu$$

se podemos provar que o erro é muito pequeno, então  $\Psi(z)$  dá uma medida do acréscimo alcançado tomando passos de Newton.

-

## VIZINHANÇA $\sum(z, \nu)$

Seja  $z \in \text{int}P$  e  $\nu > 0$ , a região  $\sum(z, \nu)$  é definida por

$$\sum(z, \nu) = \{x \in \mathbb{R}^n : \left| \frac{a_i^T (x-z)}{a_i^T z - b_i} \right| \leq \nu, i=1, \dots, m\}$$

que é uma generalização de uma bola em torno do ponto  $z$ . É evidente que para  $\nu$

$\leq 1$ ,  $\sum(z, \nu) \subseteq P$ . A forma de  $\sum(z, \nu)$  depende dos planos que formam o polítopo  $P$  e da distância de  $z$  a cada faceta de  $P$ . (ver Fig. 3.1)

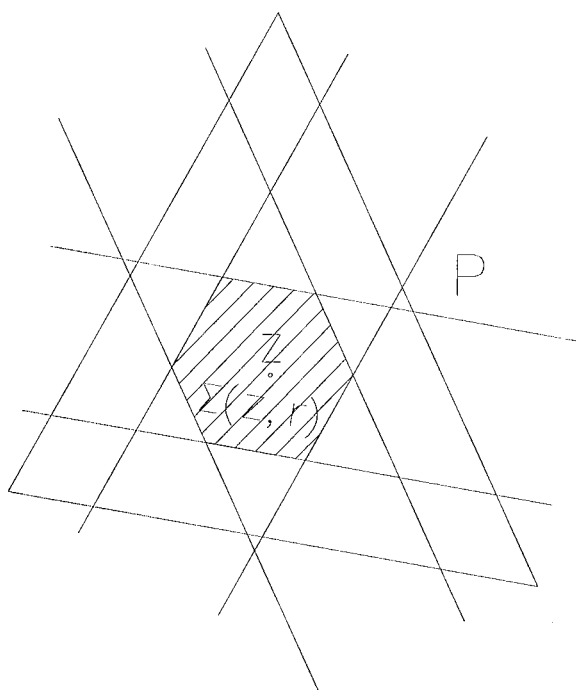


Fig. 3.1

LEMA 3.3

Para  $z \in \text{int.P}$  e  $\alpha > 0$ , temos que  $E(\nabla^2 F(x), z, \alpha) \subset \Sigma(z, \alpha)$

Prova

Seja  $y \in E(\nabla^2 F(x), z, \alpha)$ , pelo lema 3.2

$$\max_{x \in E(\nabla^2 F(x), z, \alpha)} (a_i^\top (x-z))^2 = \alpha^2 a_i^\top (\nabla^2 F(x))^{-1} a_i \quad \text{para } i = 1, \dots, \alpha$$

então  $(a_i^\top (y-z))^2 \leq \alpha^2 a_i^\top (\nabla^2 F(x))^{-1} a_i,$

$$\text{isto é } |\alpha_i^T(y-z)| \leq \alpha \sqrt{\alpha_i^T(\nabla^2 F(z))^{-1}\alpha_i} \quad (3.1)$$

Por outro lado

$$- \max_{x \in E(\nabla^2 F(z), z, 1)} \alpha_i^T(x-z) = \sqrt{\alpha_i^T(\nabla^2 F(z))^{-1}\alpha_i}$$

Se  $x \in E(\nabla^2 F(x), z, 1)$ , então  $\alpha_i^T(x-z) \leq \alpha_i^T z - b_i$  (ver prova do lema 3.1),

$$\text{logo } \max_{x \in E(\nabla^2 F(z), z, 1)} \alpha_i^T(x-z) \leq \alpha_i^T z - b_i$$

portanto

$$(*) \quad \sqrt{\alpha_i^T(\nabla^2 F(z))^{-1}\alpha_i} \leq \alpha_i^T z - b_i, \text{ para cada } i = 1, \dots, m$$

e assim de (3.1) segue-se que

$$- |\alpha_i^T(y-z)| \leq \alpha (\alpha_i^T z - b_i)$$

$$\text{e como } \alpha_i^T z - b_i > 0, \text{ conclui-se que } \left| \frac{\alpha_i^T(y-z)}{\alpha_i^T z - b_i} \right| \leq \alpha \text{ para cada } i = 1, \dots,$$

m

o que implica que  $y \in \sum (z, \alpha)$ .

Criar uma região na qual a Hessiana não mude muito é muito útil na demonstração da convergência de algoritmos de minimização barreira logarítmica.

#### LEMA 3.4

Se  $y \in \sum (z, \alpha)$  onde  $\alpha < 1$ ,  $z \in \text{int.P}$ , então para cada  $\xi \in \mathfrak{R}^n$

$$- \frac{\xi^T \nabla^2 F(z) \xi}{(1+\alpha)^2} \leq \xi^T \nabla^2 F(y) \xi \leq \frac{\xi^T \nabla^2 F(z) \xi}{(1-\alpha)^2}$$

Quando  $\alpha \geq 1$  a desigualdade do lado esquerdo ainda vale.

Prova

$$\text{Temos que } \nabla^2 F(\mathbf{x}) = \sum_{i=1}^m \frac{\mathbf{a}_i \mathbf{a}_i^T}{(\mathbf{a}_i^T \mathbf{x} - b_i)^2}$$

Assim para cada  $\xi \in \mathfrak{R}^n$

$$\xi^T \nabla^2 F(\mathbf{x}) \xi = \sum_{i=1}^m \frac{(\mathbf{a}_i^T \xi)^2}{(\mathbf{a}_i^T \mathbf{y} - b_i)^2} = \sum_{i=1}^m \frac{(\mathbf{a}_i^T \xi)^2}{(\mathbf{a}_i^T \mathbf{z} - b_i)^2} \cdot \frac{(\mathbf{a}_i^T \mathbf{z} - b_i)^2}{(\mathbf{a}_i^T \mathbf{y} - b_i)^2}$$

Como  $\mathbf{y} \in \sum (\mathbf{z}, \alpha)$ , então  $\left| \frac{\mathbf{a}_i^T (\mathbf{y} - \mathbf{z})}{\mathbf{a}_i^T \mathbf{z} - b_i} \right| \leq \alpha$ . Também,

-

$$- \left| \frac{\mathbf{a}_i^T (\mathbf{y} - b_i)}{\mathbf{a}_i^T \mathbf{z} - b_i} \right| = \left| \frac{\mathbf{a}_i^T (\mathbf{y} - \mathbf{z})}{\mathbf{a}_i^T \mathbf{z} - b_i} + 1 \right| \geq \left| \frac{\mathbf{a}_i^T (\mathbf{y} - \mathbf{z})}{\mathbf{a}_i^T \mathbf{z} - b_i} - 1 \right|$$

-

de onde,

$$- 1 - \left| \frac{\mathbf{a}_i^T (\mathbf{y} - \mathbf{z})}{\mathbf{a}_i^T \mathbf{z} - b_i} \right| \leq \left| \frac{\mathbf{a}_i^T (\mathbf{y} - b_i)}{\mathbf{a}_i^T \mathbf{z} - b_i} \right| \leq \left| \frac{\mathbf{a}_i^T (\mathbf{y} - \mathbf{z})}{\mathbf{a}_i^T \mathbf{z} - b_i} \right| + 1$$

Assim,

$$- 1 - \alpha \leq \left| \frac{\mathbf{a}_i^T (\mathbf{y} - b_i)}{\mathbf{a}_i^T \mathbf{z} - b_i} \right| \leq 1 + \alpha, \quad 1 - \alpha > 0$$

$$\text{o que dá } 0 < \frac{1}{(1 + \alpha)} \frac{1}{|\mathbf{a}_i^T \mathbf{z} - b_i|} \leq \frac{1}{|\mathbf{a}_i^T \mathbf{y} - b_i|} \leq \frac{1}{(1 - \alpha)} \frac{1}{|\mathbf{a}_i^T \mathbf{z} - b_i|}$$

portanto,

$$\frac{1}{(1 + \alpha)^2} \sum_{i=1}^m \frac{(\mathbf{a}_i^T \xi)^2}{(\mathbf{a}_i^T \mathbf{z} - b_i)^2} \leq \sum_{i=1}^m \frac{(\mathbf{a}_i^T \xi)^2}{(\mathbf{a}_i^T \mathbf{y} - b_i)^2} \leq \frac{1}{(1 - \alpha)^2} \sum_{i=1}^m \frac{(\mathbf{a}_i^T \xi)^2}{(\mathbf{a}_i^T \mathbf{z} - b_i)^2}$$

o que significa,

$$\frac{1}{(1 + \alpha)^2} \xi^T \nabla^2 F(\mathbf{z}) \xi \leq \xi^T \nabla^2 F(\mathbf{y}) \xi \leq \frac{1}{(1 - \alpha)^2} \xi^T \nabla^2 F(\mathbf{z}) \xi \quad (3.2)$$

para cada  $\xi \in \mathfrak{R}^n$ .



No caso  $\alpha \geq 1$ ,  $0 < \frac{1}{(1+\alpha)} \frac{1}{|\alpha_i^T z - b_i|} \leq \frac{1}{|\alpha_i^T y - b_i|}$ , portanto

a desigualdade do lado esquerdo de (3.2) é ainda válido.

Antes de estabelecer um corolário útil, do lema 3.4, apresentamos o seguinte lema.

### LEMA 3.5

Se  $A$  e  $B$  são matrizes de ordem  $n \times n$  definidas positivas, tal que

$\xi^T A \xi \geq \theta \xi^T B \xi$  para algum  $\theta > 0$  e para todo  $\xi \in \mathfrak{R}^n$ . Então  $\xi^T A^{-1} \xi \leq \frac{1}{\theta} \xi^T B^{-1}$

$\xi$  para todo  $\xi \in \mathfrak{R}^n$ .

#### Prova

Se  $\xi^T A \xi \leq 1$  para cada  $\xi \in \mathfrak{R}^n$ , então pela hipótese,

$\xi^T B^{-1} \xi \leq \frac{1}{\theta}$ . Isto quer dizer que  $E(A, 0, 1) \subset E(B, 0, \frac{1}{\theta})$

logo  $\max_{z \in E(A, 0, 1)} \xi^T z \leq \max_{y \in E(B, 0, \frac{1}{\theta})} \xi^T y$

que pelo lema 3.2, equivale a

$$\sqrt{\xi^T A^{-1} \xi} \leq \frac{1}{\sqrt{\theta}} \sqrt{\xi^T B^{-1} \xi}$$

portanto  $\xi^T A^{-1} \xi \leq \frac{1}{\theta} \xi^T B^{-1} \xi$

### COROLÁRIO 3.1 (do lema 3.4)

Se  $y \in \sum (z, \alpha)$  onde  $\alpha < 1$ ,  $z \in \text{int.P}$ , então para cada  $\xi \in \mathfrak{R}^n$  tem-se que

$$(1+\alpha)^2 \xi^T (\nabla^2 F(z))^{-1} \xi \geq \xi^T (\nabla^2 F(y))^{-1} \xi \geq (1-\alpha)^2 \xi^T (\nabla^2 F(z))^{-1} \xi$$

Quando  $\alpha \geq 1$  a desigualdade do lado direito é ainda válida.

### OBSERVAÇÃO

O lema 3.4 e o corolário 3.1 são válidos para  $y \in E(\nabla^2 F(x), z, \alpha)$ , com

$\alpha < 1$ ,  $z \in \text{int.P}$ , pois  $E(\nabla^2 F(x), z, \alpha) \subset \sum(z, \alpha)$

-  
-  
-  
-  
-  
-  
-  
-  
-  
-

### 3.3.4 VALORES DA FUNÇÃO BARREIRA LOGARÍTMICA PERTO DO CENTRO ANALÍTICO

- Agora nossa intenção é mostrar que o valor  $\Psi(z)$  é de fundamental utilidade para dar uma medida local da magnitude  $F(z) - F(w)$  onde  $w$  é o centro analítico exato de  $P$  e  $z \in \text{int.P}$ . Mostraremos que quando  $\Psi(z)$  é pequeno, a diferença  $F(z) - F(w)$  é limitada, com limitantes superior e inferior dependendo de  $\Psi(z)$ . Assim para  $\Psi(z)$  suficientemente pequeno, o ponto  $z$  será uma aproximação aceitável de  $w$ .

- Para mostrar ao resultado mencionado acima precisamos de vários lemas anteriores que estabeleceremos a seguir.

Para um ponto fixo  $w \in \mathfrak{R}^n$  definimos a função

- 
$$\Phi : \mathfrak{R}^{n+1} \rightarrow \mathfrak{R}^n$$

- 
$$\Phi(x, t) = \nabla F(x) - tw \text{ para } x \in \text{int.P.}$$

Como a função  $x \rightarrow \nabla F(x)$  é diferenciável e  $\det(\nabla^2 F(x)) \neq 0$  em cada ponto do conjunto aberto  $\text{int.P}$ , então a função  $\Phi(x, t)$  satisfaz as condições do teorema da função implícita no aberto  $(\text{int.P}) \times \mathfrak{R}$ . Portanto a equação  $\Phi(x, t) \equiv 0$  ou seja  $\nabla F(x) = tw$  define implicitamente uma função  $x = x(t)$  diferenciável em um adequado domínio aberto e sua derivada  $\dot{x}(t)$  satisfaz

$$\nabla^2 F(x) \dot{x}(t) = w, \text{ por tanto } \dot{x}(t) = (\nabla^2 F(x))^{-1} w$$

-

LEMA 3.6

Seja  $w \in \mathbb{R}^n$  um vetor fixo e  $0 < \nu < 1$ . Seja  $x = x(t)$  a trajetória definida implicitamente pela equação  $\nabla F(x) = tw$ , e seja  $\hat{x}$  um ponto tal que  $\nabla F(\hat{x}) = \hat{t}(w)$

para algum  $\hat{t}$ . Se  $x = x(t) \in \sum (\hat{x}, \nu)$ , então

$$\left| \frac{\alpha_i^T \dot{x}(t)}{\alpha_i^T x - b_i} \right| \leq (1+\nu) \sqrt{w^T (\nabla^2 F(\hat{x}))^{-1} w}$$

Prova:

Como  $\dot{x}(t) = (\nabla^2 F(x))^{-1} w$  temos que

$$\left( \frac{\alpha_i^T \dot{x}(t)}{\alpha_i^T z - b_i} \right)^2 = \left( \frac{\alpha_i^T (\nabla^2 F(x))^{-1} w}{\alpha_i^T z - b_i} \right)^2$$

Se

$$u = \frac{(\nabla^2 F(x))^{-1/2} \alpha_i}{\sqrt{\alpha_i^T x - b_i}} \quad e \quad v = \frac{(\nabla^2 F(x))^{-1/2} w}{\sqrt{\alpha_i^T x - b_i}}$$

$$\left( \frac{\alpha_i^T (\dot{x}(t))}{\alpha_i^T z - b_i} \right)^2 = (u^T v)^2 \leq \|u\|_2^2 \|v\|_2^2$$

$$= \left( \frac{\alpha_i^T (\nabla^2 F(x))^{-1} \alpha_i}{\alpha_i^T z - b_i} \right)^2 \left( \frac{w^T (\nabla^2 F(x))^{-1} w}{\alpha_i^T z - b_i} \right)^2$$

$$= \frac{\alpha_i^T (\nabla^2 F(x))^{-1} \alpha_i}{(\alpha_i^T z - b_i)^2} (w^T (\nabla^2 F(x))^{-1} w)$$

por outro lado, pelo corolário 3.1. para  $x \in \sum (\hat{x}, \nu)$ ,  $\nu < 1$

temos que  $w^T (\nabla^2 F(w))^{-1} w \leq (1+\nu)^2 w^T (\nabla^2 F(\hat{x}))^{-1} w$

Também  $\frac{\alpha_i^T (\nabla^2 F(x))^{-1} \alpha_i}{(\alpha_i^T z - b_i)^2} \leq 1$  para  $x \in \text{int}P$

portanto, segue que

$$- \left( \frac{\alpha_i^T(\dot{x}(t))}{\alpha_i^T z - b_i} \right)^2 \leq 1. (1+\nu)^2 w^T (\nabla^2 F(\hat{x}))^{-1} w$$

de onde conclui-se o resultado.

### LEMA 3.7

Seja  $w \in \mathfrak{R}^n$ , um ponto fixo, e  $0 < \nu < 1$ . Seja  $\hat{x} \in \mathfrak{R}^n$  um ponto tal que  $\nabla F(\hat{x}) = \hat{t} w$  para algum  $\hat{t}$ . Seja  $\bar{x} = x(\bar{t})$  um ponto sobre a trajetória  $x = x(t)$  definida implicitamente por  $\nabla F(x) = tw$ , tal  $\bar{x}$  que não está no interior do conjunto  $\sum(\hat{x}, \nu)$ . Então

$$|\bar{t} - \hat{t}| \geq \frac{\nu - \frac{\nu^2}{2}}{(1 + \nu)\sqrt{w^T(\nabla^2 F(\hat{x}))^{-1}w}}$$

Prova: (ver Fig. 3.2)

Seja  $\tilde{x} = x(\tilde{t})$  o primeiro ponto sobre a trajetória  $x=x(t)$ , tal que  $\tilde{x}$  está na fronteira de  $\sum(\hat{x}, \nu)$  quando nos movimentamos desde  $\hat{x}$  ate  $\bar{x}$

(o ponto  $x(\bar{t})$ , existe pois a trajetória  $x = x(t)$  é contínua).

Suponhamos que  $\tilde{x} \neq \bar{x}$  (ver Fig. 3.2).

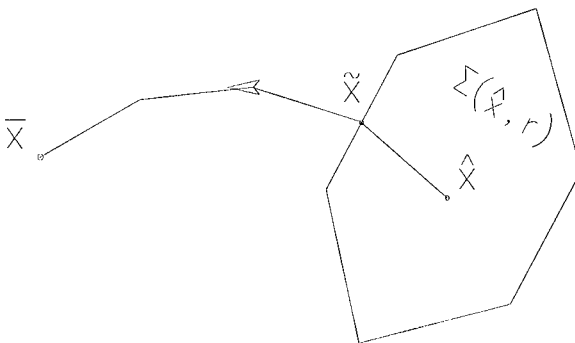


Fig. 3.2

observar que todos os pontos sobre a trajetória entre  $\hat{x}$  e  $\tilde{x}$  estão em  $\sum(\hat{x}, \nu)$ . Como  $\tilde{x}$  pertence a fronteira de  $\sum(\hat{x}, \nu)$ , então existe algum índice

$j, 1 \leq j \leq m$  tal que  $\left| \frac{\alpha_j^T(\tilde{x} - \hat{x})}{\alpha_j^T \hat{x} - b_j} \right| = \nu$  assim

$$\frac{\alpha_j^T(\tilde{x} - b_j)}{\alpha_j^T \hat{x} - b_j} - 1 = \frac{(\alpha_j^T \tilde{x} - b_j) + (b_j - \alpha_j^T \hat{x})}{\alpha_j^T \hat{x} - b_j} = \frac{\alpha_j^T(\tilde{x} - \hat{x})}{\alpha_j^T \hat{x} - b_j} = \pm \nu$$

o que da  $\frac{\alpha_j^T(\tilde{x} - b_j)}{\alpha_j^T \hat{x} - b_j} = 1 \pm \nu$  para  $0 < \nu < 1$

Sabemos que para  $\lambda \in (-1,1)$ ,

$$\ln(1+\lambda) = \lambda - \left(\frac{\lambda^2}{2} - \frac{\lambda^3}{3}\right) - \left(\frac{\lambda^4}{4} - \frac{\lambda^5}{5}\right) - \dots$$

com os termos entre parenteses positivos

Portanto tomando  $\lambda = \pm \nu \in (-1,1)$ , teremos que

$$\left| \ln \left( \frac{\alpha_j^T(\tilde{x} - b_j)}{\alpha_j^T \hat{x} - b_j} \right) \right| = |\ln(1 \pm \nu)| \geq \nu - \frac{\nu^2}{2}$$

Por outro lado,

$$\left| \ln \left( \frac{\alpha_j^T(\tilde{x} - b_j)}{\alpha_j^T \hat{x} - b_j} \right) \right| = \left| \int_{\hat{t}}^{\tilde{t}} \frac{\alpha_j^T \dot{x}(t)}{\alpha_j^T x(t) - b_j} dt \right| \leq \int_{\hat{t}}^{\tilde{t}} \left| \frac{\alpha_j^T \dot{x}(t)}{\alpha_j^T x(t) - b_j} \right| dt$$

e pelo lema 3.6,

$$\left| \frac{\alpha_j^T \dot{x}(t)}{\alpha_j^T x(t) - b_j} \right| \leq (1 + \nu) \sqrt{w^T (\nabla^2 F(\hat{x}))^{-1} w}$$

segue-se que,

$$\nu - \frac{\nu^2}{2} \leq \left| \ln \left( \frac{\alpha_j^T(\tilde{x} - b_j)}{\alpha_j^T \hat{x} - b_j} \right) \right| \leq (1 + \nu) \sqrt{w^T (\nabla^2 F(\hat{x}))^{-1} w} \int_{\hat{t}}^{\tilde{t}} dt$$

portanto,

$$\frac{\nu - \frac{\nu^2}{2}}{(1 + \nu) \sqrt{w^T (\nabla^2 F(\hat{x}))^{-1} w}} \leq |\tilde{t} - \hat{t}| \leq |\bar{t} - \hat{t}|.$$

-

### LEMA 3.8

Seja  $z \in \text{int.P}$  fixo, consideremos a trajetória  $x = x(t)$ , definida implicitamente pela equação  $\nabla F(x) = t \nabla F(z)$ . Se  $\rho \leq 0,16$  e  $\Psi(z) \leq 2\rho^2$ , então  $x(t) \in \sum (z, \sqrt{\rho})$  para  $t \in [0,1]$ . Em particular o centro analítico  $w$  de  $P$ ,  $w \in \sum (z, \sqrt{\rho})$ .

### Prova

Observar que  $z$  e o centro analítico  $w$  pertencem a trajetória  $x = x(t)$ ,

pois,  $\nabla F(z) = 1 \cdot \nabla F(z)$  e  $\nabla F(w) = \bar{\mathbf{0}} = 0 \cdot \nabla F(z)$ . Assim  $x(1) = z$  e  $x(0) = w$ .

Suponhamos que exista algum  $\bar{t} \in [0,1]$  tal que  $\bar{x} = x(\bar{t}) \in \sum (z, \sqrt{\rho})$  então por lema 3,7 com  $\hat{x}=z$ ,  $\hat{t}=1$ ,  $\bar{x} = x(\bar{t})$ ,  $r = \sqrt{\rho} < 1$

$$1 \geq |\bar{t}-1| \geq \frac{\sqrt{\rho} - \frac{\rho}{2}}{(1+\sqrt{\rho})\sqrt{\nabla F(z)^T (\nabla^2 F(z))^{-1} \nabla F(z)}}$$

logo,

$$1 \geq \frac{\sqrt{\rho} - \frac{\rho}{2}}{(1+\sqrt{\rho})} \frac{1}{\sqrt{\Psi(z)}}$$

assim,

$$\Psi(z) \geq \frac{\sqrt{\rho} - \frac{\rho}{2}}{(1+\sqrt{\rho})} > 2\rho^2, \text{ pois } \rho \leq 0,16.$$

o que contradiz a condição  $\Psi(z) \leq 2\rho^2$  dada na hipótese.

Portanto  $x(t) \in \sum (z, \sqrt{\rho})$  para cada  $t \in [0,1]$ . Em particular o centro analítico

$$w = x(0) \in \sum (z, \sqrt{\rho}).$$

Agora mostraremos o resultado principal desta seção que servirá para definir um centro analítico aproximado de um polítopo.

### LEMA 3.9

Seja  $\rho \leq 0,16$  e  $z \in \text{int.P}$  tal que  $\Psi(z) \leq 2\rho^2$ , e  $w$  o centro analítico exato do polítopo  $P$ . Então  $\frac{1}{2}(1-\sqrt{\rho})^2 \Psi(z) \leq F(z) - F(w) \leq \frac{1}{2}(1+\sqrt{\rho})^2 \Psi(z)$

### Prova

Seja  $x = x(t)$  a trajetória definida implicitamente pela equação

$$\nabla F(x) = t\nabla F(z).$$

Temos que

$$F(z) - F(w) = \int_{x=w}^{x=z} \nabla F(x) dx = \int_{t=0}^{t=1} \nabla F(x(t)) \dot{x}(t) dt$$

e como

$$\dot{x}(t) = (\nabla^2 F(x(t)))^{-1} \nabla F(z), \nabla F(x(t)) = t\nabla F(z)$$

resulta que

$$F(z) - F(w) = \int_0^1 t \nabla F(z) (\nabla^2 F(x(t)))^{-1} \nabla F(z) dt$$

Sabemos pelo lema 3.8 que,  $x(t) \in \sum (z, \sqrt{\rho})$  para cada  $t \in [0,1]$ , e como

$\sqrt{\rho} < 1$ ,  $z \in \text{int.P}$  pelo corolário 3.1 para  $\xi = \nabla f(z)$  tem-se que,

-

isto é

$$(1 - \sqrt{\rho})^2 (\nabla F(z))^T (\nabla^2 F(z))^{-1} \nabla F(z) \leq (\nabla F(z))^T (\nabla^2 F(x(t)))^{-1} \nabla F(z) \leq$$

$$(1 + \sqrt{\rho})^2 (\nabla F(z))^T (\nabla^2 F(z))^{-1} \nabla F(z)$$

portanto,

$$(1 - \sqrt{\rho})^2 \Psi(z) \int_0^1 t dt \leq \int_0^1 t \nabla F(z)^T (\nabla^2 F(x(t)))^{-1} \nabla F(z) dt \leq$$

$$(1 + \sqrt{\rho})^2 \Psi(z) \int_0^1 t dt$$

o que dá,

$$- \quad \frac{1}{2} (1 - \sqrt{\rho})^2 \Psi(z) \leq F(z) - F(w) \leq \frac{1}{2} (1 + \sqrt{\rho})^2 \Psi(z).$$

### LEMA 3.10

Seja  $z \in \text{int.P}$ ,  $x = x(t)$  a trajetória definida implicitamente pela equação  $\nabla F(x) = t \nabla F(z)$ ,  $w$  o centro analítico exato do polítopo  $P$ . Se  $v \leq 0,008$  e  $F(z) - F(w) \leq v$ ,

então  $x(t) \in \sum (z, 5\sqrt{v})$  para todo  $t \in [0,1]$ . Em particular  $x(0) = w$

$\in \sum (z, 5\sqrt{v})$ .

### Prova

Suponhamos que toda trajetória  $x(t)$  para  $t \in [0,1]$  não esteja contida em  $\sum (z, 5\sqrt{v})$ , então sendo ela contínua, corta a fronteira de



$\sum (z, 5\sqrt{v})$ . Suponhamos que  $\bar{x} = x(\bar{t})$  com  $0 \leq \bar{t} \leq 1$  seja o primeiro ponto da trajetória desde  $z$  a  $w$  que pertence a fronteira de  $\sum (z, 5\sqrt{v})$ . Então pelo lema 3.7, com  $\hat{t}=1, r = 5\sqrt{v}, \hat{x} = z$ , temos que

$$1 - \bar{t} = |\bar{t} - 1| \geq \frac{5\sqrt{v} - \frac{25v}{2}}{(1 + 5\sqrt{v})\sqrt{\nabla F(z)^T (\nabla^2 F(z))^{-1} \nabla F(z)^T}}$$

isto é,

$$(1 - \bar{t}) \geq \frac{5\sqrt{v} - \frac{25v}{2}}{(1 + 5\sqrt{v})\sqrt{\Psi(z)}}$$

de onde,

$$(1 - \bar{t}) \Psi(z) \geq \left( \frac{5\sqrt{v} - \frac{25v}{2}}{(1 + 5\sqrt{v})} \right)^2 \quad (3.3)$$

Assim,

$$\begin{aligned} F(z) - F(w) &= F(x(1)) - F(x(0)) \geq F(x(1)) - F(x(\bar{t})) && \text{(pois } w \text{ é minimizador)} \\ &= \int_{\bar{t}}^1 t \nabla F(z)^T (\nabla^2 F(x(t)))^{-1} \nabla F(z) dt && \text{(ver lema 3 a 9)} \\ &\geq (1 - 5\sqrt{v})^2 \Psi(z) \int_{\bar{t}}^1 t dt && \text{(pelo corolário 3.1)} \\ &= (1 - 5\sqrt{v})^2 \Psi(z) \frac{(1 - \bar{t})^2}{2} \\ &\geq \frac{(1 - \bar{t})^2}{2} (1 - 5\sqrt{v})^2 \Psi(z) && \text{(pois } \bar{t} \in [0, 1]) \\ &\geq \frac{(1 - 5\sqrt{v})^2}{2} \left( \frac{5\sqrt{v} - \frac{25v}{2}}{(1 + 5\sqrt{v})} \right)^2 > v && \text{(por 3.3)} \\ &&& \text{(pois } v \leq 0,008) \end{aligned}$$

ou seja  $F(z) - F(w) > v$  o que contradiz o que foi afirmado na hipótese, portanto  $x(t)$

$\in \sum (z, 5\sqrt{v})$  para todo  $t \in [0, 1]$ .

O lema que apresentamos a seguir estabelece um resultado análogo ao do lema 3.9 mostrando a relação entre pequeno valor de  $F(z)-F(w)$  (onde  $w$  centro analítico exato) e o valor  $\Psi(z)$ .

**LEMA 3.11**

Seja  $z \in \text{int}.P$ ,  $v \leq 0,008$ ,  $w$  o centro analítico exato de  $P$ . Se  $F(z)-F(w) \leq v$  então

$$\frac{1}{2}(1-5\sqrt{v})^2 \Psi(z) \leq F(z)-F(w) \leq \frac{1}{2}(1+5\sqrt{v})^2 \Psi(z) \quad (3.4)$$

**Prova**

Como na demonstração do lema 3.8 , chegamos a

$$F(z)-F(w) = \int_0^1 t \nabla F(z) (\nabla^2 F(x(t)))^{-1} \nabla F(z) dt$$

pelo lema 3.10 temos que,

$x(t) \in \sum (z, 5\sqrt{v})$  para todo  $t \in [0,1]$ , então pelo corolário 6 com  $\alpha = 5\sqrt{v}$ ,  $\xi = \nabla F(z)$  segue o resultado (3.4).

O próximo lema, mostra que quando  $\Psi(z) < 1$ , o centro analítico exato  $w$  está num pequeno elipsóide centrado em  $z$ .

**LEMA 3.12**

Seja  $z \in \text{int}.P$ , tal que  $\Psi(z) < 1$ , e  $w$  o centro analítico exato do polítopo  $P$ , e seja

$$\rho = \frac{\sqrt{\Psi(z)}}{1-\sqrt{\Psi(z)}} . \text{ Então } \alpha = \sqrt{(w-z)^T (\nabla^2 F(z))^{-1} (w-z)} , \alpha \leq \rho, \text{ isto é}$$

$$w \in E((\nabla^2 F(z), z, \rho))$$

**Prova**

Consideremos o segmento da reta

$$r(s) = z + s(w-z), \quad s \in [0,1]$$

então

$$(r(s)-z)^T (\nabla^2 F(z))(r(s)-z) = s^2 (w-z)^T \nabla^2 F(z)(w-z) = s^2 \alpha^2 .$$

Como  $r(s) \in \text{int}.(P)$  (pois  $P$  é convexo) e  $r(s) \in E((\nabla^2 F(z), z, \alpha))$

então pelo lema 3,3  $r(s) \in \sum (z, \alpha)$  para cada  $s \in [0,1]$ .

e como no lema 3.4, a desigualdade do lado esquerdo é válido para todo  $\alpha \in \mathfrak{R}$  ( $\alpha < 1, \alpha \geq 1$ ), por tanto é válido para  $s \alpha \in \mathfrak{R}$  temos que para cada  $\xi \in \mathfrak{R}^n$

$$\frac{1}{(1+s\alpha)^2} \xi^T \nabla F(z) \xi \leq \xi^T \nabla^2 F(r(s)) \xi, \text{ para } s \in [0,1] \quad (3.5)$$

Seja

$$f(s) = F(r(s)),$$

então para  $s \in [0,1]$

$$f'(s) = (\nabla F(r(s)))^T \cdot r'(s) = (\nabla F(r(s)))^T (w-z)$$

$$f''(s) = (w-z)^T \nabla^2 F(r(s))(w-z)$$

e como  $\nabla F(w) = \bar{0}$  (pois é minimizador de F)

podemos escrever

$$-\nabla F(z)^T (w-z) = \nabla F(w)^T (w-z) - \nabla F(z)^T (w-z)$$

$$= f(1) - f(0) = \int_0^1 f'(s) ds$$

$$= \int_0^1 (w-z)^T \nabla^2 F(r(s))(w-z) ds$$

o que da

$$\nabla F(z)^T (w-z) + \int_0^1 (w-z)^T \nabla^2 F(r(s))(w-z) ds = 0 \quad (3.6)$$

Por outro lado sabemos que

$$\max_{x \in (\nabla^2 F(z), z, \alpha)} -\nabla F(z)^T (x-z) = \alpha \sqrt{(-\nabla F(z))^T (\nabla^2 F(z))^{-1} (-\nabla F(z))} = \alpha \sqrt{\Psi(z)},$$

então para

$$w = r(1) \in E(\nabla^2 F(z), z, \alpha) \text{ (tomando } s=1)$$

$$(-\nabla F(z))^T (w-z) \leq \alpha \sqrt{\Psi(z)}$$

ou seja  $\nabla F(z)^T (w-z) \geq -\alpha \sqrt{\Psi(z)}$ . logo de (3.6) segue que

$$0 \geq -\alpha \sqrt{\Psi(z)} + \int_0^1 (w-z)^T \nabla^2 F(r(s))(w-z) ds = 0$$

pela validade de (3.5) para  $\xi = (w-z)$ , temos que

$$\frac{\alpha^2}{(1+s\alpha)^2} = \frac{1}{(1+s\alpha)^2} (w-z)^T \nabla^2 F(z)(w-z) \leq (w-z)^T \nabla F(r(s))(w-z)$$

logo

$$0 \geq -\alpha \sqrt{\Psi(z)} + \int_0^1 (w-z)^\top \nabla^2 F(r(s))(w-z) ds \geq -\alpha \sqrt{\Psi(z)} + \int_0^1 \frac{\alpha^2}{(1+s\alpha)^2} ds$$

o que dá 
$$0 \geq -\alpha \sqrt{\Psi(z)} + \frac{\alpha^2}{1+s\alpha}$$

isto é  $\frac{\alpha}{1+\alpha} \leq \sqrt{\Psi(z)}$  e como  $1+\alpha > 0$ , e como  $\Psi(z) < 1$

conclui-se que  $\alpha \leq \frac{\sqrt{\Psi(z)}}{1-\sqrt{\Psi(z)}}$ .

### 3.3.5 MEDIDA DE PROXIMIDADE AO CENTRO ANALÍTICO

Vimos na seção 3.3.4 que quando  $\Psi(z)$  é pequeno, podemos estimar o valor de  $F(z)-F(w)$ , onde  $w$  é o centro analítico exato do polítopo  $P$ , e que  $w$  está num pequeno elipsóide hessiano com centro  $z$ . Assim, usando o lema 3.9, podemos definir uma medida de proximidade ao centro analítico de  $P$ .

**DEFINIÇÃO 3.1** Seja  $w$  o centro analítico (exato) do polítopo  $P$ , e  $z \in \text{int}P$  um ponto tal que  $\Psi(z) \leq 2\rho^2$ , com  $\rho \leq 0,016$ . Então, dizemos que  $z$  é um centro analítico  $\Psi$ -aproximado de  $P$ , e que  $\Psi$  da uma  $\Psi$ -medida ao centro analítico.

Tendo em conta que  $\nabla F(x) = -\mathbf{A}^T \mathbf{S}^{-1} \mathbf{e}$  e que  $\nabla^2 F(x) = +\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A}$  onde  $\mathbf{S}^{-1}$  é a matriz diagonal com elementos na diagonal principal  $\frac{1}{s_i}$ ,  $s_i = \alpha_i^T x - b_i > 0$

podemos expressar

$$\begin{aligned} \Psi(z) &= \nabla F(x)^T (\nabla^2 F(x))^{-1} \nabla F(x) \text{ de outra maneira:} \\ \Psi(x) &= (\mathbf{A}^T \mathbf{S}^{-1} \mathbf{e})^T (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{S}^{-1} \mathbf{e}) \\ &= \mathbf{e}^T (\mathbf{S}^{-1} \mathbf{A}) (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A})^{-1} (\mathbf{S}^{-1} \mathbf{A})^T \mathbf{e} \end{aligned} \quad (3.7)$$

fazendo  $\mathbf{g} = \mathbf{A}^T \mathbf{S}^{-1} \mathbf{e}$  e  $\mathbf{H} = \mathbf{A}^T \mathbf{S}^{-2} \mathbf{A}$  tem-se que

$$\Psi(z) = \mathbf{g}^T \mathbf{H}^{-1} \mathbf{g} = \|\mathbf{g}\|_{\mathbf{H}}^2$$

onde  $\|\alpha\|_{\mathbf{H}} = \sqrt{\alpha^T \mathbf{H}^{-1} \alpha}$  é a norma induzida pela matriz simétrica positiva

definida  $\mathbf{H}$ .

A função  $\Psi(z)$  nos permite comparar a  $\Psi$ -medida ao centro analítico, com outras medidas estudadas, por exemplo comparar com a  $\delta$ -medida usada por Roos e Vial [34] para ter centros analíticos aproximados em LP e por D. Hertog [5] para PNL.

Na continuação comentamos brevemente sobre a  $\delta$ -medida.

$\delta$ - medida ao centro analítico:

Sabemos que o centro analítico  $w$  é o único minimizador da função barreira

logarítmica  $F(x) = - \sum_{i=1}^m \ln (\alpha_i^T x - b_i)$ , e isto acontece se, e somente se  $\nabla F(w) =$

$0$ , ou seja quando  $\mathbf{A}^T \mathbf{S}^{-1} \mathbf{e} = 0$  com  $\mathbf{S}^{-1} = \text{diag}(1/s_i)$ ,  $s_i = \alpha_i^T w - b_i > 0$

Se  $\mathbf{y} = \mathbf{S}^{-1} \mathbf{e}$ , tem-se que  $\mathbf{A}^T \mathbf{y} = 0$ ,  $\mathbf{Y} \mathbf{s} = \mathbf{e}$ ,  $\mathbf{y} > 0$

onde  $\mathbf{Y} = \text{diag}(y_i)$ ,  $\mathbf{s} = \mathbf{A}w - \mathbf{b} > 0$ ,

Portanto, as condições necessárias e suficientes de otimalidade para que  $x=w$  seja o centro analítico de  $P$  são:

$$(C.O) \begin{cases} \mathbf{A}x - \mathbf{s} = \mathbf{b} \\ \mathbf{A}^T \mathbf{y} = 0 \\ \mathbf{Y} \mathbf{s} = \mathbf{e} \\ \mathbf{y} > 0, \mathbf{s} > 0 \end{cases}$$

estas condições expressam primal e dual viabilidade, mais uma condição de centramento dado por  $\mathbf{Y} \mathbf{s} = \mathbf{e}$  ou equivalentemente  $\mathbf{S} \mathbf{y} = \mathbf{e}$ , onde  $\mathbf{S} = \text{diag}(s_i)$  com  $\mathbf{s} = \mathbf{A}x - \mathbf{b} > 0$ .

Usando a norma euclídea  $\| \cdot \|$ , definimos o número

$$\delta(x) = \min_{\mathbf{y}} \{ \| \mathbf{S} \mathbf{y} - \mathbf{e} \| : \mathbf{A}^T \mathbf{y} = 0 \} \quad (3.7)$$

Seja  $\mathbf{Y}(x) = \mathbf{S}^{-1} (\mathbf{I} - \mathbf{S}^{-1} \mathbf{A} (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A}) \mathbf{A}^T \mathbf{S}^{-1}) \mathbf{e}$

pode-se mostrar que  $\mathbf{y}(x)$  é a solução do problema (3.8),

logo

$$\delta(x) = \| \mathbf{S} \mathbf{y}(x) - \mathbf{e} \| \text{ com } \mathbf{A}^T \mathbf{y}(x) = 0.$$

Se  $\mathbf{y}(x) > 0$ , o valor  $\delta(x)$  dá uma medida de proximidade do ponto  $x$  ao centro analítico do polítopo  $P$ , o que podemos chamar a  $\delta$ -medida ao centro analítico.

Por outro lado, temos que

$$\begin{aligned} \delta^2(x) &= \| (\mathbf{e} - \mathbf{S}^{-1} \mathbf{A} (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{S}^{-1} \mathbf{e}) - \mathbf{e} \|^2 \\ &= \| \mathbf{S}^{-1} \mathbf{A} (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{S}^{-1} \mathbf{e} \|^2 \\ &= [\mathbf{S}^{-1} \mathbf{A} (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{S}^{-1} \mathbf{e}]^T [\mathbf{S}^{-1} \mathbf{A} (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{S}^{-1} \mathbf{e}] \\ &= [\mathbf{e}^T \mathbf{S}^{-1} \mathbf{A} (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{S}^{-1}]^T [\mathbf{S}^{-1} \mathbf{A} (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{S}^{-1} \mathbf{e}] \\ &= \mathbf{e}^T \mathbf{S}^{-1} \mathbf{A} (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A}) (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{S}^{-1} \mathbf{e} \end{aligned}$$

$$\begin{aligned}
&= \mathbf{e}^T \mathbf{S}^{-1} \mathbf{A} (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{S}^{-1} \mathbf{e} \\
&= (\mathbf{A}^T \mathbf{S}^{-1} \mathbf{e})^T (\mathbf{A}^T \mathbf{S}^{-2} \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{S}^{-1} \mathbf{e})
\end{aligned}$$

e logo usando a definição de  $\Psi(\mathbf{x})$ , temos que

$$\delta^2(\mathbf{x}) = \Psi(\mathbf{x}) \quad \text{ou} \quad \delta(\mathbf{x}) = \sqrt{\Psi(\mathbf{x})}$$

portanto a medida  $\delta$  e a  $\Psi$ -medida estão relacionadas, então pode-se usar para a  $\Psi$ -medida os resultados obtidos para a  $\delta$ -medida. (ver Fig. 2.1)

# CAPÍTULO 4

## O MÉTODO DE ELIPSÓIDES E GENERALIDADES SOBRE MÉTODOS DE PLANOS DE CORTE COM ELIMINAÇÃO DE CORTES

### 4.1 INTRODUÇÃO

Neste Capítulo comentaremos sobre algoritmos de planos de corte-centro analítico que contêm alguma estratégia de eliminação de cortes.

Em particular comentaremos em forma mais detalhada o algoritmo de D. Atkinson-M. Vaidya[1] para o problema de viabilidade convexa que usa uma estratégia de eliminação de corte para provar a convergência polinomial do algoritmo.

Também neste mesmo capítulo apresentaremos sobre o método de elipsóides, que usaremos para construir nosso algoritmo para o problema de viabilidade convexa com uma estratégia de identificação de cortes redundantes, o que será exposto no capítulo 5.

### 4.2 ALGORÍTMOS DE PLANOS DE CORTE-CENTRO ANALÍTICO COM ELIMINAÇÃO DE CORTES.

Quando são incorporados no algoritmo de planos de corte-centro analítico, estratégias de eliminação de cortes, as demonstrações de convergência e complexidade do algoritmo tornam-se mais árduas, pois tem que ser analisado o efeito da eliminação de cortes nas diferentes etapas do algoritmo. Também a rapidez de identificação de cortes não comprometidos na solução ótima é crucial, se a identificação é feita somente na última fase do algoritmo, neste caso a redução na computação é insignificante. Sendo por estas dificuldades que não existem na literatura um amplo número de trabalhos relativos ao processo de eliminação de cortes.

Entre as publicações que apresentam estratégias de eliminação de cortes ressaltamos [2] onde é implementado um algoritmo de plano de corte-centro analítico para a solução de um problema convexo diferenciável. É analisado o comportamento total do algoritmo e os efeitos de várias estratégias de implementação, baseados na eliminação de restrições redundantes. O teste para descartar planos de cortes redundantes é feito usando elipsóides  $E_{int}$  e  $E_{ext}$  que providenciam uma aproximação interna e externa, do polítopo localizador  $F(Z_u)$  de soluções ótimas ( $E_{int} \subset F(Z_u) \subset \rho^{-1} E_{ext}$ ,  $\rho > 1$ ), elipsóides que são centrados num centro analítico aproximado de  $F(Z_u)$ . Um hiperplano  $P_t$  que é um hiperplano suporte ao epígrafo da função objetivo ou é suporte ao conjunto de restrições, é eliminado, se  $P_t$  não intersecta ao elipsóide  $\rho^{-1} E_{ext}$ . A escolha de  $\rho > 1$  não tem nenhuma base teórica, é escolhido usando uma heurística apropriada. É testado varias versões do algoritmo, fazendo comparações de eficiência computacional entre elas.



Também cabe mencionar que em [6], [7] é estudado os efeitos de adicionar e eliminar restrições no método barreira logarítmica para programação linear (PL). O algoritmo é analisado teoricamente é provado de ser polinomial, mostrando que o algoritmo necessita  $O(q^*, L)$  Newton iterações para encontrar uma solução aproximada do problema linear, onde  $q^*$  denota o máximo número de restrições no subsistema linear durante todo o processo. Estas estratégias, chamadas “Build-Up and Down Strategies”, são estendidas em [8] para programação convexa não diferenciável que considerando subsequentes relaxações lineares do problema convexo (PC), (como no caso de [6], [7]) são efetuadas buscas lineares na direção de Newton da função barreira logarítmica da (PL) relaxação. Se o novo iterado viola uma restrição ou é ‘fechado’ a fronteira, então detém-se no prévio ponto, é um novo hiperplano suporte é adicionado para melhorar a aproximação. Se o corrente iterado é centrado, ou seja está próximo a trajetória central e o valor da folga de uma restrição neste iterado é suficientemente grande, então com uma cuidadosa regra esta restrição é removida do corrente sistema, pois é provável que esta restrição não é influente na solução ótima. Depois de eliminar restrições redundantes procede-se a recentrar se for necessário. Não foi conseguida uma rigorosa demonstração para a convergência do algoritmo sendo o maior problema, mostrar que um ciclo interno é finito, isto é provar que depois de um número finito de iterações se conseguirá centrar novamente.

A seguir apresentamos o algoritmo de Atkinson-Vaidya [1] que contém um critério de eliminação de cortes.

### 4.3 O ALGORÍTMO DE ATKINSON-VAIDYA

Atkinson-Vaidya apresentam em [1] um algoritmo para o problema de viabilidade convexa com um critério de eliminação de cortes chamados ‘não importantes’.

No algoritmo assume-se que o conjunto convexo  $S$ , ( $S \subseteq \mathcal{R}^n$ ) está contido na caixa

$$B = \{x \in \mathcal{R}^n : -2^L \leq x_i \leq 2^L, 1 \leq i \leq n\} = [-2^L, 2^L]^n$$

sendo  $L$  uma constante fixada tal que  $L \geq \log_2 n$  e que quando  $S$  é não vazia,  $S$  contém uma bola de raio  $2^{-L}$ .

Fazendo sucessivos cortes através da caixa  $B$  consegue-se uma seqüência de polítopos  $P$  contendo  $S$ , cujos centros analíticos (ou  $\Psi$ -aproximados centros analíticos) (ver capítulo 3) servem como pontos testes. Um oráculo para  $S$  recebe como entrada um centro analítico  $\Psi$ -aproximado  $z$  de  $P$ , testa  $z$ , se  $z \in S$  o algoritmo pára e o problema de viabilidade está resolvido, em caso contrário se  $z \notin S$  o oráculo retorna um vetor  $c$  tal que  $\{x \in \mathcal{R}^n : c^T x \geq c^T z\} \supset S$ . Nesta etapa do

algoritmo é quando é criado critérios para determinar se cada hiperplano  $\mathcal{A}_i^T x = b_i$

da definição do plítopo  $P$ , que não seja da caixa  $B$  é considerado “importante” ou não. No caso que seja não importante o hiperplano é eliminado da definição de  $P$  dando lugar a um novo polítopo limitado  $P_d$  cujo centro analítico  $z_d$  (ou  $\Psi$ -

aproximado) é o novo ponto teste. Se o hiperplano  $\mathcal{A}_i^T x = b_i$  é ainda importante é mantido na definição de  $P$ . Este procedimento é feito até

hiperplanos considerados não importantes, e é então quando é construído um novo hiperplano  $c^T x = \beta$  é adicionado ao sistema que define P dando lugar a um novo polítopo  $P_a = P \cap \{x \in \mathbb{R}^n : c^T x \geq \beta\}$  e o novo ponto teste é o centro analítico (ou aproximado)  $z_a$  de  $P_a$ .

Os critérios de parada do algoritmo são elaborados baseados em estratégias de volume e largura dos polítopos. Estes argumentos tem um rol importante na demonstração da convergência do algoritmo conjuntamente com a eliminação de hiperplanos não importantes. É mostrado que se algoritmo experimenta em adicionar hiperplanos mais do que um limitante superior  $\nu nL$  onde  $\nu > 0$  é uma constante adequada então o volume de P torna-se muito pequeno para conter uma bola de raio  $2^{-L}$ . Portanto a região S é declarada vazia. Em caso contrário, se o número de hiperplanos não cresce mais longe do que  $\nu nL$ , então a menos que se haja encontrado previamente um ponto em S dentro de  $O(nL^2)$  iterações o polítopo torna-se muito estreita (raso) em alguma direção para conter uma bola de raio  $2^{-L}$ , então S é declarada vazia. É provado a convergência do algoritmo dentro de  $O(nL^2)$  iterações com um total  $O((T+nL+n^3)nL^2)$  operações aritméticas, onde T é o número de operações aritméticas requeridas em cada iteração.

Para apresentar o Algoritmo de Atkinson-Vaidya (A.A.V) precisamos dar algumas definições prévias, o que daremos a seguir. ( ver [1] )

Seja

$$P = \{x \in \mathbb{R}^n : Ax \geq b\} = \{x \in \mathbb{R}^n : \mathcal{A}_i^T x \geq b_i, 1 \leq i \leq m\}$$

um polítopo, onde,  $A \in \mathbb{R}^{m \times n}$  e  $b \in \mathbb{R}^m$ , e

$$F(x) = - \sum_{i=1}^m \ln(\mathcal{A}_i^T x - b_i)$$

é a função barreira logarítmica sobre P,  $w$  o centro analítico (exato) de P, e  $z$  um centro analítico  $\Psi$ -aproximado de P

Denota-se ao hiperplano  $\mathcal{A}_i^T x = b_i$  pelo par ordenado  $(\mathcal{A}_i, b_i)$ . Correspondendo a

cada plano  $(\mathcal{A}_i, b_i)$  define-se uma quantidade  $K(\mathcal{A}_i, b_i) = \mathcal{A}_i^T z - b_i$ , que quando é

considerado  $\|\mathcal{A}_i\| = 1$  vem a ser a distância de  $z$  ao hiperplano  $(\mathcal{A}_i, b_i)$ .

Para cada  $x \in \text{int.P}$  define-se os valores

$$\sigma_i(x) = \frac{\mathcal{A}_i^T (\nabla^2 F(x))^{-1} \mathcal{A}_i}{(\mathcal{A}_i^T x - b_i)^2} \text{ para } 1 \leq i \leq m$$

verifica-se que  $0 < \sigma_i(x) \leq 1$

No algoritmo (A.A.V) são elaborados critérios para eliminação de hiperplanos. Os  $2n$  hiperplanos que formam a caixa B permanecem intocáveis não são nunca eliminados, formando sempre parte dos hiperplanos que definem o polítopo P gerados no algoritmo, para assegurar que os polítopos sejam sempre limitados.

Assim, se  $m$  hiperplanos são adicionados aos  $2n$  da caixa para gerar o polítopo  $P$ , então  $\bar{m} = 2n + m$  é o número total de hiperplanos de  $P$ .

Define-se então

$$\mu_i(z) = \begin{cases} 1 & \text{para } 1 \leq i \leq 2n \\ \frac{\alpha_i^T z - b_i}{K(\alpha_i, b_i)} & \text{para } 2n+1 \leq i \leq \bar{m} \end{cases}$$

$\mu_i(z) = 1$  para os hiperplanos da caixa assegura que os hiperplanos da caixa não são eliminados nunca no algoritmo.

Os subscritos “d” e “a” significaram ápos que ‘um hiperplano é eliminado’ ou após que ‘um hiperplano é adicionado’ respectivamente.

## ALGORÍTMO (A.A.V)

### 1. Inicialização

$$P = B, w = 0, z = 0, K(\alpha_i, b_i) = 2^L \text{ para } 1 \leq i \leq 2n$$

### 2. Iteração Genérica

Caso 1:  $\max_{1 \leq i \leq \bar{m}} \mu_i(z) > 2$

**Seja  $j$**  um índice  $j$  tal que  $\mu_j(z) > 2$

**Se**  $\sigma_j(z) < 0,04$

**então** subcaso 1.1

subcaso 1.1 O  $j$ -éssimo hiperplano não é importante. Eliminar

$(\alpha_j, b_j)$  do conjunto de hiperplanos definindo  $P$ , produzindo

um novo polítopo  $P_d$ . Tomar  $O(1)$  passos de Newton para conseguir uma adequada aproximação  $z_d$  ao novo centro analítico  $w_d$  de  $P_d$ .

**de outro modo** se  $\sigma_j(z) \geq 0,04$ ; então subcaso 1.2

Subcaso 1.2 Seja  $(\alpha_j, b_j)$  um hiperplano tal que  $\mu_j(z) > 2$ .

O  $j$ -éssimo hiperplano é ainda importante simplesmente atualizar

$$K(\alpha_j, b_j) \text{ como } K(\alpha_j, b_j) = \alpha_j^T z - b_j$$

Caso 2:  $\max_{1 \leq i \leq \bar{m}} \mu_i(z) \leq 2$

Chamar ao oráculo com  $z$  como ponto teste.

**Se** O oráculo retorna ‘sim’ (ou seja  $z \in S$ )

**então** parar.

**Se não** ( $z \notin S$ ) então

o oráculo retorna um vetor  $c$  tal que

$$S \subset \{x \in \mathbb{R}^n : c^T x \geq c^T z\}$$

Calcular  $\beta$  tal que  $\gamma^2 = \frac{c^T (\nabla^2 F(z))^{-1} c}{(c^T z - \beta)^2} = 1/16$

Adicionar  $(c, \beta)$  ao conjunto de hiperplanos definindo  $P$  isto é fazer

$$P_a = P \cap \{x \in \mathbb{R}^n : c^T x \geq \beta\}$$

Tomar  $O(1)$  passos de Newton para conseguir uma adequada aproximação  $z_a$  ao novo centro analítico  $w_a$  de  $P_a$ ,

Fazer  $K(c, \beta) = c^T z_a - \beta$ .

### 3. Condição de Parada

condição 1  $\bar{m} \geq \nu nL$ ,  $\nu$  constante adequada  $\nu \geq 1$

condição 2  $\min_{1 \leq i \leq \bar{m}} \{a_i^T x - b_i\} \leq \frac{2^{-(L+1)}}{\nu nL}$

Se a condição de parada 1 ou 2 é satisfeita, então é mostrado que a região  $S$  é vazia.

A demonstração da convergência do algoritmo (A.A.V) baseia-se em dois teoremas fundamentais que dependem do fato que hiperplanos não importantes são eliminados. Estes dois teoremas demonstrados em [1] estabelecem dois fatos muito importantes:

Uma, que a função barreira logarítmica normalizada

$$N(x) = - \sum_{i=1}^{\bar{m}} \ln \left( \frac{a_i^T x - b_i}{K(a_i, b_i)} \right)$$

verifica  $N_{\text{nov}}(w_{\text{nov}}) - N_{\text{ant}}(w_{\text{ant}}) > \theta = 0,01$  onde  $w_{\text{ant}}$  e  $w_{\text{nov}}$  são centros analíticos do anterior e do novo polítopo gerado respectivamente, e a outra que o determinante da hessiana da função barreira cresce exponencialmente com o número de hiperplanos isto é que

$$\det(\nabla^2 F(x)) > \frac{(1,01)^m}{2^{n(1+2L)}}$$

Enunciaremos a seguir os teoremas que estabelecem a convergência do algoritmo (A.A.V), suas demonstrações estão contidas em [1].

**TEOREMA 4.1** Existe uma constante  $\nu > 0$ , independente de  $m$ ,  $n$  e  $L$  tal que se o número total de hiperplanos é  $\bar{m} = \nu nL$ , então a região  $S$  é vazia.

A demonstração do teorema 4.1 é baseado no fato que o determinante da hessiana da barreira logarítmica cresce exponencialmente com o número de hiperplanos. O teorema 4.1 demonstra que a condição de parada dada pela condição 1 é legítima.

**TEOREMA 4.2** Se o algoritmo (A.A.V) não encontra antes um ponto em  $S$  ou pára pela condição 1 de parada, então dentro de  $O(nL^2)$  iterações a condição 2 de parada

deve ser encontrado. Se a condição 2 de parada é encontrada, então a região viável  $S$  é vazia.

A demonstração do teorema 4.2 usa o fato que a função barreira normalizada cresce no sentido indicado acima. A validade da condição 2 de parada está justificada pelo teorema 4.2, pois quando a condição 2, de parada é encontrada, o polítopo é muito estreita e não pode conter uma bola de raio  $2^{-L}$ , portanto declara-se que a região viável é vazia.

## 4.4 O MÉTODO DE ELIPSÓIDES

Nesta seção apresentamos as idéias básicas do método de elipsóides. Depois que o método simplex for classificado de complexidade exponencial, a maior questão teórica era descobrir a existência de um algoritmo de tempo polinomial para programação linear PL. Uma resposta afirmativa foi finalmente provada por Khachiyan em 1979 [25]. Ele mostrou como pode-se adaptar o método elipsoidal para programação convexa (da qual PL é um caso especial) desenvolvido por Shor, Yudin e Nemirovsky, para programação linear com um tempo de complexidade polinomial. Mais precisamente, Yudin e Nemirovsky mostraram que o método elipsoidal relacionado ao trabalho de Shor's trabalha com uma solução aproximada dentro de uma  $\epsilon > 0$  tolerância; em um número de iterações a qual é polinomial no tamanho da entrada e  $\log(1/\epsilon)$  Khachiyan mostrou que quando o método é aplicado a problemas de programação linear com coeficientes inteiros ainda pode-se ter uma solução exata num tempo polinomial.

É mostrado que o método de elipsóides é eficiente pelas suas implicações teóricas, mas que é ineficiente na prática, não é competitivo com o método simplex para cálculos práticos. Contudo, o método de elipsóides tem uma fundamental potência teórica, pois é uma ferramenta elegante para mostrar o tempo polinomial de solução de muitos problemas de otimização geométrica e combinatorial.

Descreveremos o método de elipsóides para o caso especial de encontrar um ponto num polítopo de dimensão plena, não vazia, definida explicitamente por um sistema de desigualdades. Ou seja, o método dos elipsóides resolve o problema de viabilidade convexa linear. ( ver [23] )

Se  $A$  é uma matriz de ordem  $n \times n$  definida positiva, definimos o elipsóide

$$E(A,a) = \{x \in \mathbb{R}^n ; (x-a)^T A^{-1} (x-a) \leq 1\} = \{x \in \mathbb{R}^n \mid \|x-a\|_A \leq 1\}$$

de centro  $a$  e raio igual a 1. Então, o elipsóide  $E(A,a)$  é uma bola unitária ao redor de  $a$  no espaço  $\mathbb{R}^n$  provido da norma  $\| \cdot \|_A$ . Em particular a bola unitária  $S(0,1)$  ao redor de zero (na norma euclideana) é o elipsóide  $E(I,0)$ .

É conhecido que uma região convexa  $K$  limitada, está contido num único elipsóide de menor volume (ver por ex. Danzer, Grünbaun - Klee (1963) p. 139). O elipsóide de menor volume contendo  $a$ , região convexa  $K$  é chamada de elipsóide de Löwner-John de  $K$ . Em geral, o elipsóide Löwner-John de uma região convexa  $K$  é difícil de calcular. No método de elipsóides e suas variantes, são usados os elipsóides de Löwner-John de certas seções elipsoidais e existem formulas explicitas para conhece-las. O que descrevemos a seguir:

$$\text{Seja } c \in \mathbb{R}^n - \{0\} \text{ e } E^*(A,a,c) = E(A,a) \cap \{x \in \mathbb{R}^n : c^T x \leq c^T a\}$$

assim  $E'(A,a,c)$  é uma seção elipsoidal obtida cortando  $E(A,a)$  pelo centro “a” usando o hiperplano  $\{x \in \mathfrak{R}^n: c^T x = c^T a\}$  (corte central). O elipsóide Löwner-John de  $E'(A,a,c)$  é o elipsóide  $E(A',a')$  dado pelas fórmulas

$$a' = a - \frac{1}{n+1} b$$

$$A' = \frac{n^2}{n^2-1} \left( A - \frac{2}{n+1} b b^T \right)$$

onde  $b = \frac{1}{\sqrt{c^T A c}} A c$

O algoritmo de Khachian faz uso do elipsóide Löwner-John de  $E'(A,a,c)$ . Existem modificações deste método usando elipsóides Löwner-John de outras seções elipsoidais que descrevemos a seguir (ver Bland, Goldfarb-Todd (1981)). Seja  $E(A,a)$  um elipsóide e  $c \in \mathfrak{R}^n - \{0\}$

Então, o hiperplano

$$H = \{x \in \mathfrak{R}^n: c^T x = \gamma\}$$

tem interseção não vazia com  $E(A,a)$  se, e somente se

$$c^T a - \sqrt{c^T A c} \leq \gamma \leq c^T a + \sqrt{c^T A c}$$

ou seja, se e somente se

$$-1 \leq \alpha \leq 1$$

onde,

$$\alpha = \frac{c^T a - \gamma}{\sqrt{c^T A c}}$$

pois,

$$\max\{c^T x : x \in E(A,a)\} = c^T a + \sqrt{c^T A c}$$

$$e \quad \min\{c^T x : x \in E(A,a)\} = c^T a - \sqrt{c^T A c}$$

Seja  $E'(A,a,c,\gamma) = E(A,a) \cap \{x \in \mathfrak{R}^n : c^T x \leq \gamma\}$

O elipsóide Löwner-John  $E(A',a')$  de  $E'(A,a,c,\gamma)$  é determinado como segue:

se  $-1 \leq \alpha \leq -1/n$  então  $E(A',a') = E(A,a)$

se  $-1/n \leq \alpha < 1$  então  $E(A',a')$  é dado por

$$a' = a - \frac{1+n\alpha}{n+1} b$$

$$A' = \frac{n^2}{n^2-1} (1-\alpha^2) \left( A - \frac{2(1+n\alpha)}{(n+1)(1+\alpha)} b b^T \right)$$

onde

$$b = \frac{1}{\sqrt{c^T A c}} A c$$

observar que: se  $\gamma = c^T a$  então  $E'(A, a, c, \gamma) = E'(A, a, c)$

Se  $0 < \alpha < 1$ , então  $E'(A, a, c, \gamma)$  está estritamente contido em  $E'(A, a, c)$ . Isto significa que cortamos fora uma peça grande de  $E(A, a)$ , e portanto  $c^T x \leq \gamma$  e chamado de corte profundo. Se  $-1/n < \alpha < 0$ , então partimos mais de  $E(A, a)$  que  $E'(A, a, c)$ , é chamamos  $c^T x \leq \gamma$  de corte raso, entretanto neste caso o elipsóide Löwner-John de  $E'(A, a, c, \gamma)$  é estritamente de volume mais pequeno que  $E(A, a)$ . Esta estratégia de redução de volume, permitirá mostrar a polinomialidade do método de elipsóides, que passamos a descrever:

### ALGORÍTMO (método básico de elipsóides)

ENTRADA: um sistema  $Cx \leq d$  de  $m \times n$  desigualdades com coeficientes inteiros. É assumido que  $P = \{x \in \mathbb{R}^n : Cx \leq d\}$  é limitada, não vazia e de dimensão plena.

### INICIALIZAÇÃO

Faça:

$$K := 0$$

$$N := 2n((n+1) \langle C \rangle + n \langle d \rangle - n^3)$$

$$A_0 = R^2 I, \text{ com } R := \sqrt{n} 2^{\langle C, d \rangle - n^2}$$

$$a_0 = 0$$

$$(\text{onde } \langle n \rangle : 1 + \lceil \log_2 (|n| + 1) \rceil, n \in \mathbb{Z})$$

### PASSO GERAL

- (1) se  $k := N$  parar (declarar  $P$  é vazio)
- (2) se  $a_k \in P$ , parar (uma solução viável é encontrado)
- (3) se  $a_k \notin P$ , então escolher uma desigualdade, digamos  $c^T x \leq \gamma$  do sistema  $Cx \leq d$  que é violada por  $a_k$

fazer:

$$b = \frac{1}{\sqrt{c^T A_k c}} A_k c$$

$$a_{k+1} = a_k - \frac{1}{n+1} b$$

$$A_{k+1} = \frac{n^2}{n^2 - 1} (A_k - \frac{2}{n+1} b b^T)$$

ir para (1). ■

Chamamos este algoritmo de algoritmo básico do método de elipsóides, pois contém todas as idéias fundamentais do procedimento.

Se

$$E_{K+1} = E(A_{K+1}, a_{k+1})$$

$$E_K = E(A_K, a_k)$$

é mostrado que

$$\frac{\text{VOL}(E_{k+1})}{\text{VOL}(E_k)} = \left( \left( \frac{n}{n+1} \right)^{n+1} \left( \frac{n}{n-1} \right)^{n-1} \right)^{1/2} < e^{-\frac{1}{2n}} < 1$$

para todo  $k$ , fato que garante a polinomialidade do algoritmo.



## CAPÍTULO 5

# UM ALGORITMO DE PLANOS DE CORTE CENTRO ANALÍTICO COM ELIMINAÇÃO DE CORTES REDUNDANTES

### 5.1 INTRODUÇÃO

Neste capítulo apresentamos nosso algoritmo (AER) de planos de corte centro analítico com eliminação de cortes para o problema de viabilidade convexa. Isto é para encontrar um ponto num conjunto convexo  $S \subset \mathbb{R}^n$ ,  $S$  limitado, com interior não vazio;

É suposto que  $S$  está contido numa caixa  $B$  centrada na origem de coordenadas e que se  $S$  é não vazia  $S$  contém uma bola de raio  $\varepsilon = 2^{-L}$ , onde  $L$  é uma constante fixa tal que  $L \geq \log_2 n$ . O algoritmo (AER) que apresentaremos na seção 5.3, é um esquema que cai dentro do algoritmo padrão de planos de corte pelo centro analítico (APC) para o problema de viabilidade convexa que foi comentado no Capítulo 3.

Em cada iteração do algoritmo (AER), é eliminado alguns hiperplanos (cortes redundantes) que não são facetas do polítopo  $P^k$  (um polítopo gerado no algoritmo). Estes hiperplanos são reconhecidos como tais se não interceptam ou são tangentes a um elipsóide  $Q^k$  que contém ao polítopo  $P^k$  (ver 2.5). A construção dos elipsóides  $Q^k$  é feita como no método das elipsóides, portanto, é garantido o decrescimento do volume dos polítopos num determinado raio, o que determina convergência polinomial do algoritmo.

Eliminação de cortes redundantes, influencia favoravelmente no tempo de computação do algoritmo, tornando-se mais eficiente, pois se trabalha somente com subconjuntos do número total de restrições, e portanto, se reduz o esforço computacional nas iterações internas. Porém não influencia muito no número total de iterações externas do algoritmo que permanecem inalteráveis, como é testado em [2] que implementa um algoritmo de plano de corte-centro analítico para programação convexa diferenciável com estratégias de eliminação de cortes redundantes usando elipsóides contendo ao conjunto “localizador” de soluções ótimas.

### 5.2 UM CRITÉRIO DE IDENTIFICAÇÃO DE RESTRIÇÕES REDUNDANTES NUM POLÍTOPO:

Seja,  $P$  um polítopo limitado. Suponhamos que  $P$  esteja contido no elipsóide  $E(B,b) = \{x \in \mathbb{R}^n : (x-b)^T B^{-1}(x-b) \leq 1\}$  onde  $B$  é uma matriz definida positiva

Seja  $H = \{x \in \mathbb{R}^n : \alpha^T x = \gamma\}$  um hiperplano tal que  $H \cap E(B,b) \neq \emptyset$ . Como

$$\max \{ a^T x / x \in E(B,b) \} = a^T b + \sqrt{a^T B a} = a^T (b + \frac{1}{\sqrt{a^T B a}} B a) \text{ e}$$

$$\min \{ a^T x / x \in E(B,b) \} = a^T b - \sqrt{a^T B a} = a^T (b - \frac{1}{\sqrt{a^T B a}} B a)$$

então o hiperplano H intersecta o elipsóide E(B,b) se, e somente se

$$a^T b - \sqrt{a^T B a} \leq \gamma \leq a^T b + \sqrt{a^T B a} \quad (5.1)$$

isto é, se e somente se

$$\left| \frac{a^T b - \gamma}{\sqrt{a^T B a}} \right| \leq 1$$

Seja  $\alpha = \frac{a^T b - \gamma}{\sqrt{a^T B a}}$ . Então H intersecta o elipsóide E(B,b) se, e somente se,

$$-1 \leq \alpha \leq 1.$$

Se  $\alpha=1$  ou  $\alpha=-1$ , H é tangente ao elipsóide E(B,b) ( ver Fig. 5.1)

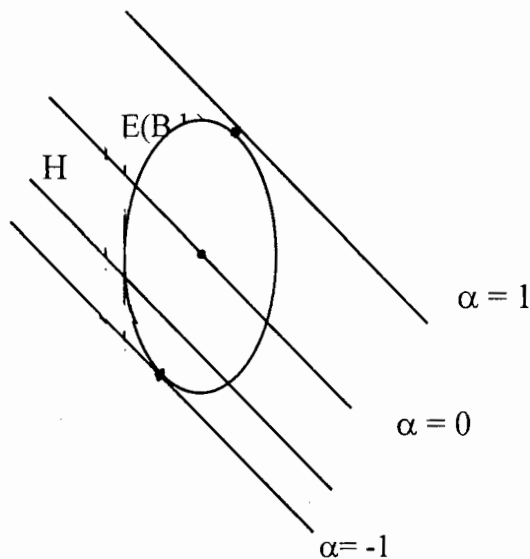


Fig. 5.1

Consideremos o semi-espaço  $H^+ = \{x \in \mathbb{R}^n : a^T x \leq \gamma\}$ . Suponhamos que  $H^+ \cap P \neq \emptyset$ .

Definimos um novo polítopo  $P' = P \cap H^+$ , então é evidente que a seção elipsoidal  $\Omega = E(B,b) \cap H^+$  contém ao polítopo  $P'$ , pois  $E(B,b) \supset P$

Queremos construir um elipsóide

$$E(B',b') = \{x \in \mathbb{R}^n : (x-b')^T (B')^{-1} (x-b') \leq 1\}$$

de menor volume que o elipsóide  $E(B,b)$  tal que  $E'(B',b')$  contenha a seção elipsoidal  $\Omega$ .

Como no método dos elipsóides (apresentado no capítulo 4) construiremos o elipsóide  $E'(B',b')$  de menor volume contendo a seção elipsoidal convexa  $\Omega$ , este elipsóide é único, e é chamado elipsóide Löwner-John de  $\Omega$ , existindo formulas explícitas para construí-lo (ver[23], o que passamos a seguir

Se  $-1 \leq \alpha \leq -1/n$ , então  $E'(B',b') = E(B,b)$

Se  $-1/n < \alpha < 1$ , então  $E'(B',b')$  é definido por

$$b' = b - \left(\frac{1+\alpha}{n+1}\right)d$$

$$B' = \frac{n^2(1-\alpha^2)}{n^2-1} \left(B - \frac{2(1+\alpha)}{(n+1)(1+\alpha)} dd^T\right)$$

onde

$$\alpha = \frac{a^T b - \gamma}{\sqrt{a^T B a}}, \quad d = \frac{B a}{\sqrt{a^T B a}} \quad (\text{ver Fig. 5.2})$$

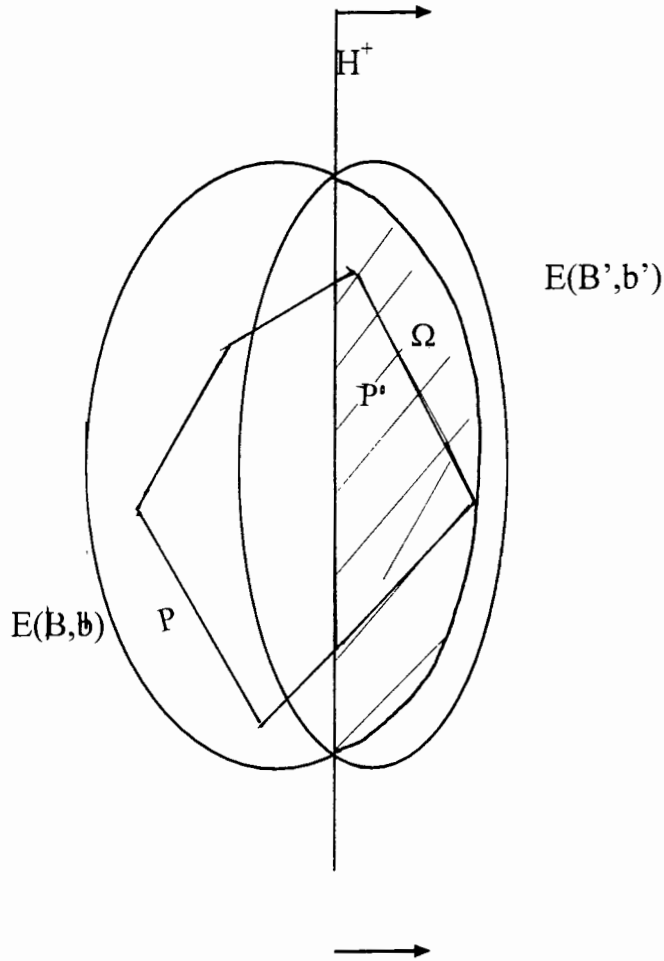


Fig. 5.2

Agora nosso desejo é eliminar o maior número possível de restrições redundantes na definição de  $P^0$ . Para este propósito fazemos uso do elipsóide

$E(B^0, b^0)$  que contém ao polítopo  $P^0$  com os valores  $\alpha$  definidos acima.

Para cada restrição  $H_i^+ : a_i^T x \leq \gamma_i$  que forma parte da definição de  $P^0$ , a distância signada desde o centro  $b^0$  do elipsóide  $E(B^0, b^0)$  ao hiperplano  $H_i : a_i^T x = \gamma_i$  calculada na métrica dada pelo elipsóide  $E(B^0, b^0)$  é dada por

$$\alpha_i = \frac{a_i^T b^0 - \gamma_i}{\sqrt{a_i^T B a_i}}$$

A folga escalada  $\alpha_i$ , indica a posição do hiperplano  $H_i$ , relativa ao elipsóide  $E(B^0, b^0)$ . Então das idéias contidas em (5.1) e de [3] segue que

Se  $\alpha_i < -1$ , então o elipsóide  $E^0$  é interior ao semiespaço  $a_i^T x \leq \gamma_i$ ;

Se  $\alpha_i = -1$ , então o hiperplano  $a_i^T x = \gamma_i$  é tangente ao elipsóide  $E^0$ ;

Se  $\alpha_i > -1$ , então o semiespaço  $a_i^T x \leq \gamma_i$  e o elipsóide  $E^0$  têm interseção não vazia.

Logo, podemos ver claramente, que cada restrição  $a_i^T x \leq \gamma_i$  para o qual  $\alpha_i \leq -1$  é redundante na definição do polítopo  $P^0$ , portanto pode ser eliminada da definição de  $P^0$ . No caso que a restrição  $a_i^T x \leq \gamma_i$  fosse redundante na definição de  $P^0$ , então por simplicidade será dito que o hiperplano  $a_i^T x = b_i$  é redundante.

O teste  $\alpha_i \leq -1$  é com freqüência muito fraco para ser eficaz, assim foram tentados várias estratégias heurísticas envolvendo também a eliminação de algumas restrições para as quais  $\alpha_i > -1$ , em relação a outro tipo de elipsóides  $E(A, c, r)$  contendo  $P^0$  onde  $c$  é um centro analítico aproximado de  $P^0$  e  $r$  um raio construído adequadamente, podendo ser  $r \neq 1$ .

### 5.3 O ALGORITMO (AER)

Seja  $S$  um conjunto convexo limitado em  $\mathcal{R}^n$ . Suponhamos que  $S$  esteja contido na caixa  $D = \{x \in \mathcal{R}^n : -2^L \leq x_i \leq 2^L, i=1, \dots, n\}$  e que se  $S$  é não vazia  $S$  contenha uma bola  $n$ -dimensional de raio  $\varepsilon = 2^{-L}$  onde  $L$  é uma constante fixa  $L \geq \log_2 n$ . A caixa  $D$  está contida na bola com centro na origem de coordenadas e raio  $R = \sqrt{n} 2^L$ . O algoritmo (AER) resolve o problema de (PVC) num tempo polinomial usando um critério de eliminação de cortes redundantes.

#### ALGORITMO (A.E.R.)

## ENTRADA

$P := D$  polítopo inicial (Caixa D),  $P \supset S$

$E := E(B, b)$  elipsóide inicial (bola com centro na origem de coordenadas e raio R)

$$R = \sqrt{n} \cdot 2^L$$

$b := 0$  (centro do elipsóide)

$B := R^2 I_n$  onde  $I_n$  é a matriz identidade de ordem  $n \times n$

$z := 0$  centro analítico aproximado de P

$N_0 = n^3 + 2n^2(1+2L) + 2n^2 \log_2 n$  (número máximo de iterações)

## ITERAÇÃO GENÉRICA (externa)

### passos:

1. Chamar ao oráculo de S em z e em b:

Se  $z \in S$  ou  $b \in S$ , parar;

Senão, se  $z \notin S$  e  $b \notin S$ :

O oráculo de S retorna um vetor  $a, c \in \mathbb{R}^n, v \in \mathbb{R}^n$  tal que

$$\{x \in \mathbb{R}^n : c^T x \leq c^T z\} \supset S \quad \text{e} \quad \{x \in \mathbb{R}^n : v^T x \leq v^T b\} \supset S$$

2. Atualizar o polítopo P:

Se  $b \notin \text{int}(P)$  fazer

$$P' = P \cap \{x \in \mathbb{R}^n : c^T x \leq c^T z\}$$

Se  $b \in \text{int}(P)$  fazer

$$P' = P \cap \{x \in \mathbb{R}^n : v^T x \leq v^T b\}$$

3. Chamar ao algoritmo (R.R) (para identificar algumas restrições redundantes do polítopo  $P'$ )

Dar como entrada ao algoritmo (R.R):

O polítopo P

O elipsóide  $E = E(B, b)$  contendo P

O hiperplano  $\{x \in \mathbb{R}^n : p^T x = \gamma\}$  sendo

$$p = v, \gamma = v^T b \text{ se } b \in \text{int}(P)$$

$$p = a, \gamma = a^T b \text{ se } b \notin \text{int}(P) \text{ onde}$$

$$\{x \in \mathbb{R}^n : a^T x \leq a^T b\} \supset P \text{ (dado por oráculo de P)}$$

$$Q = P \cap \{x \in \mathbb{R}^n : p^T x \leq \gamma\} \text{ (} Q \supset P')$$

O algoritmo (R.R) retorna o elipsóide  $E(B', b')$  contendo  $P'$ .

4. Eliminar as restrições redundantes de  $P'$  identificadas pelo algoritmo (RR)

Seja  $P'_d$  o polítopo  $P'$  sem as restrições redundantes identificadas pelo algoritmo (R.R)

5. Calcular um centro analítico aproximado  $z_d$  do polítopo  $P'_d$

Fazer

$$z := Z_d$$

$$P := P'_d$$

$$E(B, b) := E(B', b')$$

6. Testar o número de iterações externas:

Se o número de iterações é maior que  $N_0$ , então parar. Declarar que S é vazia.

Caso contrário, ir ao passo 1.

No passo 1 do algoritmo (AER), se  $z \notin S$  e  $b \notin S$ , no caso que  $b \notin \text{int}(P)$  a restrição  $a^T x \leq a^T b$  não forma parte da definição do novo polítopo  $P'$ , o corte é usado somente para gerar a seção elipsoidal.  $E(B, b) \cap \{x \in \mathbb{R}^n : a^T x \leq a^T b\} \neq \emptyset$  é construído o novo elipsóide  $E(B', b')$  que contém  $P'$  com a finalidade de eliminar cortes redundantes de  $P'$ .

A seguir apresentamos o algoritmo (RR) que identifica algumas restrições redundantes de  $P'$  usado no passo 3 do algoritmo principal (AER).

### ALGORITMO (R.R)

0. entrada

P polítopo de  $\mathbb{R}^n$

$E = E(B, b)$  elipsóide contendo P

Hiperplano  $\{x \in \mathbb{R}^n : p^T x = \gamma\}$  tal que

$$E(B, b) \cap \{x \in \mathbb{R}^n : p^T x \leq \gamma\} \neq \emptyset$$

$$Q = P \cap \{x \in \mathbb{R}^n : p^T x \leq \gamma\}$$

1. atualizar o elipsóide  $E(B, b)$

$$\text{Calcular } \alpha = \frac{p^T b - \gamma}{\sqrt{p^T B p}}$$

Caso 1. Se  $-1 \leq \alpha \leq -1/n$

Fazer:

$$B' := B, \quad b' := b$$

Definir

$$E' = E(B', b') = E(B, b)$$

Caso 2. Se  $-1/n < \alpha < 1$

Calcular

$$\delta = \frac{n^2(1 - \alpha^2)}{n^2 - 1}$$

$$\sigma = \frac{2(1 + n\alpha)}{(1 + \alpha)(1 + n)}$$

$$\tau = \frac{1 + n\alpha}{n + 1}$$

$$d = \frac{1}{\sqrt{p^T B p}} B p$$

Faça:

$$B' = \delta (B - \sigma d d^T)$$

$$b' = b - \tau d$$

$$E' = E(B', b') \quad (\text{elipsóide contendo } Q)$$

2. Identificação de algumas restrições redundantes do polítopo Q  
para cada hiperplano  $a_i^T x = b_i$  da definição do polítopo Q calcular

$$\lambda_i = \frac{a_i^T b - b_i}{\sqrt{a_i^T B a_i}}$$

se  $\lambda_i \leq -1$  então a restrição  $a_i^T x \leq b_i$  é redundante na definição de Q (é portanto redundante na definição de P')

3. parar.

No algoritmo (R.R) quando é dado como entrada a caixa  $P = D$ , considera-se como o elipsoide  $E(B, b)$  inicial a bola com centro na origem de coordenadas e

raio  $R = \sqrt{\sum_{i=1}^n 2^{2L}} = \sqrt{n} 2^L$ , pois esta bola contém a caixa D, portanto, neste caso  $b=0$  e  $B=R^2 I_n$  onde  $I_n$  é a matriz identidade de ordem n, (Nev Fig. 5.3).

## 5.4 DECREMENTO DO VOLUME DOS ELIPSÓIDES USADOS NO ALGORITMO (R.R)

No algoritmo (R.R.) exposto na seção anterior, consideramos o elipsoide  $E(B, b) = \{x \in \mathbb{R}^n : (x-b)^T B^{-1} (x-b) \leq 1\}$  com B uma matriz definida positiva, contendo ao polítopo P. Usando o elipsoide  $E(B, b)$  foi construído outro elipsoide.

$$E(B', b') = \{x \in \mathbb{R}^n : (x - b')^T (B')^{-1} (x - b') \leq 1\}$$

Contendo a seção elipsoidal

$$\Omega = E(B, b) \cap H^+ \neq \emptyset$$

onde  $H^+ = \{x \in \mathbb{R}^n ; p^T x \leq \gamma\}$ , é evidente que o elipsoide  $E(B', b')$  contém  $Q = P \cap \{x \in \mathbb{R}^n ; p^T x \leq \gamma\}$  por tanto contém ao polítopo P' gerado no algoritmo (AER) (ver passo 2)

Mostraremos que se  $E(B', b') \neq E(B, b)$  o volume do elipsoide  $E' = E'(B', b')$  é estritamente menor que o volume de  $E = E(B, b)$ , por estimativa

do quociente  $\frac{\text{Vol}(E')}{\text{Vol}(E)}$  (ver [23]).

Se  $\alpha = \frac{p^T b - \gamma}{\sqrt{p^T B p}}$ , como  $E(B, b) \cap H^+ \neq \emptyset$ , sabemos que  $-1 < \alpha < 1$ . No algoritmo

(RR)  $E(B', b')$  é diferente de  $E(B, b)$  se  $-1/n < \alpha < 1$ . O caso  $\alpha = -1$  não é analisado, pois neste caso o hiperplano  $H = \{x \in \mathbb{R}^n ; p^T x \leq \gamma\}$  é tangente ao elipsoide  $E(B, b)$  e portanto a seção elipsoidal  $\Omega$  é constituído somente por um ponto.

Analizaremos somente o caso,  $0 \leq \alpha < 1$  já que segundo nosso algoritmo (AER) teremos sempre  $\alpha = 0$  e valores  $0 < \alpha < 1$  pode ser usado numa modificação do algoritmo (AER).



O seguinte lema será usado na demonstração do teorema 5.1 que estabelece a complexidade polinomial do algoritmo (AER)

LEMA 5.1 Se  $0 \leq \alpha < 1$ , no algoritmo (R.R) então

$$\frac{\text{Vol}(E^1)}{\text{Vol}(E)} \leq \left( \left( \frac{n}{n+1} \right)^{n+1} \left( \frac{n}{n-1} \right)^{n-1} \right)^{1/2} < e^{-\frac{1}{2n}} < 1$$

Prova : A demonstração para o caso  $\alpha = 0$  está contida em [23]

Para estimar o quociente  $\frac{\text{Vol}(E^1)}{\text{Vol}(E)}$ , primeiro assumimos que o elipsóide inicial é

$E_0 = S(0, 1) = E(I, 0)$  é a bola unitária ao redor do origem de coordenadas, e  $d = (-1, 0 \dots 0)^T$ .

Neste caso temos as fórmulas (5.2) usadas para construir  $E^1 = E(B^1, b^1)$  tornam-se

$$\begin{aligned} d &= (-1, 0, \dots, 0)^T \\ b^1 &= 0 - \frac{1+n\alpha}{(n+1)} d = \left( \frac{1+n\alpha}{(n+1)}, 0, \dots, 0 \right)^T \\ B^1 &= \frac{n^2(1-\alpha^2)}{n^2-1} \left( I - \frac{2(1+n\alpha)}{(n-1)(1+\alpha)} d^T d \right), \end{aligned}$$

Onde  $I$  é a matriz identidade de ordem  $n \times n$ .

Como  $D = d^T d$  é uma matriz diagonal de ordem  $n \times n$  com o primeiro elemento da diagonal principal igual a 1, e as outras iguais a zero, temos que

$$\begin{aligned} B^1 &= \left( \frac{n^2(1-\alpha^2)}{n^2-1} I - \frac{2n^2(1-\alpha^2)(1+n\alpha)}{(n^2-1)(n+1)(1+\alpha)} D \right) \\ &= \text{diag} \left( r, \frac{n^2(1-\alpha^2)}{n^2-1}, \dots, \frac{n^2(1-\alpha^2)}{n^2-1} \right) \end{aligned}$$

$$\begin{aligned} \text{onde } r &= \frac{n^2(1-\alpha^2)}{n^2-1} - \frac{2n^2(1-\alpha^2)(1+n\alpha)}{(n^2-1)(n+1)(1+\alpha)} \\ &= \frac{1}{(n^2-1)(n+1)(1+\alpha)} [n^2(1-\alpha^2)((n+1)(1+\alpha) - 2(1+n\alpha))] \\ &= \frac{n^2(1-\alpha^2)(1-\alpha)(n-1)}{(n^2-1)(n+1)(1+\alpha)} = \frac{n^2(1-\alpha)^2}{(n+1)^2(1+\alpha)} \end{aligned}$$

Assim finalmente  $B^1$  é a matriz diagonal

$$B^1 = \text{diag} \left( \frac{n^2(1-\alpha)^2}{(n+1)^2}, \frac{n^2(1-\alpha^2)}{n^2-1}, \dots, \frac{n^2(1-\alpha^2)}{n^2-1} \right) \quad (5.3)$$

Seja :  $E_0^a := E(B^a, b^a)$  com  $B^a$  e  $b^a$  dados acima.

É conhecido que para qualquer elipsóide

$$E = E(A, a) = \{x \in \mathfrak{R}^n : (x-a)^T A^{-1} (x-a) \leq 1\}$$

Seu volume que denotamos por  $\text{Vol}(E)$ , depende somente do determinante de  $A$  e da dimensão do espaço e é dado por

$$\text{Vol}(E) = \sqrt{\det A} \cdot V_n \quad (5.4)$$

onde  $V_n$  é o volume da bola unitária  $S(0, 1)$  em  $\mathfrak{R}^n$ , sendo

$$V_n = \frac{\Pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + 1)} \sim \frac{1}{\sqrt{\Pi n}} \left(\frac{2e\Pi}{n}\right)^{n/2}$$

com  $\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt$ , para  $x > 0$ , (função Gamma).

portanto usando a fórmula dada em (5. 4) temos que

$$\text{Vol}(E_0) = \text{Vol}(S(0, 1)) = V_n$$

$$\text{Vol}(E_0^a) = \sqrt{\det B^a} \cdot V_n$$

então

$$\frac{\text{Vol}(E_0^a)}{\text{Vol}(E_0)} = \sqrt{\det B^a} \text{ onde } B^a \text{ é dada por (5.3)}$$

como  $B^a$  é uma matriz diagonal, então o determinante dela é produto dos elementos de sua diagonal principal, assim

$$\det B^a = \frac{n^2(1-\alpha)^2}{(n+1)^2} \left(\frac{n^2(1-\alpha^2)}{n^2-1}\right)^{n-1}$$

logo

$$\frac{\text{Vol}(E_0^a)}{\text{Vol}(E_0)} = \left[ \frac{n^2(1-\alpha)^2}{(n+1)^2} \left(\frac{n^2(1-\alpha^2)}{n^2-1}\right)^{n-1} \right]^{1/2}$$

sendo  $0 \leq \alpha < 1$ , temos que  $0 < 1-\alpha \leq 1$ , e que  $0 < 1-\alpha^2 \leq 1$

o que dá  $(1-\alpha)^2 \leq 1$  e  $(1-\alpha^2)^{n-1} \leq 1$

então

$$\begin{aligned} \frac{\text{Vol}(E^l_0)}{\text{Vol}(E_0)} &\leq \left( \frac{n^{2n}}{(n-1)^2 (n-1)^{n-1} (n+1)^{n-1}} \right)^{1/2} \\ &= \left( \frac{n^{n+1} n^{n-1}}{(n+1)^{n+1} (n-1)^{n-1}} \right)^{1/2} \\ &= \left( \left( \frac{n}{n+1} \right)^{n+1} \left( \frac{n}{n-1} \right)^{n-1} \right)^{1/2} \end{aligned}$$

isto é

$$\frac{\text{Vol}(E^l_0)}{\text{Vol}(E_0)} \leq \left( \left( \frac{n}{n+1} \right)^{n+1} \left( \frac{n}{n-1} \right)^{n-1} \right)^{1/2} \quad (5.5)$$

portanto

$$\frac{\text{Vol}(E^l_0)}{\text{Vol}(E_0)} < e^{-\frac{1}{2n}} \quad (5.6)$$

se  $\left( \frac{n}{n+1} \right)^{n+1} \left( \frac{n}{n-1} \right)^{n-1} < e^{-\frac{1}{n}}$  (5.7)

desde que

$$\begin{aligned} \left( \frac{n}{n+1} \right)^{n+1} \left( \frac{n}{n-1} \right)^{n-1} < e^{-\frac{1}{n}} &\Leftrightarrow e^{\frac{1}{n}} < \left( \frac{n+1}{n} \right)^{n+1} \left( \frac{n-1}{n} \right)^{n-1} \\ &\Leftrightarrow \frac{1}{n} < (n+1)\ln\left(1+\frac{1}{n}\right) + (n-1)\ln\left(1-\frac{1}{n}\right) \end{aligned}$$

Logo basta mostrar que

$$\frac{1}{n} < (n+1)\ln\left(1+\frac{1}{n}\right) + (n-1)\ln\left(1-\frac{1}{n}\right) \quad (5.8)$$

para ter (5.6)

Para mostrar que (5.8) é verdadeira, faremos uso da série de potências

$$\ln x = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(x-1)^k}{k} \text{ para } 0 < x \leq 2$$

tomando  $x = 1 + \frac{1}{n}$

$$\ln\left(1 + \frac{1}{n}\right) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{kn^k}$$

e tomando  $x = 1 - \frac{1}{n}$ ,  $0 < 1 - 1/n < 2$ ,  $n \neq 1$

$$\ln\left(1 - \frac{1}{n}\right) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1} (-1)^k}{kn^k} = - \sum_{k=1}^{\infty} \frac{1}{kn^k}$$

logo

$$\begin{aligned} (n+1) \ln\left(1 + \frac{1}{n}\right) + (n-1) \ln\left(1 - \frac{1}{n}\right) &= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(n+1)}{kn^k} - \sum_{k=1}^{\infty} \frac{(n-1)}{kn^k} \\ &= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{2}{kn^k} - \sum_{k=1}^{\infty} \frac{2(n-1)}{2kn^{2k}} \\ &= \sum_{k=1}^{\infty} \frac{2}{(2k-1)n^{2k-1}} - \sum_{k=1}^{\infty} \frac{2}{2kn^{2k}} - \sum_{k=1}^{\infty} \frac{1}{kn^{2k-1}} + \sum_{k=1}^{\infty} \frac{1}{kn^{2k}} \\ &= \sum_{k=1}^{\infty} \frac{1}{(2k-1)k n^{2k-1}} > \frac{1}{n} \end{aligned}$$

assim é provado (5.8), e portanto (5.6) é válido.

Isto é o lema 5.1, é válido para o caso especial dos elipsóides  $E_0$  e  $E_0^l$  definidos acima. Usando este resultado mostraremos a validade do lema 5.1. Para  $E = E(B, b)$  e  $E' = E(B', b')$  construído usando as fórmulas dadas em (5.2) do algoritmo (R. R.).

É conhecido que para cada matriz definida positiva  $A$ , existe uma única matriz definida positiva, denotada por  $A^{1/2}$  tal que  $A = A^{1/2} A^{1/2}$ .

Também é conhecido que cada elipsoide  $E(A, a)$  pode-se expressar como  $E(A, a) = A^{1/2} S(0, 1) + a$ . Assim cada elipsóide é a imagem da bola unitária  $S(0, 1)$  mediante uma transformação  $T$  afim bijetora.

Então, sendo a matriz  $B$  do elipsóide  $E = E(B, b)$  é definida positiva, temos que

$$E = E(B, b) = B^{1/2} S(0, 1) + b = B^{1/2} E_0 + b$$

Como na definição de  $E' = E(B', b')$ , (ver (5.2)) temos que

$$d = \frac{1}{\sqrt{p^T B p}} B p = B^{1/2} \left( \frac{1}{\sqrt{p^T B p}} B^{1/2} p \right)$$

Existe uma matriz ortogonal, digamos  $Q$  ( $Q^T Q = I$ ) a qual rota o vetor  $B^{1/2} p$  num múltiplo positivo do vetor  $(-1, 0, \dots, 0)$ , isto é verifica-se que

$$(-1, 0, \dots, 0)^T = \left( \frac{1}{\left\| B^{1/2} p \right\|} Q B^{1/2} p \right) = \left( \frac{1}{\sqrt{p^T B p}} Q B^{1/2} p \right)$$

Então

$$T(x) = B^{1/2} Q^T x + b \quad (5.9)$$

é a transformação afim bijetora tal que  $E(B, b)$  é a imagem da bola unitária  $S(0, 1)$  mediante  $T$ .

$$\begin{aligned} \text{Observar que } T((-1, 0, \dots, 0)^T) &= B^{1/2} Q^T \left( \frac{1}{\sqrt{p^T B p}} Q B^{1/2} p \right) \\ &= \left( \frac{1}{\sqrt{p^T B p}} B^{1/2} (Q^T Q) B^{1/2} p \right) \\ &= \left( \frac{1}{\sqrt{p^T B p}} B p \right) = d \end{aligned}$$

e que  $T^{-1}(x) = Q B^{-1/2} (x - b)$

$$\begin{aligned} \text{Agora: } T(E_0) &= \{ T(y) : y^T y \leq 1 \} \\ &= \{ x : (T^{-1} x)^T (T^{-1} x) \leq 1 \} \\ &= \{ x : (x - b)^T B^{-1/2} Q^T Q B^{-1/2} (x - b) \leq 1 \} \\ &= \{ x : (x - b)^T B (x - b) \leq 1 \} = E(B, b) = E \end{aligned}$$

Analogamente pode-se mostrar que  $T(E'_0) = E(B', b') = E_0$ . Por outro lado, é conhecido que se  $T: x \rightarrow D x + d$  é uma transformação afim bijetora, então

$$\text{Vol}(T(E(A, a))) = (\det. D) \sqrt{\det A} \cdot V_n = (\det. D) \text{Vol}(E(A, a))$$

para qualquer elipsoide  $E(A, a)$ .

Portanto usando esta propriedade da transformação (5.9) e o fato que  $T(E_0) = E$  e  $T(E'_0) = E'$ , temos que

$$\frac{\text{Vol}(E^l)}{\text{Vol}(E)} = \frac{\text{Vol}(T(E_0^l))}{\text{Vol}(T(E_0))} = \frac{\det(B^{1/2}Q^T)\text{Vol}(E_0^l)}{\det(B^{1/2}Q^T)\text{Vol}(E_0)} = \frac{\text{Vol}(E_0^l)}{\text{Vol}(E_0)}$$

e usando o resultado (5.5) e (5.7) concluímos que

$$\frac{\text{Vol}(E^l)}{\text{Vol}(E)} \leq \left(\frac{n}{n+1}\right)^{n+1} \left(\frac{n}{n-1}\right)^{n-1} < e^{-\frac{1}{n}}.$$

O algoritmo (RR) é o lema 5.1 foi estabelecido usando os elipsóides  $E(B,b) = \{x \in \mathfrak{R}^n : (x-b)^T B^{-1} (x-b) \leq 1\}$  com raio igual a 1. Também pode-se mostrar a validade do algoritmo e do lema para elipsóides

$E(B,b) = \{x \in \mathfrak{R}^n : (x-b)^T B^{-1} (x-b) \leq r\}$  com raio  $r$  diferente de 1, fazendo uma mudança de escala que leve a ter raio igual a 1.

## 5.5 POLINOMIALIDADE DO ALGORITMO (AER)

Em cada iteração externa do algoritmo (AER) quando no passo 3, é chamado o algoritmo interno (RR) tendo como entrada um elipsóide  $E(B,b)$  contendo o polítopo  $P$ , o hiperplano  $\{x \in \mathfrak{R}^n : p^T x = \gamma\}$ , o valor de  $\gamma$  e  $p$  dependem da condição de pertença de  $b$  ao interior de  $P$ , do seguinte modo:

Se  $b \notin \text{int}(P)$  então  $\gamma = a^T b$ ,  $p = a$ , logo

$$\alpha = \frac{p^T b - \gamma}{\sqrt{p^T B p}} = 0$$

e se  $b \in \text{int}(P)$  então  $\gamma = v^T b$ ,  $p = v$ , logo

$$\alpha = \frac{p^T b - \gamma}{\sqrt{p^T B p}} = 0$$

portanto  $\gamma = 0$  em cada iteração do algoritmo, (AER), que é fundamental para provar a polinomialidade de nosso algoritmo, que estabelecemos no seguinte teorema.

**TEOREMA 5.1.** Se o algoritmo (AER) não encontra primeiro um ponto em  $S$ , então depois de

$$N_0 = n^3 + 2n^2(1 + 2L) + 2n^2 \log_2 n$$

iterações, o algoritmo determina que o conjunto  $S$  é vazio.

prova:

Denotemos por  $E_k = E(B^k, b^k)$  o elipsóide gerado pelo algoritmo (RR) na  $k$ -ésima iteração do algoritmo (AER) para identificar restrições redundantes do corrente polítopo  $P_k$ .

Seja  $N$  alguma iteração, denotemos por  $E_0 = E(B_0, b_0)$  o elipsóide inicial, que é uma bola com centro na origem de coordenadas e raio

$$R = \sqrt{\sum_{i=1}^n 2^{2L}} = \sqrt{n 2^{2L}} = \sqrt{n} 2^L$$

Temos que

$$\frac{\text{Vol}(E_N)}{\text{Vol}(E_0)} = \frac{\text{Vol}(E_N)}{\text{Vol}(E_{N-1})} \frac{\text{Vol}(E_{N-1})}{\text{Vol}(E_{N-2})} \cdots \frac{\text{Vol}(E_1)}{\text{Vol}(E_0)}$$

como  $\alpha = 0$  em cada iteração, então pelo lema 5.1

$$\frac{\text{Vol}(E_{N-i})}{\text{Vol}(E_{N-i-1})} < e^{-\frac{1}{2^n}} \text{ para } 0 \leq i \leq N-1$$

logo

$$\frac{\text{Vol}(E_N)}{\text{Vol}(E_0)} < e^{-\frac{N}{2^n}}$$

ou seja  $\text{vol}(E_N) < e^{-\frac{N}{2^n}} \text{vol}(E_0)$

sabemos que  $\text{vol}(E_0) \leq \text{vol}(C) = 2^n R^n$

onde  $C = \{x \in \mathbb{R}^n : \|x\|_\infty \leq R\}$ ,  $\|x\|_\infty = \max\{|x_i|\}$

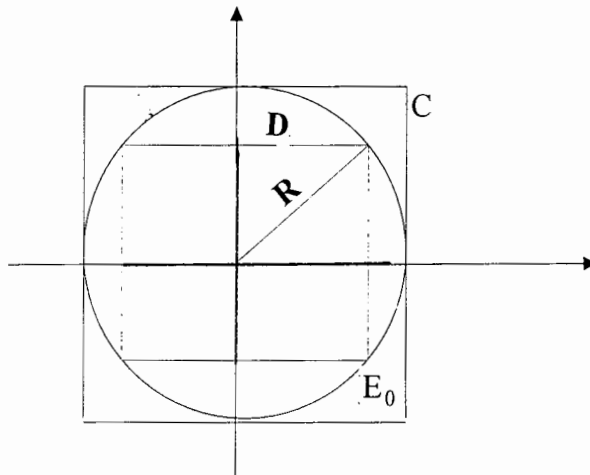


Figura 5.3

por outro lado, como  $n < 2^n$  para todo  $n \geq 1$ , tem-se que

$$R = \sqrt{n} 2^L < 2^{L+\frac{n}{2}}$$

e como  $e^{-\frac{N}{2^n}} < 2^{-\frac{N}{2^n}}$

então  $\text{Vol}(E_N) < e^{-\frac{N}{2^n}} \text{Vol}(E_0) \leq 2^{-\frac{N}{2^n}} (2^n R^n) \leq 2^{-\frac{N}{2^n}} \cdot 2^n \cdot (2^{L+n/2})^n$

ou seja

$$\text{Vol}(E_N) \leq 2^{-\frac{N}{2^n} + n + nL + \frac{n^2}{2}}$$

pelas condições assumidas para o conjunto convexo  $S$ , Se  $S \neq \emptyset$ ,  $S$  contém uma bola  $\tilde{B}$  de raio  $2^{-L}$ . Então, os polítopos  $P_k$  gerados durante o algoritmo devem conter a bola  $\tilde{B}$ , e portanto os elipsóides  $E_k$  gerados contem a bola  $\tilde{B}$ .

Em  $\mathcal{R}^n$  sabemos que o volume de uma bola de raio  $2^{-L}$  é limitada inferiormente por  $(\frac{2^{-L}}{n})^n$ . Assim

$$\text{Vol}(\tilde{B}) \geq (\frac{2^{-L}}{n})^n$$

Mostraremos agora que para  $N > N_0$  tem-se que  $\text{Vol}(E_N) < (\frac{2^{-L}}{n})^n$ . O que implicará que

$$\text{Vol}(E_N) < \text{Vol}(\tilde{B})$$

o que é uma contradição, pois  $E_N \supset \tilde{B}$ . Logo deve-ser  $S$  vazia. Ou seja depois da iteração  $N_0$ , o algoritmo determinará que  $S$  é vazia.

De fato:

$$\text{Se } 2^{\frac{-N}{2n} + n + nL + \frac{n^2}{2}} < (\frac{2^{-L}}{n})^n = 2^{-nL} n^{-n}$$

$$\text{Teremos que } \text{Vol}(E_N) < (\frac{2^{-L}}{n})^n$$

Temos que

$$\begin{aligned} 2^{\frac{-N}{2n} + n + nL + \frac{n^2}{2}} < 2^{-nL} n^{-n} &\Leftrightarrow -N/2n + n(1+2L) + n^2/2 < -n \log_2 n \\ &\Leftrightarrow n \log_2 n + n(1+2L) + n^2/2 < N/2n \\ &\Leftrightarrow N > n^3 + 2n^2(1+2L) + 2n^2 \log_2 n = N_0 \end{aligned}$$

Do teorema 5.1 vemos que algoritmo (AER) apresenta complexidade  $O(n^2L)$ .

Uma modificação ao algoritmo (AER) pode ser considerado no passo 2. Fazendo

cortes pelo centro analítico  $z$  se o respectivo valor de  $\alpha = \frac{a^t b - a^t z}{\sqrt{a^t B a}} \in [0,1)$ , onde

$a^T x = a^T z$  é o hiperplano separador de  $z$  e  $S$ . Caso contrario considera-se o corte  $a^T x = a^T b$  de  $b$  e  $S$ . Observe que neste caso e preservado em cada iteração  $\alpha \in [0,1)$ , e portanto o algoritmo com esta modificação tem a mesma complexidade do algoritmo (AER), à demonstração é a mesma dada para o teorema 5.1.



# CAPÍTULO 6

## UMA APLICAÇÃO AO PROBLEMA DE PROGRAMAÇÃO CONVEXA RESTRITA

### 6.1 INTRODUÇÃO

Sabemos que um problema de programação convexa é um caso particular de um problema de viabilidade convexa. Portanto pode-se aplicar para resolvê-la os algoritmos existentes para o problema geral de viabilidade convexa.

Neste capítulo aplicaremos o algoritmo de Goffin-Luo-Ye [21] para resolver um problema de programação convexa restrita. Uma aplicação deste algoritmo para o caso de programação convexa com restrição caixa foi desenvolvido por Kiwiel [27], especificamente para o problema seguinte:

$$\begin{aligned} \text{Min. } f(x) &= f^* \\ 0 &\leq x \leq e \end{aligned}$$

Onde  $f: \mathfrak{R}^n \rightarrow \mathfrak{R}$  é convexo,  $f^* \geq 0$ .

Nos estendemos o algoritmo de Kiwiel para o caso de programação convexa restrita, e damos uma estimativa de complexidade do algoritmo sob algumas condições.

Antes de apresentar nosso algoritmo damos a conhecer o algoritmo de Goffin-Luo-Ye [21]; e o algoritmo de K.. C. Kiwiel [27].

#### 6.1.1 O ALGORITMO DE GOFFIN-LUO-YE PARA O PROBLEMA GERAL DE VIABILIDADE CONVEXA

Dado um conjunto convexo  $D \subset \mathfrak{R}^n$ , com interior não vazio, tal que  $D$  está contido na caixa

$$\Omega^0 = \{y \in \mathfrak{R}^n : 0 \leq y \leq e\} = [0,1]^n$$

o problema de viabilidade considerado é encontrar um ponto no interior de  $D$ .

O conjunto  $D$  é definido implicitamente, é assumido que existe um oráculo tal que para cada  $\bar{y} \in \Omega^0$  ou responde que  $\bar{y} \in \text{int}D$  ou gera um hiperplano separador de  $\bar{y}$  e  $D$

$$\{y \in \mathfrak{R}^n : \alpha^T y \leq \alpha^T \bar{y}\} \supset D$$

É assumido que  $\|\alpha\| = 1$ . Também é assumido que  $\text{int}D$  contém uma bola fechada  $n$ -dimensional com raio  $\varepsilon < 1/2$ .

Antes de apresentar o algoritmo de Goffan L. Y. damos uma definição de proximidade ao centro analítico.

### Uma Medida de Proximidade ao Centro Analítico.

Dado um polítopo  $\Omega = \{y \in \mathfrak{R}^n : A^T y + s = c, s \geq 0\}$  limitado em  $\mathfrak{R}^n$ , e  $y \in \text{int}\Omega$ , tal que  $s = A^T y - c > 0$ , para uma variável primal  $x > 0$  e  $Ax = 0$ , define-se a medida

$$\eta(x, y) = \|X(c - A^T y) - e\| = \|Xs - e\| = \|Sx - e\|$$

Se  $\|Xs - e\| < \eta < 1$ , então o ponto  $y$  é dito um  $\eta$ -aproximado centro analítico do polítopo  $\Omega$ ;

Outra medida ao centro analítico é definido por

$$\delta(y) = \|Sx(y) - e\| = \min_x \{ \|Sx - e\| : Ax = 0 \}$$

onde

$$\begin{aligned} x(y) &= S^{-1} e - S^{-2} A^T (AS^{-2} A^T)^{-1} AS^{-1} e \\ &= S^{-1} (I - S^{-1} A^T (AS^{-2} A^T)^{-1} AS^{-1}) e \end{aligned}$$

e verifica  $Ax(y) = 0$ .

Se  $x(y) > 0$ , então  $\delta(y) = \eta(x(y), y)$ .

Se  $\delta(y) < 1$ , então é mostrado que

$$\delta(y) = \eta(x(y), y)$$

Também é mostrado que se

$$\delta(y) < 1, \text{ então o ponto } y^+ = y + \Delta(y) \in \text{int}(\Omega)$$

com  $\Delta(y) = -(\nabla F^2(y))^{-1} (\nabla F(y)) = -(AS^{-2} A^T)^{-1} AS^{-1} e$

e  $\delta(y^+) \leq \delta(y)^2$ . Fato que da convergência quadrática do método de Newton para minimizar a função barreira logarítmica.

Usando a  $\eta$ -medida de proximidade, apresentamos a seguir o algoritmo de Goffin-Luo Ye.

### Algoritmo (AGLY)

#### Passo 0. (Inicialização)

Seja  $0 < \eta < 1$

$$A^0 = (I, -I) \in \mathfrak{R}^{n \times 2n}$$

$$c^0 = \begin{pmatrix} e \\ 0 \end{pmatrix} \in \mathfrak{R}^{2n}$$

$$y^0 = \frac{1}{2} e \in \mathfrak{R}^n$$

$$s^0 = c^0 - (A^0)^T y^0 = \frac{1}{2} e \in \mathfrak{R}^{2n}$$

$$x^0 = e \in \mathfrak{R}^{2n}$$

$$k = 1$$

(Observar que  $\Omega^0 = \{ y \in \mathfrak{R}^n : c^0 - (A^0)^T y \geq 0 \}$ ,  $A^0 x^0 = 0$ ,  $x^0 > 0$ ,  $\| X^0 s^0 - e \| = 0$ ,  $y^0$  centro analítico de  $\Omega^0$ )

#### Passo 1 (cálculo do centro analítico aproximado)

Começando no ponto  $(x^{k-1}, y^{k-1}, s^{k-1})$ , com  $y^{k-1}$  um  $\eta$ -aproximado de  $\Omega^{k-1}$ , calcular  $(x^k, y^k, s^k)$ , com  $y^k$  um  $\eta$ -aproximação centro analítico de

$$\Omega^k = \{ y \in \mathfrak{R}^n : c^k - (A^k)^T y \geq 0 \}$$

#### Passo 2 (critério de parada)

Perguntar ao oráculo de D

se  $y^k \in \text{int}D$ .

Se é sim, parar.

Se não, continue.

#### Passo 3 (geração de corte)

Se  $y^k \notin \text{int}D$ , o oráculo devolve um hiperplano

$$\{ y \in \mathfrak{R}^n : a_{k+1}^T y \leq a_{k+1}^T y^k \} \supset D \text{ com } \|a_{k+1}\| = 1$$

fazer  $A^{k+1} = (A^k, a_{k+1})$

$$c^{k+1} = \begin{pmatrix} c^k \\ a_{k+1}^T y^k \end{pmatrix}$$

Considerar  $\Omega^{k+1} = \{y \in \mathbb{R}^n : C^{k+1} - (A^{k+1}) y \geq 0\}$

#### Passo 4

fazer  $k := k+1$  logo retornar ao passo 1.

O seguinte teorema estabelece a estimativa de complexidade do algoritmo (AGLY)

#### TEOREMA 6.1

O algoritmo (AGLY) termina com uma solução viável ( $y^k \in \text{int}D$ ) logo que  $k$  satisfaz

$$\frac{\varepsilon^2}{n} \leq \frac{\frac{1}{2} + 2n \ln(1 + \frac{k+1}{8n^2})}{2n+k+1} \exp(-2\alpha \frac{k+1}{k+1+2n})$$

onde  $\alpha$  é uma constante.

O teorema 6.1 implica que a complexidade do método é de  $O^*(n^2 / \varepsilon^2)$  iterações, onde a notação  $O^*$  significa que os termos de menor ordem são ignorados. É observado que  $\alpha$  positiva ou não positiva, não altera a estimativa de complexidade significativamente.

### **6.1.2 O ALGORITMO DE KIWIEL PARA UM PROBLEMA DE PROGRAMAÇÃO CONVEXA.**

O algoritmo que apresentaremos a seguir, é o algoritmo de Kiwiel [27], que é uma aplicação do algoritmo (AGLY) para resolver o problema convexo seguinte:

$$\text{Min. } f(x) = f^*$$

$$0 \leq x \leq e$$

para  $f: \mathfrak{R}^n \rightarrow \mathfrak{R}$  convexo, com  $f^* \geq 0$ .

É assumido que  $f^* \geq 0$  e que  $f(\frac{1}{2} e) = 1$ ,

define-se

$$X = [0,1]^n,$$

$$f_X(x) = \begin{cases} f(x) & \text{se } x \in X, \\ \infty & \text{se } x \notin X \end{cases}$$

dada uma tolerancia  $\bar{\varepsilon} > 0$ , uma  $\bar{\varepsilon}$ -solução (isto é  $x$  tal que  $f_X(x) \leq f^* + \bar{\varepsilon}$ ) é encontrado no conjunto

$$Y_{\bar{\varepsilon}} = \{ y \in [0,1]^{\tilde{n}} : y = (x, \xi), f(x) \leq \xi, \xi \leq f^* + \bar{\varepsilon} \}$$

onde  $\tilde{n} = n+1$  será usada as normas

$$\|x\| = \left( \sum_{i=1}^n x_i^2 \right)^{1/2}$$

$$\|x\|_1 = \sum_{i=1}^n |x_i|$$

$\|z\|_{\infty} = \max \{ |x_i| : 1 \leq i \leq n \}$ ,  $I$  denotará a matriz identidade e  $e$  o vetor com componentes todos iguais a 1.

### ALGORITMO (AK)

#### Passo 0 (inicialização)

Fazer

$$A^0 = (I, -I) \in \mathfrak{R}^{\tilde{n} \times 2\tilde{n}}$$

$$c^0 = \begin{pmatrix} e \\ 0 \end{pmatrix} \in \mathfrak{R}^{2\tilde{n}}$$

$$k = 0$$

$$j_0 = 0 \text{ (} j_k \text{ é o número de cortes subgradientes realizadas antes da iteração } k \text{)}$$

#### Passo 1 (cálculo do centro analítico)

Achar o centro analítico  $y^k = (x^k, \xi_k)$  do polítopo

$$\Omega^k = \{ y \in \mathfrak{R}^{\tilde{n}} : (A^k)^T y \leq c^k \}$$

dado por  $m_k = 2\tilde{n} + j_k$  desigualdades.

Seja  $s^k = c^k - (a^k)^T y^k > 0$

Passo 2 (critério de parada)

se  $x^k$  é uma solução desejada, então parar  
senão, continuar.

Passo 3 (geração de corte)

Encontrar  $f(x^k)$  e  $g^k \in \partial f(x^k)$ .

Se  $f(x^k) < \xi_k$ , então faça corte horizontal:

Fazer 
$$c_{\tilde{n}}^{k+1} = f(x^k)$$
$$c_j^{k+1} = c_j^k \text{ para } j = 1, \dots, m_k \text{ e } j \neq \tilde{n}$$
$$A^{k+1} = A^k$$
$$j_{k+1} = j_k$$

Se  $f(x^k) \geq \xi_k$ , então gerar um corte subgradiente:

Fazer 
$$a_{m_{k+1}} = \begin{pmatrix} g^k \\ -1 \end{pmatrix}$$
$$c_{m_{k+1}} = (g^k)^T x^k - f(x^k)$$
$$A^{k+1} = A^k$$
$$c^{k+1} = \begin{pmatrix} c^k \\ c_{m_{k+1}} \end{pmatrix}$$
$$j_{k+1} = j_k + 1$$
$$k(j_{k+1}) = k$$

Passo 4

Fazer  $k := k + 1$  e retornar para passo 1.

A estimativa da complexidade do algoritmo (A.K.) é dada no seguinte teorema. (ver [27])

Teorema 6.2

Seja  $L_\infty = \sup \{ \|g\|_1 : g \in \partial f(x), 0 < x < e \}$

e  $f_{\text{rec}}^k = \min_{j < k} f(x^j)$ .

Então,

$$f_{\text{rec}}^k - f^* \leq 2(1 + L_\infty) \sqrt{\frac{\frac{\tilde{n}}{2} + \tilde{n} \ln(1 + \frac{j_k}{8\tilde{n}}) / 8 \ln \frac{9}{8}}{2\tilde{n} + j_k}} \exp(-2 \frac{k - (1 - \tilde{\alpha})j_k}{2\tilde{n} + j_k})$$

corolário  $f_{\text{rec}}^k \downarrow f^*$

## 6.2 O PROBLEMA CONVEXO

Consideremos o seguinte problema convexo.

$$(P) \begin{cases} \text{Min. } f(x) \\ \text{s.a} \\ h_j(x) \leq 0, \text{ para } j=1, \dots, m \\ 0 \leq x \leq e \end{cases}$$

Onde  $h_j: \mathfrak{R}^n \rightarrow \mathfrak{R}$  para  $j=1, \dots, m$  são funções convexas em  $\mathfrak{R}^n$ , e

$f: \mathfrak{R}^n \rightarrow \mathfrak{R}$  é convexa sobre conjunto convexo

$$A = \{x \in \mathfrak{R}^n: 0 \leq x \leq e, h_j(x) \leq 0, j=1, \dots, m\}$$

Seja

$$h(x) = \max_{1 \leq j \leq m} h_j(x)$$

Como as funções  $h_j$  são convexas sobre  $\mathfrak{R}^n$ , então a função  $h$  é convexa sobre  $\mathfrak{R}^n$ .

Sem perda de generalidade, por simplicidade, podemos considerar em lugar do problema convexo (P), o seguinte problema convexo.

$$(PC) \begin{cases} \text{Min. } f(x) \\ \text{s.a} \\ h(x) \leq 0 \\ 0 \leq x \leq e \end{cases}$$

Seja  $f^* = \min\{f(x): h(x) \leq 0, 0 \leq x \leq e\}$ . Suponhamos que  $f^* \geq 0$ ,  $h(\frac{1}{2}e) \leq 0$  e que

$f(\frac{1}{2}e) = 1$ . Também assumimos que é possível calcular  $g = g(x) \in \partial f(x)$ ,

$\mu = \mu(x) \in \partial h(x)$  nos pontos de definição de  $f$  e  $h$ , e que  $\|g\|_1 \leq L_\infty$ ,  $\|\mu\|_1 \leq L_\infty$ .

## 6.3 ALGUMAS DEFINIÇÕES

Dado um Conjunto finito  $I \subset \mathbb{N}$  onde  $\mathbb{N} = \{ 1, 2, \dots \}$  e uma seqüência de pontos  $\{x^i\}_{i \in I}$  em  $\mathfrak{R}^n$ , as funcionais

$$f^i(x) = f(x^i) + g(x^i)(x-x^i) \text{ para } i \in I$$

são suportes afins de  $f$  em  $x^i$ . Então, definimos

$$f_I(x) = \max\{f^i(x) : x \in I\}$$

$f_I$  é uma relaxação de  $f$ , isto é  $f_I \leq f$ . Similarmente dado um conjunto finito  $J \subset \mathbb{N}$  e uma seqüência  $\{x^j\}_{j \in J}$  de pontos em  $\mathfrak{R}^n$ , definimos as funcionais

$$h^j(x) = h(x^j) + \mu_j^T(x-x^j), j \in J.$$

onde  $\mu_j \in \partial h_{\ell}(x^j)$  para algum  $\ell \in J$  tal que  $h(x^i) = h_{\ell}(x^j)$

A função

$$h_J(x) = \max\{h^j(x) : j \in J\}$$

é uma relaxação de  $h$ , isto é  $h_J \leq h$ , então tem-se que  $D_J \supset D$  onde

$$D = \{x \in \mathfrak{R}^n : h(x) \leq 0\}, D_J = \{x \in \mathfrak{R}^n : h_J(x) \leq 0\}$$

Assim, o seguinte problema convexo

$$\begin{cases} \text{Min. } f_I(x) \\ \text{s.a} \\ h_J(x) \leq 0 \\ 0 \leq x \leq e \end{cases}$$

é uma relaxação do problema convexo(PC) que pode ser tratado como um problema de programação linear. A sua solução  $f_I^* = f_I(x_I^*)$  é limitante inferior do valor  $f^* = f(x^*)$  otimal do problema (PC). O valor

$$\xi_{\text{up}} = \min\{f(x^i) : i \in I, h(x^i) \leq 0, 0 \leq x^i \leq e\}$$

é um limitante superior do valor optimal  $f^*$ .

Definimos o conjunto de localização  $\Omega_{\xi_{\text{up}}}$  por

$$\Omega_{\xi_{\text{up}}} = \{y = (x, \xi) \in \mathfrak{R}^{n+1} : f_I(x) \leq \xi, \xi \leq \xi_{\text{up}}, h_J(x) \leq 0, 0 \leq x \leq e\}$$

que é um subconjunto limitado do epígrafo de  $f_I$ .

O polítopo  $\Omega_{\xi_{\text{up}}}$  contém cada ponto ótimal  $(x^*, f^*)$  do problema (PC). Pode-se observar que quando o número de pontos  $\{x_i\}$  com índices em  $I$  ou  $J$  crescem os limitantes superiores  $\xi_{\text{up}}$  de  $f^*$  decrescem, então o conjunto localizador é uma melhor aproximação exterior do conjunto.

$$\Omega_{\epsilon} = \{y = (x, \xi) \in [0,1]^{n+1} : h(x) \leq 0, 0 \leq x \leq e, f(x) \leq \xi, \xi \leq f^* + \epsilon\} \text{ onde}$$



$\varepsilon > 0$ , que chamamos de conjunto  $\varepsilon$ -ótimal do problema (PC). Um ponto  $\bar{x}$  tal que  $(\bar{x}, \bar{y}) \in \Omega_\varepsilon$  é chamado uma  $\varepsilon$ -solução do problema(PC). Embora o valor ótimal  $f^*$  seja não conhecido, apresentamos um algoritmo para encontrar um ponto em  $\Omega_\varepsilon$ , para valores arbitrariamente pequenos de  $\varepsilon$ .

Em nosso algoritmo, assumimos que o conjunto convexo

$$S = \Omega_\varepsilon \text{ esteja contido na caixa}$$

$$\Omega^0 = \{ y = (x, \xi) \in \mathfrak{R}^{\tilde{n}} : 0 \leq y \leq e \},$$

onde  $\tilde{n} = n + 1$ . Baseados no algoritmo do Goffin-Luo-Ye [21] e modificando a análise feita por K.C. Kiwiel [27] damos uma estimativa de complexidade do método.

## 6.4 ALGORITMO (p.v. $\varepsilon$ )

### Passo 0 (inicialização)

$$A^0 = (I, -I) \in \mathfrak{R}^{\tilde{n} \times 2\tilde{n}}, I \text{ matriz identidade}$$

$$C^0 = \begin{pmatrix} e \\ 0 \end{pmatrix} \in \mathfrak{R}^{2\tilde{n}},$$

$$y^0 = \frac{1}{2} e = (x^0, \frac{1}{2}) \in \mathfrak{R}^{\tilde{n}}, x^0 = \frac{1}{2} e \in \mathfrak{R}^n$$

$$s^0 = c^0 - (A^0)^T y^0 = \frac{1}{2} e \in \mathfrak{R}^{2\tilde{n}}$$

$$k = 0$$

$$j_0 = 0 \text{ (} j_k \text{ número de cortes não horizontais efetuadas antes da iteração } k \text{)}$$

$$\Omega^0 = \{ y = (x, \xi) \in \mathfrak{R}^{\tilde{n}} : c^0 \geq (A^0)^T y \}$$

### Passo 1 (cálculo do centro analítico)

Calcular o centro analítico  $y^k = (x^k, \xi_k)$  do polítopo

$$\Omega^k = \{ y = (x, \xi) \in \mathfrak{R}^{\tilde{n}} : c^k \geq (A^k)^T y \}$$

dado por  $m_k = 2\tilde{n} + j_k$  desigualdades.

$$\text{Fazer } s^k = c^k - (A^k)^T y^k > 0$$

### Passo 2 (critério de parada)

Se  $x^k$  é uma solução desejada do problema (PC).

então parar,

senão vai para o passo 3.

### Passo 3 (geração de corte)

Caso 1 Se  $h(x^k) \leq 0$  e  $f(x^k) < \xi_k$

Fazer: corte horizontal:

$$C_{\tilde{n}}^{k+1} = f(x^k)$$

$$C_j^{k+1} = C_j^k \text{ para } j = 1, \dots, m_k \text{ com } j \neq \tilde{n}$$

$$A^{k+1} = A^k \text{ (não houve geração de coluna)}$$

$$j_{k+1} = j_k \text{ (não houve corte subgradiente, nem vertical)}$$

Caso 2 Se  $h(x^k) \leq 0$  e  $\xi_k \leq f(x^k)$

Fazer corte subgradiente f:

$$a_{m_k+1} = \begin{pmatrix} g^k \\ -1 \end{pmatrix}$$

$$C_{m_k+1} = \begin{pmatrix} g^k \\ -1 \end{pmatrix}^T \begin{pmatrix} x^k \\ f(x^k) \end{pmatrix} = g^k x^k - f(x^k)$$

$$A^{k+1} = (A^k, a_{m_k+1})$$

$$C^{k+1} = \begin{pmatrix} c^k \\ C_{m_k+1} \end{pmatrix}$$

$$j_{k+1} = j_k + 1 \text{ (houve corte subgradiente na iteração k)}$$

$$k(j_{k+1}) = k \text{ (o último corte não horizontal aconteceu na iteração k)}$$

### caso 3

Se  $h(x^k) > 0$  (não definido  $f(x^k)$ )

Subcaso 3.1 Se  $h(x^k) \geq \xi_k$ . Fazer corte subgradiente h:

$$a_{m_k+1} = \begin{pmatrix} \mu^k \\ -1 \end{pmatrix}$$

$$C_{m_k+1} = \begin{pmatrix} \mu^k \\ -1 \end{pmatrix} \begin{pmatrix} x^k \\ h(x^k) \end{pmatrix} = \mu^k x^k - h(x^k)$$

$$A^{k+1} = (A^k, a_{m_k+1})$$

$$C^{k+1} = \begin{pmatrix} c^k \\ C_{m_k+1} \end{pmatrix}$$

$$j_{k+1} = j_k + 1$$

$$k(j_{k+1}) = k$$

subcaso 3.2 Se  $0 < h(x^k) < \xi_k$ .

Fazer corte horizontal:

$$a_{m_k+1} = \begin{pmatrix} \mu^k \\ -1 \end{pmatrix}$$

$$C_{m_k+1} = \begin{pmatrix} \mu^k \\ 0 \end{pmatrix}^T \begin{pmatrix} x^k \\ \xi_k \end{pmatrix} = \mu^k x^k$$

$$A^{k+1} = (A^k, a_{m_k+1})$$

$$C^{k+1} = \begin{pmatrix} c^k \\ C_{m_k+1} \end{pmatrix}$$

$$j_{k+1} = j_k$$

Passo 4

$$k := k+1,$$

ir para passo 1.

**OBSERVAÇÕES:**

1) No caso 3 do algoritmo quando  $h(x^k) \geq \xi_k$  temos o corte:

$$H^+ : a_{m_k+1}^T y \leq a_{m_k+1}^T \hat{y}^k$$

$$\text{onde } a_{m_k+1} = \begin{pmatrix} \mu^k \\ -1 \end{pmatrix}, y = \begin{pmatrix} x \\ \xi \end{pmatrix}, \hat{y}^k = \begin{pmatrix} x^k \\ h(x^k) \end{pmatrix}$$

contém ao epígrafo de  $f$

$$\text{epi } f = \{y = (x, \xi) \in \mathfrak{R}^n : h(x) \leq 0, f(x) \leq \xi, 0 \leq x \leq e\}$$

De fato:

Seja  $\tilde{y} = (\tilde{x}, \tilde{\xi}) \in \text{epi } f$ , então  $h(\tilde{x}) \leq 0$  e  $f(\tilde{x}) \leq \tilde{\xi}$ ,  $0 \leq \tilde{x} \leq e$

$$H^+ = \{y \in \mathfrak{R}^n : a_{m_k+1}^T y \leq a_{m_k+1}^T \hat{y}^k\} = \{y \in \mathfrak{R}^n : \mu^k (x - x^k) + h(x^k) \leq \xi\}$$

pela convexidade de  $h$  em  $\mathfrak{R}^n$ , temos que

$$\mu^k (\tilde{x} - x^k) + h(x^k) \leq h(x) \text{ para todo } x \in \mathcal{R}^n .$$

$$\text{logo } \mu^k (\tilde{x} - x^k) + h(x^k) \leq h(\tilde{x})$$

$$\text{mas } h(\tilde{x}) \leq 0 \leq f(\tilde{x}) \leq \tilde{\xi} \text{ pois } f^* \geq 0$$

$$\text{logo } \mu^k (\tilde{x} - x^k) + h(x^k) \leq \tilde{\xi}$$

portanto  $(\tilde{x}, \tilde{\xi}) \in H^+$ .

2) No caso 3, quando  $0 < h(x^k) < \xi_k$ , o corte vertical  $H^+ : a_{m_k+1}^T y \leq a_{m_k+1}^T y^k$

$$\text{onde } a_{m_k+1} = \begin{pmatrix} \mu^k \\ 0 \end{pmatrix}, y = \begin{pmatrix} x \\ \xi \end{pmatrix} y^k = \begin{pmatrix} x^k \\ \xi_k \end{pmatrix}$$

contém ao epígrafo de  $f$ , de fato :

$$\text{Se } (\tilde{x}, \tilde{\xi}) \in \text{epif}, \text{ então } h(\tilde{x}) \leq 0, 0 \leq \tilde{x} \leq e, f(\tilde{x}) \leq \tilde{\xi}$$

pela convexidade de  $h$  temos que  $h(\tilde{x}) \geq h(x^k) + \mu^k (\tilde{x} - x^k)$  então

$$0 \geq h(\tilde{x}) \geq h(x^k) + \mu^k (\tilde{x} - x^k) \geq \mu^k (\tilde{x} - x^k) \quad (\text{pois } h(x^k) > 0)$$

assim  $\mu^k (\tilde{x} - x^k) \leq 0$ , por tanto  $(\tilde{x} - x^k) \in H^+$ .

## 6.5 REDUÇÃO DO MAX-POTENCIAL

Seja  $\Omega = \{y \in \mathbb{R}^{\tilde{n}} : c \geq A^T y\}$

$$= \{y \in \mathbb{R}^{\tilde{n}} : s = c - A^T y \geq 0\}$$

um polítopo  $\tilde{n}$ -dimensional, onde  $c \in \mathbb{R}^m$ ,  $A \in \mathbb{R}^{\tilde{n} \times m}$ , e seja  $y^a$  o centro analítico do polítopo  $\Omega$ , ( $y^a \in \text{int}(\Omega)$ ).

Sabemos que  $y^a$  é o ponto que minimiza a função barreira logarítmica

$$F(y) = - \sum_{j=1}^m \ln(c_j - \alpha_j^T y) = - \sum_{j=1}^m \ln s_j$$

$y \in \text{int}(\Omega)$ , onde  $\alpha_j$  denota a coluna  $j$  de  $A$ ,  $s_j = c_j - \alpha_j^T y > 0$

O potencial de  $\Omega$  é definido por

$$p(\Omega) = -F(y^a) = \sum_{j=1}^m \ln(c_j - \alpha_j^T y^a) = \sum_{j=1}^m \ln s_j^a$$

onde  $s^a = c - A^T y^a > 0$ .

Observar que  $p(\Omega) = \max \{-F(y) : y \in \text{int}(\Omega)\}$  daí que  $p(\Omega)$  é também chamado de max-potencial de  $\Omega$ .

Se no polítopo

$$\Omega = \{y \in \mathbb{R}^{\tilde{n}} : \alpha_j^T y \leq c_j, j = 1, \dots, m\}$$

mudamos o lado direito da última desigualdade da definição de  $\Omega$ , por exemplo,

por  $\alpha_m^T y \leq \alpha_m^T y^a + \beta s_m^a$  onde  $\beta$  é um parâmetro, gera-se um novo polítopo.

$$\Omega_\beta^+ = \{y \in \mathbb{R}^{\tilde{n}} : \alpha_j^T y \leq c_j, j = 1, \dots, m-1, \alpha_m^T y \leq \alpha_m^T y^a + \beta s_m^a\}$$

No lema que estabelecemos a continuação; mostramos redução do máximo potencial de  $\Omega_\beta^+$  em relação ao máximo potencial de  $\Omega$ .

As demonstrações dos lemas 6.3 e 6.4 estão contidas em [47].

**Lema 6.3** Seja  $\Omega = \{y \in \mathbb{R}^{\tilde{n}} : \alpha_j^T y \leq c_j, j = 1, \dots, m\} = \{y \in \mathbb{R}^{\tilde{n}} : c \geq A^T y\}$

$$\Omega_\beta^+ = \{y \in \mathbb{R}^{\tilde{n}} : \alpha_j^T y \leq c_j \text{ para } j = 1, \dots, m-1, \alpha_m^T y \leq \alpha_m^T y^a + \beta s_m^a\}$$

então

$$p(\Omega_\beta^+) \leq p(\Omega) - 1 + \beta \text{ para } \beta \geq 0 \text{ ou } \beta < 0.$$

prova

Se  $y^a$  é o centro analítico de  $\Omega$ , então  $-\mathbf{A}(\mathbf{S}^a)^{-1} \mathbf{e} = \nabla_y F(y^a) = 0$ , ( $\mathbf{e}$  é um vetor de componentes todos iguais a 1) onde  $(\mathbf{S}^a)^{-1}$  é uma matriz diagonal.  $(\mathbf{S}^a)^{-1} = \text{diag}(1/s_j^a)$   
 Se definimos  $x^a = (\mathbf{S}^a)^{-1} \mathbf{e} > 0$ , temos que

$$\mathbf{A}x^a = 0 \quad \text{é} \quad X^{-a} s^a = \mathbf{e}, \quad (6.1)$$

onde  $X^{-a}$  é a matriz diagonal  $X^{-a} = \text{diag}(x_j^a)$

Seja  $\bar{c}_j = c_j$  para  $j = 1, \dots, m-1$

$$\bar{c}_m = \alpha_m^T y^a + \beta s_m^a$$

$\bar{y}^a$  o centro analítico de  $\Omega_\beta^+$

$$\bar{s}^a = \bar{c} - \mathbf{A}^T \bar{y}^a$$

então temos

$$\begin{aligned} e^T X^a \bar{s}^a &= e^T X^a (\bar{c} - \mathbf{A}^T \bar{y}^a) = e^T X^a \bar{c} \quad (\text{pois } \mathbf{A}x^a = 0) \\ &= e^T X^a \bar{c} - (1-\beta)x_n^a s_n^a = n-1 + \beta \end{aligned}$$

Sabemos que

$$P(\Omega) = \sum_{j=1}^m \ln s_j^a = \ln\left(\prod_{j=1}^m s_j^a\right)$$

$$P(\Omega_\beta^+) = \sum_{j=1}^m \ln \bar{s}_j^a = \ln\left(\prod_{j=1}^m \bar{s}_j^a\right)$$

então,

$$\exp(p(\Omega)) = \prod_{j=1}^m s_j^a, \quad \exp(p(\Omega_\beta^+)) = \prod_{j=1}^m \bar{s}_j^a$$

logo,

$$\begin{aligned} \frac{\exp(p(\Omega_\beta^+))}{\exp(p(\Omega))} &= \prod_{j=1}^m \frac{\bar{s}_j^a}{s_j^a} = \prod_{j=1}^m \bar{s}_j^a x_j^a \quad (\text{pois } X^a s^a = \mathbf{e}) \\ &\leq (1/m \sum_{j=1}^m \bar{s}_j^a x_j^a)^m = ((m-1+\beta)/m)^m \leq \exp(\beta-1) \end{aligned}$$

isto é,

$$\exp(p(\Omega_\beta^+)) \leq \exp(p(\Omega)) \exp(\beta-1)$$

ou equivalentemente,

$$p(\Omega_\beta^+) \leq p(\Omega) + \beta - 1$$

para qualquer  $\beta \geq 0$  ou  $\beta < 0$ .

#### Lema (6.4)

Seja  $\Omega = \{y \in \mathbb{R}^n: a_j^T y \leq c_j, j = 1, \dots, m\}$  e

$$\Omega_\beta^+ = \{y \in \mathbb{R}^n: a_j^T y \leq c_j \text{ para } j = 1, \dots, m, a_{m+1}^T y \leq a_{m+1}^T y^a + \beta \bar{r}\}$$

onde  $\beta$  é um parâmetro ( $\beta \geq 0$  ou  $\beta < 0$ ) e

$$\bar{r} = \sqrt{a_{m+1}^T (A(S^a)^{-2} A^T)^{-1} a_{m+1}}$$

onde  $s^a = c - A^T y^a > 0$ ,  $y^a$  o centro analítico de  $\Omega$

$$(S^a)^{-2} = \text{diag}(1/(s_j^a)^2).$$

Então,

$$P(\Omega_\beta^+) \leq P(\Omega) + \ln \bar{r} - \bar{\alpha} + \max\{\beta, 0\}$$

para  $\beta \geq 0$  ou  $\beta < 0$ , sendo  $\bar{\alpha} = 1, 5 - \ln 4 > 0 (\approx 0,1137)$

#### prova

Seja  $y^a$  o centro analítico de  $\Omega$ , então existe  $x^a > 0$  tal que

$$Ax^a = 0 \text{ e } X^{-a} s^a = e \text{ onde}$$

$$s^a = c - A^T y^a > 0, \text{ e seja } \beta \geq 0.$$

Seja  $\bar{y}^a$  o centro analítico de  $\Omega_\beta^+$ ;

$$\bar{s}^a = c - A^T \bar{y}^a > 0 \text{ (as primeiras } m \text{ folgas do novo centro } \bar{y}^a)$$

$$\bar{s}_{m+1}^a = a_{m+1}^T \bar{y}^a + \beta \bar{r} - a_{m+1}^T \bar{y}^a = a_{m+1}^T (y^a - \bar{y}^a) + \beta \bar{r}$$

Então, temos

$$\bar{s}_{m+1}^a = a_{m+1}^T (y^a - \bar{y}^a) + \beta \bar{r}$$

$$\begin{aligned}
&= \alpha_{m+1}^T (A(X^a)^2 A^T)^{-1} (A(X^a) A^T) (y^a - \bar{y}^a) + \beta \bar{r} \\
&= \alpha_{m+1}^T (A(X^a)^2 A^T)^{-1} A(X^a)^2 (A^T y^a - A^T \bar{y}^a) + \beta \bar{r} \\
&= \alpha_{m+1}^T (A(X^a)^2 A^T)^{-1} A(X^a)^2 (-c + A^T y^a + c - A^T \bar{y}^a) + \beta \bar{r} \\
&= \alpha_{m+1}^T (A(X^a)^2 A^T)^{-1} A(X^a)^2 (\bar{s}^a - s^a) + \beta \bar{r} \\
&= \alpha_{m+1}^T (A(X^a)^2 A^T)^{-1} A(X^a) (X^a \bar{s}^a - e) + \beta \bar{r} \\
&\leq \|\alpha_{m+1}^T (A(X^a)^2 A^T)^{-1} A X^a\| \|X^a \bar{s}^a - e\| + \beta \bar{r} \\
&= (\|X^a \bar{s}^a - e\| + \beta) \bar{r}
\end{aligned}$$

isto é,

$$\bar{s}_{m+1}^a \leq (\|X^a \bar{s}^a - e\| + \beta) \bar{r} \quad (6.2)$$

Temos que,

$$\begin{aligned}
e^T X^a \bar{s}^a &= e^T X^a (c - A^T \bar{y}^a) \\
&= e^T X^a c - (e^T X^a A^T \bar{y}^a) \\
&= e^T X^a c - (x^{aT} A^T) \bar{y}^a && \text{(pois } e^T X^a = \\
& && x^{aT}) \\
&= e^T X^a c - (A x^a)^T y^a \\
&= e^T X^a (c - A^T y^a) && \text{(pois } A x^a = 0) \\
&= e^T X^a c \\
&= e^T X^a s^a \\
&= e^T e = m
\end{aligned}$$

ou seja

$$e^T X^a \bar{s}^a = m \quad (6.3)$$

Assim, temos que



$$\begin{aligned}
\frac{\exp p(\Omega_{\beta}^+)}{\bar{r} \exp p(\Omega)} &= \frac{\prod_{j=1}^{m+1} \bar{s}_j^a}{\bar{r} \prod_{j=1}^m s_j^a} = \left( \frac{\bar{s}_{m+1}^a}{\bar{r}} \right) \prod_{j=1}^m \left( \frac{\bar{s}_j^a}{s_j^a} \right) = \left( \frac{\bar{s}_{m+1}^a}{\bar{r}} \right) \prod_{j=1}^m \bar{s}_j^a x_j^a \\
&= (\|X^a \bar{s}^a - e\| + \beta) \prod_{j=1}^m \bar{s}_j^a x_j^a \quad \text{por (6.2)} \\
&= (\|\alpha - e\| + \beta) \prod_{j=1}^m \alpha_j = \Psi(\alpha) \quad (6.4)
\end{aligned}$$

onde  $\alpha = X^a \bar{s}^a > 0$ , que pelo resultado(6.3)

satisfaz  $e^T \alpha = m$ .

Consideremos o seguinte problema

$$\begin{cases} \text{maximizar } \Psi(\alpha) \\ \text{s.a} \\ e^T \alpha = m \\ \alpha > 0 \end{cases}$$

Seja  $\mu = \|\alpha - e\|$  fixado. Então este problema é relacionado ao problema

$$\begin{cases} \text{maximizar } (\mu + \beta) \prod_{j=1}^m \alpha_j \\ \text{s.a} \\ e^T \alpha = m \\ \|\alpha - e\| = \mu \\ \alpha > 0 \end{cases}$$

cujos máximo é atingido, sem perda de generalidade, em

$$\hat{\alpha} = (\alpha_1, \dots, \alpha_m) \text{ tal que}$$

$$\alpha_1 = \delta > 1, \alpha_2 = \dots = \alpha_m = \frac{m - \delta}{m - 1} > 0$$

Assim,

$$\alpha - e = (\delta - 1, \frac{1 - \delta}{m - 1}, \dots, \frac{1 - \delta}{m - 1}), \text{ com } (\delta - 1) > 0$$

que implica,

$$\|\alpha - e\| = (\delta - 1) \sqrt{1 + \frac{m-1}{(m-1)^2}} = (\delta - 1) \sqrt{1 + \frac{1}{m-1}} = (\delta - 1) \sqrt{\frac{m}{m-1}},$$

$$\prod_{j=1}^m \alpha_j = \alpha_1 \left( \prod_{j=1}^{m-1} \alpha_j \right) = \delta \left( \frac{m-\delta}{m-1} \right)^{m-1}$$

Assim,

$$\begin{aligned} \Psi(\alpha) \leq \Psi(\hat{\alpha}) &= ((\delta - 1) \sqrt{\frac{m}{m-1}} (\beta)) \delta \left( \frac{m-\delta}{m-1} \right)^{m-1} \\ &= \left( 4 \sqrt{\frac{m}{m-1}} \right) \frac{(\delta - 1) + \sqrt{\frac{m-1}{m}} (\beta)}{2} (\delta/2) \left( \frac{m-\delta}{m-1} \right)^{m-1} \\ &\leq 4 \sqrt{\frac{m}{m-1}} \left( \frac{(m-0,5) + \sqrt{\frac{m-1}{m}} (\beta)}{m+1} \right)^{m+1} \\ &\leq \frac{4}{\exp(1,5 - \sqrt{\frac{m-1}{m}} \beta)} < \frac{4}{\exp(1,5 - \beta)} \end{aligned}$$

e tendo em conta (6.4), temos

$$\frac{\exp p(\Omega_{\beta}^+)}{\bar{r} \exp p(\Omega)} = \Psi(\alpha) \leq \frac{4}{\exp(1,5 - \beta)}$$

ou seja

$$\exp p(\Omega_{\beta}^+) \leq (4 \bar{r} \exp p(\Omega)) \frac{4}{\exp(1,5 - \beta)}$$

o que implica:

$$p(\Omega_{\beta}^+) \leq p(\Omega) + \ln 4 \bar{r} - 1,5 + \beta = p(\Omega) + \ln \bar{r} - \bar{\alpha} + \beta$$

onde  $\bar{\alpha} = 1,5 - \ln 4$ , para  $\beta \geq 0$

Veamos agora o caso  $\beta < 0$ . Para  $\beta = 0$ ,

$$\Omega_0^+ = \{ y \in \mathfrak{R}^n : \alpha_j^T y \leq c_j, j = 1, \dots, m-1, \alpha_m^T y \leq \alpha_m^T y^a \}$$

Assim para  $\beta < 0$ , temos que  $\Omega_{\beta}^+ \subset \Omega_0^+$

logo

$$p(\Omega_0^+) = \max_{y \in \text{int}(\Omega_0^+)} \{-F(y)\} \geq \max_{y \in \text{int}(\Omega_\beta^+)} \{-F(y)\} = p(\Omega_\beta^+)$$

e também para  $\beta = 0$  pelo caso anterior

$$p(\Omega_0^+) \leq p(\Omega) + \ln \bar{r} - \bar{\alpha}$$

portanto,

$$p(\Omega_\beta^+) \leq p(\Omega_0^+) \leq p(\Omega) + \ln \bar{r} - \bar{\alpha}.$$

Veremos agora para cada caso do passo 3 do algoritmo(p.v.ε) conseguimos redução do max-potencial.

i) No caso1. quando  $h(x^k) \leq 0$  e  $f(x^k) < \xi_k$ . O corte prévio horizontal  $\xi \leq c_{\bar{n}}^k$

(correspondente a coluna  $\mathcal{A}_{\bar{n}} = e^{\bar{n}} = (0, \dots, 1)$  de  $A^k$ ,  $\mathcal{A}_{\bar{n}}^T y \leq c^k$ ) é substituído pelo novo corte horizontal  $\xi \leq f(x^k)$ .

Temos que

$$\xi \leq f(x^k) = \xi_k + \frac{(f(x^k) - \xi_k)}{s_{\bar{n}}^k} s_{\bar{n}}^k = \xi_k + \beta_k s_{\bar{n}}^k$$

onde  $\beta_k = \frac{(f(x^k) - \xi_k)}{s_{\bar{n}}^k} < 0$ , e

$$s_{\bar{n}}^k = c_{\bar{n}}^k - \mathcal{A}_{\bar{n}}^T y^k = c_{\bar{n}}^k - y_{\bar{n}}^k = c_{\bar{n}}^k - \xi_k > 0$$

isto é, o novo corte profundo (pois  $\beta_k < 0$ ) é

$$\xi \leq \xi_k + \beta_k s_{\bar{n}}^k$$

logo o novo polítopo é

$$\begin{aligned} \Omega^{k+1} &= \{y \in \mathfrak{R}^{\bar{n}} : (A^{k+1})^T y \leq C^{k+1}\} = \\ &= \{y \in \mathfrak{R}^{\bar{n}} : \mathcal{A}_j^T y \leq C_j^k \text{ para } j=1, \dots, m_k, j \neq \bar{n}, \xi \leq \xi_k + \beta_k s_{\bar{n}}^k\} \end{aligned}$$

onde  $\xi \leq \xi_k + \beta_k s_{\bar{n}}^k$  corresponde à restrição

$$\mathcal{A}_{\bar{n}}^T y \leq \mathcal{A}_{\bar{n}}^T y^k + \beta_k s_{\bar{n}}^k \text{ com } \mathcal{A}_{\bar{n}} = e^{\bar{n}}, y^k = \begin{pmatrix} x^k \\ \xi_k \end{pmatrix}$$

portanto pelo lema 6.3, Temos redução do max-potencial:

$$p(\Omega^{k+1}) \leq p(\Omega^k) - 1 + \beta_k \leq p(\Omega^k) - 1 \quad (\text{pois } \beta_k < 0)$$

Observar também que no caso 1

$$C_{\bar{n}}^{k+1} = f(x^k) < \xi_k < C_{\bar{n}}^k \text{ e como } c_{\bar{n}}^0 = 1 \quad (\text{pelo passo 0})$$

então por indução

$$C_{\bar{n}}^k \leq 1 \text{ para todo } k$$

ii) No caso 2, temos  $h(x^k) \leq 0$  e  $\xi_k \leq f(x^k)$

$$\begin{aligned} C_{m_k+1} &= (g^k)^T x^k f(x^k) = \begin{pmatrix} g^k \\ -1 \end{pmatrix}^T \begin{pmatrix} x^k \\ \xi_k \end{pmatrix} + \xi_k - f(x^k) \\ &= a_{m_k+1}^T y^k + \xi_k - f(x^k) \leq a_{m_k+1}^T y^k \end{aligned}$$

isto é

$$C_{m_k+1} \leq a_{m_k+1}^T y^k$$

Seja

$$\begin{aligned} r_k &= \sqrt{a_{m_k+1}^T (A^k (S^k)^{-2} (A^k)^T)^{-1} a_{m_k+1}} \\ &= \sqrt{a_{m_k+1}^T (\nabla^2 F(y^k))^{-1} a_{m_k+1}} \\ &= \|a_{m_k+1}\|_{H_k^{-1}} \end{aligned}$$

Onde  $H_k = \nabla^2 F(y^k)$  é uma matriz simétrica definida positiva. A matriz  $(S^k)^{-2} = \text{diag}(\frac{1}{(s_i^k)^2})$  onde  $s^k = C^k - (A^k)^T y^k > 0$  logo  $r_k > 0$ .

Então, podemos expressar

$$C_{m_k+1} = a_{m_k+1}^T y^k + \left( \frac{C_{m_k+1} - a_{m_k+1}^T y^k}{r_k} \right) r_k = a_{m_k+1}^T y^k + \beta_k r_k$$

onde

$$\begin{aligned} \beta_k &= \frac{C_{m_k+1} - a_{m_k+1}^T y^k}{r_k} \\ &= \frac{a_{m_k+1}^T y^k + \xi_k - f(x^k) - a_{m_k+1}^T y^k}{r_k} \end{aligned}$$

$$= \frac{\xi_k - f(x^k)}{r_k} \leq 0$$

assim

$$C_{m_k+1} = a_{m_k+1}^T y^k + \beta_k r_k \quad \text{com } \beta_k \leq 0$$

logo

$$\Omega^{k+1} = \{y \in \mathfrak{R}^n : (A^k)^T y \leq C^k, a_{m_k+1}^T y \leq a_{m_k+1}^T y^k + \beta_k r_k\}$$

onde  $\beta_k \leq 0$

Então como  $\beta_k \leq 0$ , pelo lema 6.4 temos que redução de max-potencial

$$\begin{aligned} p(\Omega^{k+1}) &\leq p(\Omega^k) + \ln r_k - \bar{\alpha} + \max(\beta_k, 0) \\ &\leq p(\Omega^k) + \ln r_k - \bar{\alpha}. \end{aligned}$$

com  $\bar{\alpha} = 1,5 - \ln 4 > 0$

### iii) No caso 3

a) Se  $h(x^k) \geq \xi_k$  : temos que

$$C_{m_k+1} = a_{m_k+1}^T y^k \frac{C_{m_k+1} - a_{m_k+1}^T y^k}{r_k} r_k = a_{m_k+1}^T y^k + \beta_k r_k$$

onde

$$\begin{aligned} r_k &= \sqrt{a_{m_k+1}^T (A^k (S^k)^{-2} (A^k)^T)^{-1} a_{m_k+1}} \\ \beta_k &= \frac{C_{m_k+1} - a_{m_k+1}^T y^k}{r_k} = \frac{a_{m_k+1}^T y^k + \xi_k - h(x^k) - a_{m_k+1}^T y^k}{r_k} \\ &= \frac{\xi_k - h(x^k)}{r^k} < 0 \quad (\text{pois } \xi_k - h(x^k) < 0) \end{aligned}$$

portanto pelo lema (6.4) para  $\beta_k < 0$  tem-se que

$$p(\Omega^{k+1}) \leq p(\Omega^k) + \ln r_k - \bar{\alpha}, \quad \text{com } \bar{\alpha} = 1,5 - \ln 4 > 0.$$

b) Se  $0 < h(x^k) < \xi_k$

$$\Omega^{k+1} = \{y \in \mathfrak{R}^n : (A^k)^T y \leq C^k, a_{m_k+1}^T y \leq a_{m_k+1}^T y^k\}$$

onde  $a_{m_{k+1}} = \begin{pmatrix} \mu^k \\ 0 \end{pmatrix}$ ,  $y^k = \begin{pmatrix} x^k \\ \xi_k \end{pmatrix}$ ,  $c_{m_{k+1}}^k = a_{m_{k+1}}^T y^k = \mu^k x^k$

$$a_{m_{k+1}}^T y \leq a_{m_{k+1}}^T y^k \quad \Leftrightarrow \mu^k x \leq \mu^k x^k$$

$$\Leftrightarrow \mu^k (x - x^k) \leq 0$$

como o corte passa pelo centro analítico temos que  $\beta_k = 0$ . Logo temos redução do max-potencial:

$$p(\Omega^{k+1}) \leq p(\Omega^k) + \ln r_k - \bar{\alpha}.$$

## 6.6 COMPLEXIDADE DO ALGORITMO(p.v.ε)

Supondo satisfeitas as condições expostas na seção 6.4, faremos uma análise da complexidade do algoritmo(PVε).

Suponhamos que o algoritmo não tenha terminado. Então na k-ésima iteração temos que

$$A^k = (I, -I, a_{2\tilde{n}+1}, \dots, a_{2\tilde{n}+j_k})$$

O número de colunas geradas depende do valor de  $j_k$ .

Observar que no passo 0, temos que  $x^0 = \frac{1}{2}e$  e

$$\xi_0 = y_{\tilde{n}}^0 = \frac{1}{2} < 1 = f(\frac{1}{2}e) = f(x^0)$$

então houve corte não horizontal na iteração  $k=1$ , assim  $j_k = 1$ .

$$\text{Seja } B^k = 8I + \sum_{j=1}^{j_k} \frac{a_{2\tilde{n}+j} a_{2\tilde{n}+j}^T}{\|a_{2\tilde{n}+j}\|_1^2}, B^0 = 8I$$

onde  $I$  é matriz identidade.

Os enunciados dos lemas (sem demonstração) estão contidas em [27]. Nas demonstrações modificamos as dadas em [21].

### LEMA 6.5.

Sejam  $y \in \text{int}(\Omega^k)$ , e  $s_j^k = c_j^k - \alpha_j^T y > 0$  para  $j = 1, \dots, 2\tilde{n} + j_k = m_k$ . Então

$$0 < s_j^k < 1 \text{ para } j = 1, \dots, 2\tilde{n} \text{ e } 0 < s_j^k < \|\alpha_j\|_1, \text{ para } j = 2\tilde{n}+1, \dots, 2\tilde{n} + j_k$$

**Prova:** Sabemos que  $0 < y < e$

i) Para  $j = 1, \dots, \tilde{n}, j \neq \tilde{n}$ , temos que  $c_j^k = 1, \alpha_j = e^j$  e  $s_j^k = c_j^k - \alpha_j^T y = 1 - y_j < 1$

ii) Para  $j = \tilde{n}$ , temos  $\alpha_j = -e^j$ . Logo  $0 < s_j^k = c_j^k - \xi \leq 1 - \xi < 1$  (pois  $c_{\tilde{n}}^k \leq 1$  para todo  $k$ ,  $\xi > 0$ )

iii) Para  $j = \tilde{n}+1, \dots, 2\tilde{n}$ , temos que  $c_j^k = 0, \alpha_j = -e^j$ , então  $0 < s_j^k = 0 - y_j < 1$

iv) Para  $j = 2\tilde{n}+1, \dots, 2\tilde{n} + j_k = m_k$ :

$$\text{Seja } \ell = k(j-2\tilde{n})$$

Temos no caso 2. do passo 3 que  $a_j = \begin{pmatrix} g^\ell \\ -1 \end{pmatrix}$ ,

$$c_j^k = \begin{pmatrix} g^\ell \\ -1 \end{pmatrix}^\top \begin{pmatrix} x^\ell \\ f(x^\ell) \end{pmatrix}$$

logo

$$\begin{aligned} s_j^k = c_j^k - a_j^\top y &= \begin{pmatrix} g^\ell \\ -1 \end{pmatrix}^\top \begin{pmatrix} x^\ell \\ f(x^\ell) \end{pmatrix} - \begin{pmatrix} g^\ell \\ -1 \end{pmatrix}^\top \begin{pmatrix} x \\ \xi \end{pmatrix} \\ &= \xi - f(x^\ell) - g^\ell(x - x^\ell) \leq \xi - \xi_1 - g^\ell(x - x^\ell) \\ &\leq - \begin{pmatrix} g^\ell \\ -1 \end{pmatrix}^\top \begin{pmatrix} x - x^\ell \\ \xi - \xi_\ell \end{pmatrix} \\ &\leq \|a_j\|_1 \|y - y^\ell\| \leq \|a_j\|_1 \end{aligned}$$

onde

$$y^\ell = \begin{pmatrix} x^\ell \\ \xi_\ell \end{pmatrix}, y = \begin{pmatrix} x \\ \xi \end{pmatrix}.$$

A última desigualdade vem de

$$\|y - y^\ell\|_1 = \max\{\|x_i - x_i^\ell\|, |\xi - \xi_\ell|\} \leq 1, \text{ pois } y \text{ e } y^\ell \text{ estão na caixa } 0 \leq y \leq e$$

No caso 3:

a) Quando  $h(x^\ell) > 0$  e  $h(x^\ell) \geq \xi_\ell$ , temos definindo

$$a_j = \begin{pmatrix} \mu^\ell \\ -1 \end{pmatrix},$$

$$c_j^k = \begin{pmatrix} \mu^\ell \\ -1 \end{pmatrix} \begin{pmatrix} x^\ell \\ h(x^\ell) \end{pmatrix}$$

logo,

$$\begin{aligned} s_j^k = c_j^k - a_j^\top y &= \xi - h(x^\ell) - \mu^\ell(x - x^\ell) \\ &\leq \xi - \mu^\ell(x - x^\ell) \\ &= \begin{pmatrix} \mu^\ell \\ -1 \end{pmatrix}^\top \begin{pmatrix} x^\ell - x \\ -\xi \end{pmatrix} \\ &\leq \|a_j\|_1 \|y - y^\ell\| \leq \|a_j\|_j \end{aligned}$$

onde  $\|y - y^\ell\|_1 \leq 1$

b) Se para algum  $1 \leq \ell \leq j_k, j = 2\tilde{n} + \ell$  tem-se que  $h(x^j) > 0$  e  $h(x^j) < \xi_j$  temos que

$$a_j = \begin{pmatrix} \mu^j \\ 0 \end{pmatrix}, c_j^k = \begin{pmatrix} \mu^j \\ 0 \end{pmatrix} \begin{pmatrix} x^j \\ \xi_j \end{pmatrix} = \mu^j x^j$$



$$\begin{aligned}
s_j^k &= c_j^k - a_j^T y \\
&= \mu^j x^j - \mu^j x \\
&= \mu^j (x^j - x) \\
&\leq \|\mu^j\|_1 \|x^j - x\|_1 \leq \|\mu^j\|_1 = \|a_j\|_1
\end{aligned}$$

**LEMA 6.6.**

Seja  $\Omega^k = \{y \in \mathbb{R}^{\tilde{n}} : (A^k)^T y \leq c^k\}$ ,  $y^k$  o centro analítico de  $\Omega^k$ ,  $s^k = c^k - (A^k)^T y^k > 0$

$$e \quad B^k = 8I + \sum_{j=1}^k \frac{a_{2\tilde{n}+j} a_{2\tilde{n}+j}^T}{\|a_{2\tilde{n}+j}\|_1^2}, \quad B^0 = 8I$$

Então

$$A^k (S^k)^{-2} (A^k)^T \succeq B^k$$

Isto é

$A^k (S^k)^{-2} (A^k)^T - B^k$  é positiva semidefinida.

**prova:**

Seja  $Y^k = \text{diag}(y^k)$ , onde  $y^k = \begin{pmatrix} x^k \\ \xi_k \end{pmatrix}$

como pelo lema 6.5 temos que

$$s_j^k = c_j^k - a_j^T y^k = 1 - y_i^k \text{ para } j=1, \dots, \tilde{n}, \text{ com } j \neq \tilde{n}$$

$$s_j^k = c_j^k - a_j^T y^k = 0 - (-e^{\tilde{n}})^T y^k = y_j^k \text{ para } j = \tilde{n}+1, \dots, 2\tilde{n}$$

$$e \quad s_{\tilde{n}}^k = c_{\tilde{n}}^k - a_{\tilde{n}}^T y^k = c_{\tilde{n}}^k - \xi_k \leq 1 - \xi_k \leq 1 = \|a_{\tilde{n}}\|_1, \text{ com } a_{\tilde{n}} = e^{\tilde{n}}$$

logo

$$\begin{aligned}
A^k (S^k)^{-2} (A^k)^T &= \nabla^2 F(y^k) \\
&= \sum_{i=1}^{2\tilde{n}+k} \frac{a_i a_i^T}{(s_i^k)^2} \\
&= (y^k)^{-2} + (I - y^k)^{-2} + \frac{e^{\tilde{n}} (e^{\tilde{n}})^T}{(s_{\tilde{n}}^k)^2} + \sum_{j=1}^k \frac{a_{2\tilde{n}+j} a_{2\tilde{n}+j}^T}{\|a_{2\tilde{n}+j}\|_1^2}
\end{aligned}$$

$$\begin{aligned}
&\succeq (y^k)^{-2} + (I - y^k)^{-2} + \sum_{j=1}^{j_k} \frac{a_{2\tilde{n}+j} a_{2\tilde{n}+j}^T}{\|a_{2\tilde{n}+j}\|_1^2} \\
&\succeq 8I + \sum_{j=1}^{j_k} \frac{a_{2\tilde{n}+j} a_{2\tilde{n}+j}^T}{\|a_{2\tilde{n}+j}\|_1^2} \quad (\text{pois } 0 \leq y^k \leq e) \\
&= B^k
\end{aligned}$$

portanto,

$$A^k (S^k)^{-2} (A^k)^T \succeq B^k.$$

Se  $j_{k+1} > j_k$ , isto é se um corte subgradiente e gerado,

Seja

$$\omega_k = \sqrt{a_{m_k+1}^T (B^k)^{-1} a_{m_k+1}} \geq \sqrt{a_{m_k+1}^T (A(S^k)^{-2} A^T)^{-1} a_{m_k+1}} = r_k$$

então pelo lema 6.6. temos que  $\omega_k^2 \geq r_k^2$  para todo  $k$ .

**LEMA 6.7.** Seja  $v_k = \|a_{m_k+1}\|_1 = \|a_{2\tilde{n}+j_{k+1}}\|_1$  para cada  $k$ . Então,

$$\sum_{j=0}^{j_{k+1}} \left( \frac{\omega_j^2}{v_j^2} \right) \leq 2\tilde{n} \ln \left( 1 + \frac{j_{k+1}}{8\tilde{n}} \right)$$

**prova:**

$$B^{k+1} = B^k + \frac{a_{2\tilde{n}+j_{k+1}} a_{2\tilde{n}+j_{k+1}}^T}{\|a_{2\tilde{n}+j_{k+1}}\|_1^2}$$

Temos que

$$\begin{aligned}
\det(B^{k+1}) &= \det \left( B^k + \frac{a_{2\tilde{n}+j_{k+1}} a_{2\tilde{n}+j_{k+1}}^T}{\|a_{2\tilde{n}+j_{k+1}}\|_1^2} \right) \\
&= \left( 1 + \frac{a_{2\tilde{n}+j_{k+1}}^T (B^k)^{-1} a_{2\tilde{n}+j_{k+1}}}{\|a_{2\tilde{n}+j_{k+1}}\|_1^2} \right) \det B^k \\
&= \left( 1 + \frac{\omega_k^2}{v_k^2} \right) \det B^k
\end{aligned}$$

logo

$$\ln(\det B^{k+1}) = \ln(\det B^k) + \ln \left( 1 + \frac{\omega_k^2}{v_k^2} \right) \quad (6.5)$$

como,  $B^k \succeq 8I$  então

$$\omega_k^2 \leq 1/8 \quad a_{2\tilde{n}+j_{k+1}}^T \cdot a_{2\tilde{n}+j_{k+1}} \leq 1/8 \quad \|a_{2\tilde{n}+j_{k+1}}\|_1^2 \leq 1/8 \quad v_k^2$$

isto é

$$0 < \frac{\omega_k^2}{v_k^2} \leq 1/8 < 1$$

Sabemos que para  $0 < x < 1$

$$\ln(1+x) = (x - \frac{x^2}{2}) + (\frac{x^3}{3} - \frac{x^4}{4}) + \dots$$

sendo cada termo não negativo. Logo

$$\ln(1+x) \geq x - \frac{x^2}{2} \geq x - \frac{x^2}{2(1-x)}$$

assim para  $x = \frac{\omega_k^2}{v_k^2} \in (0,1)$  temos

$$\begin{aligned} \ln(1 + \frac{\omega_k^2}{v_k^2}) &\geq \frac{\omega_k^2}{v_k^2} - \frac{(\frac{\omega_k^2}{v_k^2})^2}{2(1 - \frac{\omega_k^2}{v_k^2})} \\ &= \frac{\omega_k^2}{v_k^2} (1 - \frac{\frac{\omega_k^2}{v_k^2}}{2(1 - \frac{\omega_k^2}{v_k^2})}) \\ &\geq \frac{\omega_k^2}{2v_k^2} \end{aligned}$$

portanto de (6.5)

$$\ln(\det B^{k+1}) \geq \ln(\det B^k) + \frac{\omega_k^2}{2v_k^2}, \quad \text{para todo } k$$

logo ,

$$\begin{aligned} \ln(\det B^{k+1}) &\geq \ln(\det B^0) + \sum_{j=0}^{j_{k+1}} \frac{\omega_j^2}{2v_j^2} \\ &= \tilde{n} \ln 8 + \sum_{j=0}^{j_{k+1}} \frac{\omega_j^2}{2v_j^2} \end{aligned} \quad (6.6)$$

É conhecido que:

$$(\det B^{k+1})^{1/\tilde{n}} \leq \frac{\text{traco } B^{k+1}}{\tilde{n}}$$

e tem-se que

$$\text{traco } B^{k+1} = 8\tilde{n} + \sum_{j=0}^{j_{k+1}} \frac{\|a_{2\tilde{n}+j}\|_2^2}{\|a_{2\tilde{n}+j}\|_1^2}$$

mas,  $\|a_{2\tilde{n}+j}\|^2 \leq \|a_j\|_1^2$  ou seja  $\frac{\|a_{2\tilde{n}+j}\|^2}{\|a_j\|_1^2} \leq 1$  para

cada  $j = 1, \dots, j_{k+1}$ . Portanto

$$(\det B^{k+1})^{1/\tilde{n}} \leq \frac{\text{traco } B^{k+1}}{\tilde{n}} \leq 8 + \frac{j_{k+1}}{\tilde{n}}$$

o que implica que

$$1/\tilde{n} \ln(\det B^{k+1}) \leq \ln\left(8 + \frac{j_{k+1}}{\tilde{n}}\right)$$

isto é,  $\ln(\det B^{k+1}) \leq \tilde{n} \ln\left(8 + \frac{j_{k+1}}{\tilde{n}}\right)$

Então de (6.6) segue que

$$\tilde{n} \ln 8 + \sum_{j=0}^{j_{k+1}} \frac{\omega_j^2}{2v_j^2} \leq \tilde{n} \ln\left(8 + \frac{j_{k+1}}{\tilde{n}}\right)$$

de onde

$$\sum_{j=0}^{j_{k+1}} \frac{\omega_j^2}{2v_j^2} \leq \tilde{n} \left(\ln\left(8 + \frac{j_{k+1}}{\tilde{n}}\right) - \ln 8\right) = \tilde{n} \ln\left(1 + \frac{j_{k+1}}{8\tilde{n}}\right)$$

isto é

$$\sum_{j=0}^{j_{k+1}} \frac{\omega_j^2}{v_j^2} \leq 2\tilde{n} \ln\left(1 + \frac{j_{k+1}}{8\tilde{n}}\right).$$

### LEMA 6.8 .

Seja

$$\Omega^k = \{y \in \mathfrak{R}^{\tilde{n}} : a_j^T y \leq c_j^k, j = 1, \dots, 2\tilde{n}+j_k = m_k\}$$

Consideramos o polítopo

$$\tilde{\Omega}^k = \{y \in \mathfrak{R}^{\tilde{n}} : \tilde{a}_j^T y \leq \tilde{c}_j^k, j = 1, \dots, m_k\}$$

onde

$$\tilde{a}_j^T = \frac{a_j^T}{\|a_j\|_1} \quad \text{e} \quad \tilde{c}_j^k = \frac{c_j^k}{\|a_j\|_1}. \text{ Seja } \varepsilon \text{ tal que}$$

$0 < \varepsilon \leq \tilde{c}_j^k - \tilde{a}_j^T \tilde{y}$  para cada  $j = 1, \dots, m_k$  para algum  $\tilde{y}$ , Então

$$p(\Omega^k) - \sum_{j=2\tilde{n}+1}^{m_k} \ln \|a_j\|_1 = p(\tilde{\Omega}^k) \geq m_k \ln \varepsilon$$

**prova:**

Pela definição do max-potencial temos que

$$\begin{aligned}
 p(\Omega^k) &= \max \left\{ \sum_{j=1}^{m_k} \ln(c_j^k - a_j^T y) : y \in \text{int}(\Omega^k) \right\} \\
 p(\tilde{\Omega}^k) &= \max \left\{ \sum_{j=1}^{m_k} \ln(\tilde{c}_j^k - \tilde{a}_j^T \tilde{y}) : y \in \text{int}(\tilde{\Omega}^k) \right\} \\
 &= \max \left\{ \left( \sum_{j=1}^{m_k} \ln(c_j^k - a_j^T y) - \sum_{j=1}^{m_k} \ln \|a_j\|_1 \right) : y \in \text{int}(\Omega^k) \right\} \\
 &= \max \left\{ \left( \sum_{j=1}^{m_k} \ln(c_j^k - a_j^T y) \right) : y \in \text{int}(\Omega^k) \right\} - \sum_{j=1}^{m_k} \ln \|a_j\|_1 \\
 &= p(\Omega^k) - \left( \sum_{j=1}^{2\tilde{n}_m} \ln(1) + \sum_{j=2\tilde{n}+1}^{m_k} \ln \|a_j\|_1 \right)
 \end{aligned}$$

portanto,

$$p(\tilde{\Omega}^k) = p(\Omega^k) - \sum_{j=2\tilde{n}+1}^{m_k} \ln \|a_j\|_1$$

por outro lado,

$$p(\tilde{\Omega}^k) \geq \sum_{j=1}^{m_k} \ln(\tilde{c}_j^k - \tilde{a}_j^T \tilde{y}) \geq \sum_{j=1}^{m_k} \varepsilon = m_k \ln \varepsilon$$

o que implica que

$$p(\Omega^k) - \sum_{j=2\tilde{n}+1}^{m_m} \ln \|a_j\|_1 \geq m_k \ln \varepsilon.$$

**LEMA 6.9 .** Seja  $\varepsilon > 0$  tal que satisfaz a condição dada no enunciado do lema 6.8. Então,

$$\varepsilon^2 / \tilde{n} \leq \frac{1}{2} + 2 \ln \left( 1 + \frac{j_{k+1}}{8\tilde{n}} \right) \exp \left( (-2) \frac{k - (1 - \bar{\alpha})j_{k+1}}{2\tilde{n} + j_{k+1}} \right)$$

**prova:**

Na secção 6.5 mostramos que temos redução do max-potencial em cada iteração do algoritmo (P.V. $\varepsilon$ ) dado por

$$p(\Omega^{k+1}) \leq p(\Omega^k) - 1$$

ou

$$p(\Omega^{k+1}) \leq p(\Omega^k) + \ln r_k - \bar{\alpha}$$

logo

$$p(\Omega^{k+1}) \leq p(\Omega^0) + \sum_{j=0}^{j_k} \ln r_j - j_{k+1} (\bar{\alpha} - 1) + k$$

como pelo lema 6.8

$$p(\Omega^{k+1}) \geq (2\tilde{n} + j_k + 1) \ln \varepsilon + \sum_{j=2\tilde{n}+1}^{2\tilde{n}+j_k+1_m} \ln \|\alpha_j\|_1$$

e como  $\sum_{j=2\tilde{n}+1}^{2\tilde{n}+j_k+1_m} \ln \|\alpha_j\|_1 = \sum_{j=0}^{j_k} \ln v_j$ . Então

$$(2\tilde{n} + j_k + 1) \ln \varepsilon + \sum_{j=0}^{j_k} \ln v_j \leq p(\Omega^0) + \sum_{j=0}^{j_k} \ln r_j + j_{k+1} (\bar{\alpha} - 1) + k$$

ou seja,

$$\begin{aligned} (2\tilde{n} + j_k + 1) \ln \varepsilon - k - j_{k+1} (\bar{\alpha} - 1) &\leq \frac{1}{2} (2\tilde{n} \ln \frac{1}{4}) + \frac{1}{2} \left[ \sum_{j=0}^{j_k} \ln r_j^2 - \sum_{j=0}^{j_k} \ln v_j^2 \right] \\ &\leq \frac{1}{2} (2\tilde{n} \ln \frac{1}{4}) + \frac{1}{2} \left[ \sum_{j=0}^{j_k} \ln \omega_j^2 - \sum_{j=0}^{j_k} \ln v_j^2 \right] \quad (\text{pois } r_j^2 \leq \omega_j^2) \\ &\leq \frac{1}{2} (2\tilde{n} \ln \frac{1}{4}) + \frac{1}{2} \left[ \sum_{j=0}^{j_k} \ln \frac{\omega_j^2}{v_j^2} \right] \\ &\leq \frac{1}{2} \left[ 2\tilde{n} \ln \frac{1}{4} + \sum_{j=0}^{j_k} \ln \frac{\omega_j^2}{v_j^2} \right] \end{aligned}$$

então,

$$\ln \varepsilon + \frac{(1-\bar{\alpha})j_{k+1}}{2\tilde{n} + j_k + 1} - \frac{k}{2\tilde{n} + j_k + 1} \leq \frac{1}{2} \left[ \frac{2\tilde{n}}{2\tilde{n} + j_k + 1} \ln \frac{1}{4} + \sum_{j=0}^{j_k} \frac{1}{2\tilde{n} + j_k + 1} \ln \left( \frac{\omega_j^2}{v_j^2} \right) \right]$$

como,

$$\frac{2\tilde{n}}{2\tilde{n} + j_k + 1} + \sum_{j=0}^{j_k} \frac{1}{2\tilde{n} + j_k + 1} = \frac{2\tilde{n}}{2\tilde{n} + j_k + 1} + \frac{j_k + 1}{2\tilde{n} + j_k + 1} = 1$$

e a função  $\ln x$  é côncava, então

$$\begin{aligned} \frac{1}{2} \ln \varepsilon^2 + \frac{(1-\bar{\alpha})j_{k+1}}{2\tilde{n} + j_k + 1} - \frac{k}{2\tilde{n} + j_k + 1} &\leq \frac{1}{2} \ln \left[ \frac{2\tilde{n}(1/4)}{2\tilde{n} + j_k + 1} + \sum_{j=0}^{j_k} \frac{1}{2\tilde{n} + j_k + 1} \left( \frac{\omega_j^2}{v_j^2} \right) \right] \\ &\leq \frac{1}{2} \ln \left( \frac{\frac{\tilde{n}}{2}}{2\tilde{n} + j_k + 1} + \frac{1}{2\tilde{n} + j_k + 1} \sum_{j=0}^{j_k} \frac{\omega_j^2}{v_j^2} \right) \end{aligned}$$

$$\leq \frac{1}{2} \ln \left( \frac{\tilde{n}}{2\tilde{n} + j_k + 1} + \frac{1}{2\tilde{n} + j_k + 1} \right)^{2\tilde{n}}$$

$$\ln \left( 1 + \frac{j_{k+1}}{8\tilde{n}} \right)$$

a última desigualdade é pelo lema (6.7). Logo,

$$\ln \varepsilon^2 \leq \left( \ln \left( \frac{\tilde{n}}{2\tilde{n} + j_k + 1} + \frac{2\tilde{n}}{2\tilde{n} + j_k + 1} \ln \left( 1 + \frac{j_{k+1}}{8\tilde{n}} \right) \right) - 2 \left( \frac{k - (1 - \tilde{\alpha})j_{k+1}}{2\tilde{n} + j_k + 1} \right) \right)$$

ou seja,

$$\varepsilon^2 / \tilde{n} \leq \frac{\frac{1}{2} + 2 \ln \left( 1 + \frac{j_{k+1}}{8\tilde{n}} \right)}{2\tilde{n} + j_{k+1}} \exp \left( (-2) \frac{k - (1 - \tilde{\alpha})j_{k+1}}{2\tilde{n} + j_{k+1}} \right)$$

A seguir damos uma estimativa da complexidade do algoritmo (p.v.ε)

**Teorema 6.10**

Seja  $f^* = f(x^*)$ , onde  $f^* = \min\{f(x) : h(x) \leq 0, 0 \leq x \leq e\}$ ,  $0 \leq f^* < 1$ .

Suponhamos que  $h(x^*) < 0$ ,  $0 < x^* < e$ . Seja  $L_\infty > 0$  tal que  $\|g\|_1 \leq L_\infty$  para  $g \in \partial f(x)$ ,  $\|\mu\|_1 \leq L_\infty$  para  $\mu \in \partial h(x)$ ,  $0 < x < e$ .

E seja  $f_{rec}^k = \min\{f(x^j) : h(x^j) \leq 0, 0 \leq x^j \leq e, j < k\}$ . Então

$$f_{rec}^k - f^* \leq 2\alpha(1 + L_\infty) \left( \sqrt{\frac{\frac{\tilde{n}}{2} + 2\tilde{n} \ln\left(1 + \frac{j_k}{8\tilde{n}}\right)}{2\tilde{n} + j_k}}{\right)} \exp\left(-2 \frac{k - (1 - \bar{\alpha})j_k}{2\tilde{n} + j_k}\right)$$

para  $\alpha > 1$  suficientemente grande, tal que  $\frac{-h(x^*)}{2L_\infty} > \frac{f(1/2 e) - f^*}{2\alpha(1 + L_\infty)} > 0$

**Prova:**

Suponhamos que  $f_{rec}^k > f^*$ . É suficiente mostrar que para  $\varepsilon = \frac{f_{rec}^k - f^*}{2(1 + L_\infty)} > 0$ ,

existe algum  $\tilde{y}$  tal que  $0 < \varepsilon \leq \tilde{c}_j^k - \tilde{a}_j^T \tilde{y}$  para  $j=1, \dots, m_k$ , (6.7)

onde  $\tilde{a} = \frac{a_j}{\|a_j\|_1}$  e  $\tilde{c}_j^k = \frac{c_j^k}{\|a_j\|_1}$

Então, aplicando o lema 6.9, teremos a conclusão do teorema.

Seja  $\tilde{y} = (\tilde{x}, \tilde{\xi})$  onde  $\tilde{\xi} = f_{rec}^k - \varepsilon$  e  $\tilde{x} \in R^n$  definido por

$$\tilde{x}_i = \begin{cases} x_i^* + \varepsilon, & \text{se } x_i^* \leq \frac{1}{2} \\ x_i^* - \varepsilon, & \text{se } x_i^* > \frac{1}{2} \end{cases} \quad \text{para } i=1, \dots, n$$

A seguir mostraremos que  $\tilde{y}$  satisfaz a condição (6.7)

Sabemos que

$$\begin{aligned} f^* &\geq \min\{f(x^0) + g^0(x - x^0) : h(x) \leq 0, 0 \leq x \leq e\} \\ &\geq \min\{f(x^0) + g^0(x - x^0) : 0 \leq x \leq e\} \end{aligned}$$

onde  $x^0 = 1/2 e \in R^n$

temos que,

$$\|x - x^0\|_1 = \max\{|x_i - 1/2| : 0 \leq x_i \leq 1, i = 1, \dots, n\} \leq 1/2$$

logo,  $-g^0(x - x^0) \leq \|g^0\|_1 \|x - x^0\|_1 \leq \frac{\|g^0\|_1}{2}$



então,

$$f^* \geq \min\{f(x^0) + g^0(x - x^0) : 0 \leq x \leq e\} \geq f(x^0) - \frac{\|g^0\|_1}{2}$$

o que implica,

$$L_\infty \geq \|g^0\|_1 \geq 2(f(x^0) - f^*) \geq 2(f_{rec}^k - f^*)$$

assim,

$$2(1 + L_\infty) > 4(f_{rec}^k - f^*)$$

logo,  $\varepsilon < 1/4$ .

Da definição de  $\tilde{x}$  e o fato que  $\varepsilon < 1/4$  tem-se que :

$\varepsilon \leq \tilde{x}_i \leq (1 - \varepsilon)$  para cada  $i=1, \dots, n$  ou seja  $\varepsilon e \leq \tilde{x} \leq (1 - \varepsilon)e$ .

Seja  $\tilde{s} = c^k - (A^k)^T \tilde{y}$ ; mostraremos que  $\tilde{s} = c^k - a_j^T \tilde{y} \geq \|a_j\|_1 \varepsilon$ , para cada  $j=1, \dots, m_k$

o que implicará  $\tilde{c}_j^k - \tilde{a}_j^T \tilde{y} \geq \varepsilon > 0$ , para cada  $j=1, \dots, m_k$

a) Para cada  $j=1, \dots, \tilde{n} - 1$  temos que  $a_j^T = e^j = (0, \dots, 0, 1, 0, \dots, 0)$ ,  $c_j^k = 1$ . Logo,

$$\tilde{s}_j = c_j^k - a_j^T \tilde{y} = 1 - \tilde{y}_j = 1 - \tilde{x}_j, \quad \text{para } j \leq \tilde{n} - 1$$

como,  $\tilde{x}_j \leq \varepsilon(1 - \varepsilon)$ , então,  $\tilde{s}_j = 1 - \tilde{x}_j \geq \varepsilon = \|a_j\|_1 \varepsilon$ , para  $j=1, \dots, \tilde{n} - 1$

pois,  $\|a_j\|_1 = 1$  e  $\tilde{c}_j^k = c_j^k = 1, \tilde{a} = a_j$

b) Para  $j = \tilde{n}$ , temos

$$\tilde{s}_{\tilde{n}} = c_{\tilde{n}}^k - a_{\tilde{n}}^T \tilde{y} = c_{\tilde{n}}^k - \tilde{y}_{\tilde{n}} = c_{\tilde{n}}^k - \bar{\xi} = c_{\tilde{n}}^k - (f_{rec}^k - \varepsilon) \geq \varepsilon$$

pois,

$$c_{\tilde{n}}^k \geq f_{rec}^k. \text{ Logo } \tilde{s}_{\tilde{n}} \geq \varepsilon = \|a_{\tilde{n}}\|_1 \varepsilon$$

já que,

$$a_{\tilde{n}} = e^{\tilde{n}} = (0, \dots, 1) \in R^n$$

c) Para  $j = \tilde{n} + 1, \dots, 2\tilde{n} - 1$ , temos que

$$a_j = -e^{j - \tilde{n}}, \text{ logo, } \|a_j\|_1 = 1, c_j^k = 0.$$

Logo,

$$\tilde{s}_j = c_j^k - a_j^T \tilde{y} = 0 + (e^{j - \tilde{n}})^T \tilde{y} = \tilde{y}_{j - \tilde{n}} = \tilde{x}_{j - \tilde{n}}$$

e como

$$\tilde{x} \geq \varepsilon e, \text{ tem-se que } \tilde{x}_{j - \tilde{n}} \geq \varepsilon$$

assim  $\tilde{s}_j \geq \varepsilon = \|a_j\|_1 \varepsilon$

d) Para  $j=2\tilde{n}$ , tem-se que  $c_{2\tilde{n}}^k = 0$ ,  $a_{2\tilde{n}}^T = -e^{\tilde{n}}$

logo,

$$\tilde{s}_{2\tilde{n}} = c_{2\tilde{n}}^k - a_{2\tilde{n}}^T \tilde{y} = \tilde{\xi} = f_{rec}^k - \varepsilon$$

$$f_{rec}^k - f^* \leq f_{rec}^k, \text{ pois } f^* \geq 0$$

assim,

$$\tilde{s}_{2\tilde{n}} = f_{rec}^k - \varepsilon \geq f_{rec}^k - f^* - \varepsilon \geq \varepsilon = \|a_{2\tilde{n}}\|_1 \varepsilon$$

pois,

$$f_{rec}^k - f^* - \varepsilon \geq \varepsilon \Leftrightarrow f_{rec}^k - f^* \geq 2\varepsilon = 2 \left( \frac{f_{rec}^k - f^*}{2(1 + L_\infty)} \right) \Leftrightarrow 1 + L_\infty \geq 1$$

e) Para  $j=2\tilde{n}+1, \dots, 2\tilde{n}+j_k=m_k$

Casos:

(i) Se houve corte subgradiente  $f$ , (passo3, caso 2 do algoritmo)

seja  $\ell = k(j - 2\tilde{n})$ . Temos que

$$a_j^T = (g^\ell, -1), \quad c_j^k = (g^\ell)^T x^\ell - f(x^\ell)$$

então

$$\tilde{s}_j = c_j^k - a_j^T \tilde{y} = \tilde{\xi} - f(x^\ell) - (g^\ell)^T (\tilde{x} - x^\ell)$$

como,

$$f(x^*) \geq f(x^\ell) + (g^\ell)^T (x^* - x^\ell)$$

$$\text{logo, } -f(x^\ell) \geq -f(x^*) + (g^\ell)^T (x^* - x^\ell)$$

$$\tilde{s}_j \geq \tilde{\xi} - f(x^*) + (g^\ell)^T (x^* - x^\ell) - (g^\ell)^T (\tilde{x} - x^\ell) = \tilde{\xi} - f(x^*) - (g^\ell)^T (\tilde{x} - x^*)$$

$$\geq \tilde{\xi} - f^* - \|g^\ell\|_1 \|\tilde{x} - x^*\|_\infty = f_{rec}^k - \varepsilon - f^* - \|g^\ell\|_1 \|\tilde{x} - x^*\|_\infty$$

mas,  $\|\tilde{x} - x^*\|_\infty = \max\{|\tilde{x}_i - x^*_i| : 1, \dots, n\} \leq \varepsilon$

(pois  $\tilde{x}_i - x^*_i = \varepsilon$  se  $x^*_i \leq 1/2$  e  $\tilde{x}_i - x^*_i = -\varepsilon$  se  $x^*_i > 1/2$ )

$$\text{portanto, } \tilde{s}_j \geq f_{rec}^k - \varepsilon - f^* - \varepsilon L_\infty = (f_{rec}^k - f^*) - \varepsilon (1 + L_\infty) \\ = 2\varepsilon(1 + L_\infty) - \varepsilon(1 + L_\infty) = \varepsilon (1 + L_\infty)$$

isto é,

$$\tilde{s}_j \geq \varepsilon (1 + L_\infty) \geq \|a_j\|_1 \varepsilon \quad \text{pois } 1 + L_\infty \geq 1 + \|g^\ell\|_1 = \|a_j\|_1$$

(ii) Se houve corte subgradiente h (ver passo 3, subcaso 3.1 do algoritmo (pve))

Seja  $\ell = k(j - 2\tilde{n})$

Neste caso  $h(x^\ell) \geq \xi_\ell$ ,  $h(x^\ell) > 0$ .

$$c_j^k = (\mu^\ell)^T x^\ell - h(x^\ell), \quad a_j^T = (\mu^\ell, -1)$$

então,

$$\tilde{s}_j = c_j^k - a_j^T \tilde{y} = \xi - h(x^\ell) - (\mu^\ell)^T (\tilde{x} - x^\ell)$$

$$-h(x^\ell) \geq -h(x^*) + (\mu^\ell)^T (x^* - x^\ell) \quad (\text{pela convexidade de } h)$$

h)

assim,

$$\tilde{s}_j \geq \xi - h(x^*) - (\mu^\ell)^T (\tilde{x} - x^*)$$

e como  $h(x^*) \leq 0 \leq f(x^*) = f^*$  temos que

$$\tilde{s}_j \geq \xi - f^* - (\mu^\ell)^T (x^* - \tilde{x})$$

$$\geq (f_{rec}^k - \varepsilon) - f^* - \|\mu^\ell\|_1 \|x^* - \tilde{x}\|_\infty \quad (\text{pela definição de } \xi)$$

$$\geq (f_{rec}^k - \varepsilon) - f^* - L_\infty \varepsilon \quad (\text{pois } \|\mu^\ell\|_1 < L_\infty)$$

em forma análoga, do caso (i) concluí-se que  $\tilde{s}_j \geq \|a_j\|_1 \varepsilon$

(iii) Se houve corte vertical (ver passo 3, subcaso 3.2 do algoritmo (pve))

seja  $j = 2\tilde{n} + \ell$ , para algum  $1 \leq \ell \leq j_k$

Neste caso temos que  $0 < h(x^j) < \xi_j$

$$c_j^k = (\mu^j)^T x^j, \quad a_j^k = (\mu^j, 0)$$

logo,  $\tilde{s}_j = c_j^k - a_j^T \tilde{y} = (\mu^j)^T (x^j - \tilde{x})$

pela convexidade de h temos que  $h(\tilde{x}) \geq h(x^j) + (\mu^j)^T (\tilde{x} - x^j)$

isto é,

$$(\mu^j)^T (\tilde{x} - x^j) \leq h(\tilde{x}) - h(x^j) \leq h(\tilde{x})$$

assim,

$$\tilde{s}_j = (\mu^j)^T (x^j - \tilde{x}) \geq -h(\tilde{x})$$

o que implica que  $\frac{\tilde{s}_j}{\|a_j\|_1} \geq \frac{-h(\tilde{x})}{\|a_j\|_1}$ . Mas  $\|a_j\|_1 = \|\mu^j\|_1 \leq L_\infty$

portanto,

$$\frac{\tilde{s}_j}{\|a_j\|_1} \geq -\frac{h(\tilde{x})}{L_\infty} \quad (6.8)$$

como foi assumido que  $h(x^*) < 0$ ,  $0 < x^* < e$ , então  $x^*$  é um ponto interior do conjunto  $D = \{x \in \mathbb{R}^n : h(x) \leq 0, 0 \leq x \leq e\}$ . Sendo  $h$  contínua em  $x^*$ , (pois  $h$  é convexo em  $\mathbb{R}^n$ ) temos que  $\lim_{x \rightarrow x^*} h(x) = h(x^*)$ . Logo usando a definição do

limite, existe algum  $r > 0$  tal que:

$$|h(x) - h(x^*)| < -h(x^*) / 2$$

para  $x \in B(x^*, r) = \{x \in \mathbb{R}^n : \|x - x^*\| < r\}$

portanto,  $3/2 h(x^*) < h(x) < h(x^*)/2$ , para  $x \in B(x^*, r)$ .

$$\text{Temos que } \|x^* - \tilde{x}\| = \sqrt{\sum_{j=1}^n (x_j^* - \tilde{x}_j)^2} = \sqrt{\sum_{j=1}^n \varepsilon^2} = \sqrt{n\varepsilon}$$

$$\text{seja } \Delta = f(1/2 e) - f^* = 1 - f^* > 0.$$

Podemos escolher  $L_\infty$  suficientemente grande tal que  $\frac{\Delta}{2(1+L_\infty)} < \frac{r}{\sqrt{n}}$

$$\text{mas, } \frac{\Delta}{2(1+L_\infty)} \geq \frac{f_{rec}^k - f^*}{2(1+L_\infty)} = \varepsilon \quad \text{para todo } k \quad (\text{pois } f(1/2 e) \geq f_{rec}^k)$$

logo,  $\varepsilon < \frac{r}{\sqrt{n}}$ . Assim,  $\|\tilde{x} - x^*\| < r$  ou seja  $\tilde{x} \in B(x^*, r)$ , de onde temos que

$$\frac{3}{2} h(x^*) < h(\tilde{x}) < \frac{h(x^*)}{2}$$

Assim temos que,

$$\frac{\tilde{s}_j}{\|a_j\|_1} \geq -\frac{h(\tilde{x})}{L_\infty} > -\frac{h(x^*)}{2L_\infty}$$

sendo  $-\frac{h(x^*)}{2L_\infty} > 0$ , podemos escolher  $\alpha > 0$  suficientemente grande tal que

$$-\frac{h(x^*)}{2L_\infty} > \frac{\Delta}{2\alpha(1+L_\infty)} > 0$$

mas,

$$\frac{\Delta}{2\alpha(1+L_\infty)} \geq \frac{f_{rec}^k - f^*}{2\alpha(1+L_\infty)} = \frac{\varepsilon}{\alpha} \quad \text{para todo } k.$$

Assim de (6.8), para tal  $\alpha > 0$  teríamos  $\frac{\tilde{s}_j}{\|a_j\|_1} \geq \frac{\varepsilon}{\alpha}$

ou seja  $\tilde{s}_j \geq \varepsilon_1 \|a_j\|_1$ , com  $\varepsilon_1 = \frac{\varepsilon}{\alpha}$ .

Portanto, em vista de toda o análise feito podemos tomar  $\varepsilon_1 = \frac{\varepsilon}{\alpha}$  no lugar de  $\varepsilon$ .

(observe que  $\varepsilon_1 < \varepsilon$  se  $\alpha > 1$ ), desta forma os casos anteriores reduz-se a:

$$\tilde{s}_j \geq \|a_j\|_1 \varepsilon \geq \|a_j\|_1 \varepsilon_1.$$

## 6.7 UM ALGORITMO PARA O PROBLEMA DE PROGRAMAÇÃO CONVEXA IRRESTRITA.

Consideremos o problema de programação convexa irrestrita seguinte:

$$(P) \quad \begin{cases} \min f(x) \\ \text{s.a.} & \text{onde } f: \mathbb{R}^n \rightarrow \mathbb{R} \text{ é uma função convexa.} \\ x \in \mathbb{R}^n \end{cases}$$

Seja  $f^* = \min \{f(x) : x \in \mathbb{R}^n\}$  suponhamos que é conhecido um limitante superior  $f_{up}$  e um limitante inferior  $f_{low}$  de  $f^*$ . Então, apresentaremos um algoritmo para resolver o problema (P) por redução do gap ( $f_{up} - f_{low}$ ), sendo  $f_{up} = f(x^k)$  onde a seqüência  $\{x^k\}$  é gerado por um oráculo que usa algum algoritmo implementável para resolver um problema de viabilidade convexa.

### 6.7.1 Algoritmo (A.I)

(0) Seja  $f_{low}^0$  e  $f_{up}^0$  dados tal que

$$f_{low}^0 \leq f^* \leq f_{up}^0$$

$$f_{up} := f_{up}^0, \quad f_{low} := f_{low}^0, \quad \varepsilon > 0.$$

(1) Faça

$$f_{lev} = \frac{f_{up} + f_{low}}{2}$$

(2) Chamar ao oráculo para encontrar um ponto viável no conjunto

$$S = \{x \in \mathbb{R}^n : f(x) \leq f_{lev}\}$$

(2.1) Se S é vazio, então faça  $f_{low} := f_{lev}$

(2.2) Se o oráculo retorna um ponto  $\bar{x} \in S$ , então faça  $f_{up} := f(\bar{x})$

(3) Se  $f_{up} - f_{low} \leq \varepsilon$  parar,  $\bar{x}$  é uma solução  $\varepsilon$ -optimal do problema (P).

Senão, retornar ao passo (1).

### 6.7.2 Complexidade do algoritmo (A.I)

Suponhamos que na  $k$ -ésima iteração temos os valores  $f_{up}^k, f_{low}^k$ . Observar

que 
$$f_{up}^{k+1} - f_{low}^{k+1} \leq 1/2 (f_{up}^k - f_{low}^k) \quad \text{para todo } k.$$

#### Teorema (6.11)

O Algoritmo (A.I) termina com uma solução  $\varepsilon$ -otimal,  $\varepsilon > 0$ , num número de

iterações não maior que  $1 + \log_2 \left( \frac{f_{up}^0 - f_{low}^0}{\varepsilon} \right)$  e a complexidade do algoritmo é

de ordem

$$O\left[\left(1 + \log_2 \left( \frac{f_{up}^0 - f_{low}^0}{\varepsilon} \right)\right)T\right] \quad \text{onde } T \text{ é a complexidade de algum algoritmo}$$

para resolver o problema de viabilidade convexa.

#### Prova:

De  $f_{up}^{k+1} - f_{low}^{k+1} \leq 1/2 (f_{up}^k - f_{low}^k)$  para todo  $k$ , temos que

$$(f_{up}^k - f_{low}^k) \leq (1/2^k)(f_{up}^0 - f_{low}^0) \quad \text{para todo } k$$

$$(1/2^k)(f_{up}^0 - f_{low}^0) \leq \varepsilon \quad \text{se, e somente se, } k \geq \log_2 \left( \frac{f_{up}^0 - f_{low}^0}{\varepsilon} \right)$$

logo  $f_{up}^k - f_{low}^k \leq \varepsilon$  para  $k \geq \log_2 \left( \frac{f_{up}^0 - f_{low}^0}{\varepsilon} \right)$ , e o algoritmo pára com uma

solução  $\varepsilon$ -otimal.

## CAPÍTULO 7

### CONCLUSÕES

Neste trabalho de tese, foi apresentado um novo algoritmo de planos de corte com uma estratégia de eliminação de cortes redundantes e com dois pontos testes: O centro analítico e o centro de elipsóide, para resolver o problema geral de viabilidade convexa. Usamos os elipsóides de Löwner-John para identificar os cortes redundantes. Mostramos também que o algoritmo proposto possui complexidade polinomial  $o(n^2L)$  onde  $n$  é a dimensão do espaço e  $L$  o tamanho de dados de entrada.

Apresentamos também um algoritmo de  $O \log_2 \left( \log_2 \left( \frac{f_{up} - f_{low}}{\varepsilon} \right) \right)$  iterações

para resolver um problema de programação convexa irrestrita, onde  $f_{up}$  e  $f_{low}$  são limitantes superior e inferior do valor ótimo respectivamente. Em cada iteração deste algoritmo os pontos testes são fornecidos por um oráculo que resolve um problema de viabilidade convexa.

Foi também desenvolvido um algoritmo de planos de corte-centro analítico para um problema de programação convexa restrita desde seu alhi [27]. Este algoritmo é uma extensão do trabalho de Kiwiel [27] para um problema de programação convexa com restrição caixa. Damos uma estimativa de complexidade do algoritmo. E cabe ressaltar que usamos vários tipos de corte para encontrar um ponto no conjunto “localizador” de  $\varepsilon$ -soluções.

Embora tenha-se apresentado um algoritmo polinomial com eliminação de cortes redundantes para PVC (ver capítulo 5), fica em aberto para futuras pesquisas criar outras estratégias para identificar cortes redundantes, como também desenvolver um método de atualização dos centros analíticos aproximados no algoritmo (AER).

Não foi possível determinar com exatidão a complexidade do algoritmo (pvc) (ver capítulo 6) para o problema de programação restrita, sendo também este outro problema para futuras pesquisas.

Um outro trabalho por realizar seria fazer uma implementação dos algoritmos aqui propostos a fim de estudar sua eficiência e robustez, em especial do algoritmo (AER).

## BIBLIOGRAFIA

- [ 1] ATKINSON, D. S., PRAVIN, V. VAIDYA, M. "A Cutting Plane Algorithm for Convex Programming That uses Analytic Centers". In: *Mathematical Programming* v. 69 pp. 1-43 1995.
  
- [ 2].BAHN,O., GOFFIN, J. L., VIAL, J. P., MERLE, O. D. "Implementation and Behavior of an Interior Point Cutting Plane Algorithm for Convex Programming: An Application to Geometric Programming" Technical Report. 1991.
  
- [ 3] BAHN, O., MERLE, O., GOFFIN, J. L, VIAL, J. P. "A Cutting Plane Method from Analytic Centers for Stochastic Programming". In: *Mathematical programming*, 69, pp. 45-73, 1995.
  
- [ 4] BLAND, R. G., GOLDFARB, D., TODD, M. J. "The ellipsoid method: A survey". In: *Operations Research* v.29, pp. 1039-1091. 1981
  
- [ 5] HERTOOG, D. "Interior Point Approach to Linear Quadratic and Convex Programming". In: *Thesis the Degree of Doctor*. Delft, Holanda 1992
  
- [ 6] DEN HERTOOG, D., ROOS, D., TERLAKY, C. "A Build Up Variant of the Path-Following Method for L. P." In: *Report N° 91*, Delft University of Technology to appear in *Operations Research Letters*. 1991
  
- [ 7] DEN HERTOOG, D., ROOS, D., TERLAKY, C. "Adding and Deleting Constraints in the Logarithmic Barrier Method for Linear Programming Problems". In: Koninklijke Shell Laboratorium Amsterdam the Netherlands, Submitted to *Mathematical Programming*. 1992
- [ 8] DEN HERTOOG, D., ROOS, D., TERLAKY, C. - "A Logarithmic Barrier Cutting Plane Method for Convex Programming". In: *Report, Support of Shell*. 1993.
  
- [ 9] DIKIN T. I. "Iterative Solution of Problems of Linear and Quadratic Programming". In: *Soviet Mathematics Doklady* 8, pp. 674-675, 1967.



- [ 10] DZIUBAN, S. T., ECKER, J. G., KUPFERSCHMID, M. “Using Deep Cuts in an Ellipsoid Algorithm for Nonlinear Programming”. In: *Mathematical Programming Study* v. 25 pp. 93-107 1985.
- [11] ELZINGA, J., MOORE, T. “A central cutting plane algorithm for convex programming”. In: *Mathematical programming* v. 8, pp. 134-145 1975.
- [12] ECKER, J. G., KUPFERSCHMID, M.- “An Ellipsoid Algorithm for Non Linear Programming”. In: *Mathematical Programming* v. 27 pp. 83-106 1983-
- [13] ECKER, J. G., KUPFERSCHMID, M “A Computational Comparison of the Ellipsoid Algorithm With Several Nonlinear Programming Algorithms”. In: *Siam Journal on Control on Optimization* v. 23 pp. 657- 674 1985.
- [14] FRISCH, K. R., “The logarithmic potential method for solving linear programming problems”. In: *Memorandum, Institute of Economics*, Oslo, Norway. 1955
- [15] FERREIRA, C. E., WAKABAYASHI, Y. “Combinatoria poliedrica e planos de corte faciais”. In: *Departamento de ciência da computação IME . Universidade de São Paulo (U.S.P.)* 1966. (texto)
- [16] GONZAGA, C. C. “An Algorithm for Solving Linear Programming Problems in  $O(n^3 L)$  Operations”, In: : *Progress in Mathematical Programming, Interior Point and Related Methods*, pp.1-28 Springer Verlag, New York (1984).
- [17] GRÖTSCHEL, M., LOVÁSZ, L., SCHRIJVER, A. “The Ellipsoid Method and its Consequences in Combinatorial Optimization”. In: *Combinatoria* v. 1. pp. 169-197 1981.
- [18] GOFFIN, J.L., “Variable metric relaxation methods, Part II : The ellipsoid method”. In: *Mathematical Programming* v. 30, pp. 147-162 1984.
- [19] GOFFIN, J. L., HAURIE, A. - VIAL, J. P. “Decomposition and Nondifferentiable Optimization with the Projective Algorithm”. In: *Management Science* v. 38, pp. 284-302 1992.

- [20] GOFFIN, J. L., HAURIE, A.-VIAL, J. P. "Cutting planes and column generation technics with projective algorithm". In: *Core discussion paper 8829. Core, Universite Catholique de Louvain, Louvain, Belgium, 1988.*
- [21] GOFFIN, J. L.-LUO,Z. Q.-YE Y. "Complexity Analysis of an Interior Cutting Plane Method for Convex Feasibility Problems". In: *Departament of Management Sciences, the University of Iowa, USA 1994.Siam Journal of Optimization 1996.*
- [22] GOFFIN, J. L. "Using the Primal Dual Infeasible Newton Method in the Analytic Center Method for Problems Defined by Deep Cutting Planes". In: *Faculty of Management, McGill University, Montreal, Quebec, Canada 1994.*
- [23] GRÖTSCHEL, M., LOVÁSZ, L, SCHRJVER, A. "Geometric Algorithms and Combinatorial Optimization", (Springer, Berlin, 1988).
- [24] HUARD, P. "Resolution of Mathematical Programming with Non Linear Constraints by the Method of Centres". In: *nonlinear Programming*. pp. 207-219 1967.
- [25] KHACHIYAN, L. G. "A Polynomial Algorithm in Linear Programming". In: *Soviet Mathematics Doklady* v. 20, pp. 191-194 1979.
- [26] KARMARKAR, N. "A New Polynomial - time Algorithm for Linear Programming". In: *Combinatoria* v. 4, pp. 373-395 1984.
- [27] KIWIEL, K. C. "Complexity of Some Cutting Plane Methods that use Analytic Centers." In: *Mathematical Programming* v. 74, pp. 47-54 1996.
- [28] KELLEY, J. E. "The Cutting - Plane Method for Solving Convex Programs". In: *Journal of the Society for Industrial and Applied Mathematics* v. 8, pp. 703-712

- [29] LANCASTER, P., TIMENESTSKY, M. "The Theory of Matrices". In: *Academy Press*, New York, 2th ad. 1985.
- [30] LEVIN, A. "An Algorithm of Minization of Convex Funtions". In: *Soviet Mathematics Doklady* v. 160,.6, pp 1244-1247 1965.
- [31] LEMARECHAL, C. "Nonsmoth optimization and descent methods". In: *Research Report 78-4*, IIASA. Luxemburg, Austria, 1978.
- [32] NEWMAN, D. "Location of Maximum on Unimodal Surfaces". In: *Journal of ACM* v 12, 3, pp. 395-368 1965..
- [33] NESTEROV, Y. "Complexity estimates of Some Cutting Plane Methods Based on the Analytic Barrier". In: *Mathematical Programming*, 69, pp. 149-176, 1995.
- [34] ROSS, C. - VIAL, J. "A Polynomial Method of Approximate Centers for Linear Programming". In: *Mathematical Programming* v. 54, pp. 295-305 1992.
- [35] SHOR, N. Z. "Utilization of the Operation of the Space Dilatation in the Minimization of Convex Funtions". In: *Cybernetics* v. 6, pp. 7-15 1970.
- [36] SHOR, N. Z. "Convergence rate of the gradient descent method with Dilatation of the space". In: *Cybernetics* v. 6, pp. 102-108 1970.
- [37] SHOR, N. Z. "A cutting plane method for solving convex programming problems". In: *Cybernetics* v. 1, pp. 42-50 1977.
- [38] SONNEVEND, G. Y. "New algorithms in convex programming based on a motion of centre (for systems of analytic inequalities) and on rational extrapolation". In: *International series of numerical mathematics* pp. 311-327 1988.
- [39] SONNEVEND, G. Y. "An Analytic Center for Polyhedrons and New Classes of Global Algorithms For Linear (Smooth, Convex) Programming". In: *Lecture notes in control and information sciences* v. 84, pp. 866-878

- [40] SCHRIJVER, A., "Theory of lines on inter programming". John Wiley - sons - New York 1989.
- [41] TARASOV, S. P., KHACHIYAN, L. G., ERLICH, I. I. "The Méthod of Inscribed Ellipsoids". In: *Soviet Mathematics Doklady* v. 37, pp. 226-230 1988.
- [42] VAIDYA, P. M. "A New Algorithms for Minimizing Convex Funcions Over Convex Sets". In: *Mathematical Programming* v. 73, pp. 291-341 1996.
- [43] ZHI - LUO, Q. "Analysis of a Cutting Pane Method that Uses Weighted Analytic Center and Multiple Cuts". In: *Department of Eletrical and Computer Engenhering*, M. C. Master University Hamilton Canadá 1994.
- [44] B. YUDIN, B., D. - NEMIROVSKY, A. "Informational Complexity and Efficient Methos for the Solution of Convex Extremal Problems." In: *Matekon* v. 13 (2), pp. 3 - 25 1976.
- [45] YUDIN, B - NEMIROVSKY, A. "Evaluation of the Informational Complexity of Mathematical Programming Problems." In: *Ekonomika Matematicheskije Metody* v. 12, pp. 128-142 ( in Russian) 1976.
- [46] YE, Y. "Karmarkar's Algorithms and the Ellipsoid Method." In: *Operations Research Letters* v. 6, pp. 177-182 1987.
- [47] YE, Y. "A Potencial Reduction Algorithm Allowing Column Generation". In: *Siam J. Optimization* v. 2, pp. 7-20 1992.
- [48] YE, Y. "Complexity Analysis of the Analytic Center Cutting Plane Method that Uses Multiple Cuts". In: *Mathematical Programming*, 78, pp. 85-104, 1997.