


USANDO A TRAJETÓRIA CENTRAL PARA CALCULAR O CENTRO
ANALÍTICO DE UM POLÍTOPO APÓS A ADIÇÃO DE UM PLANO DE
CORTE PROFUNDO

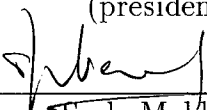
Marli Cardia

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PRO-
GRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE
FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NE-
CESSÁRIOS PARA OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS
EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

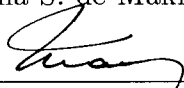
Aprovada por:



Prof. Clóvis Caesar Gonzaga, D. Sc.
(presidente)



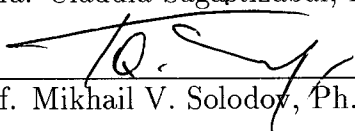
Profa. Susana S. de Makler, D. Sc.



Prof. Nelson Maculan Filho, D. Sc.



Profa. Cláudia Sagastizábal, D. Sc.



Prof. Mikhail V. Solodov, Ph. D.

RIO DE JANEIRO, RJ – BRASIL
ABRIL DE 1998

CARDIA, MARLI

Usando a Trajetória Central para Calcular o Centro Analítico de um Politopo após a Adição de um Plano de Corte Profundo [Rio de Janeiro] 1998

VIII, 106 p., 29.7 cm, (COPPE/UFRJ, D. Sc., Engenharia de Sistemas e Computação, 1998)

Tese - Universidade Federal do Rio de Janeiro, COPPE

1. Ponto Interior 2. Trajetória Central 3. Plano de Corte 4. Otimização Convexa

I. COPPE/UFRJ II. Título(Série).

*À memória
de minha mãe
Maria Ximenes Cardia*

Agradecimentos

Quero exprimir minha profunda gratidão ao Professor Clóvis C. Gonzaga pela amizade, otimismo e incentivo demonstrados ao longo do meu curso de doutorado e pela orientação segura que resultou no desenvolvimento e na conclusão deste trabalho.

Sou especialmente grata à amiga e Professora Susana S. de Makler por ter acompanhado com interesse essa minha luta, que durou alguns bons anos, sem nunca deixar faltar sua ajuda, apoio, conforto e estímulo. Também não esqueço todos os demais professores, funcionários, amigos e colegas do Programa de Engenharia de Sistemas e Computação que de perto ou de longe sempre colaboraram.

Agradeço ainda aos Professores Van Hien Nguyen e Jean-Jacques Strodiot por terem me aceitado em seu grupo de pesquisa no Departamento de Matemática da Facultés Universitaire de Namur (Bélgica), durante a fase do meu Doutorado Sanduíche, onde comecei a estudar o assunto que deu origem a esta tese.

Meus agradecimentos à CAPES e CNPq pelo apoio financeiro e à Universidade Federal do Paraná, particularmente ao Departamento de Matemática, por garantir o meu afastamento.

Finalmente, não posso deixar de mencionar que, teria sido impossível chegar ao fim da luta sem o constante suporte emocional e afetivo do meu companheiro, Luis Alberto de Almeida Pinheiro, para vencer os inúmeros momentos de incertezas.

Resumo da Tese apresentada à COPPE como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D. Sc.)

Usando a Trajetória Central para Calcular o Centro Analítico de um Polítopo após a Adição de um Plano de Corte Profundo

Marli Cardia

Abril de 1998

Orientador: Clóvis Caesar Gonzaga

Programa: Engenharia de Sistemas e Computação

Neste trabalho desenvolvemos um algoritmo para resolver o problema de adicionar um corte muito profundo a um polítopo, baseado na obtenção do centro analítico do polítopo. Este problema é o passo central em métodos de plano de corte para resolver problemas de otimização convexa. Quando cortes muito profundos são considerados, a grande dificuldade é a recuperação da viabilidade para iniciar um algoritmo de centralização. A nossa estratégia para vencer tal dificuldade é partir de um centro analítico aproximado do antigo polítopo, que é conhecido, e seguir a trajetória central de um problema de programação linear auxiliar até obter um centro analítico aproximado do novo polítopo.

Abstract of Thesis presented to COPPE as partial fulfillment of the requirements for the degree of Doctor of Science (D. Sc.)

Using the Central Path to Compute the Analytic Center of a Polytope after
Adding a Deep Cutting Plane.

Marli Cardia

April 1998

Thesis Supervisor: Clóvis Caesar Gonzaga

Department: Systems Engineering and Computer Science

In this work we develop an algorithm to solve the problema of adding a very deep cut to a polytope based on obtaining the analytic center of the polytope. This problem is the central step in cutting plane methods for solving convex optimization problems. When very deep cuts are considered the major difficulty is to retrieve feasibility to initialize a centralization algorithm. Our strategy to overcome such difficulty is to start from an approximate analytic center of the old polytope that is known and then to follow the central path of an auxiliary linear programming until obtaining an approximate analytic center of the new polytope.

Conteúdo

1	Introdução	1
1.1	Apresentação do Problema	1
1.2	Aplicação do Problema	2
1.3	Os Métodos de Plano de Corte	2
1.4	O Objetivo do Trabalho	5
1.5	Notação	6
2	Preliminares	8
2.1	Um Pouco de Álgebra Linear	8
2.2	Otimização Não-Linear	12
2.3	O Problema de Programação Linear	18
3	Centro Analítico e Trajetória Central	23
3.1	A Trajetória Central Associada ao Problema (P)	24
3.2	Centro Analítico	38
3.3	Procedimentos de centralização	41
3.4	Variação da Função Barreira ao longo da Trajetória Central . . .	43
4	O Algoritmo de Goffin-Vial para Calcular o Centro Analítico de um Politopo	51
4.1	Os Problemas Dual e Primal	52

4.2	O Algoritmo de Centralização Projetivo Primal	55
4.3	Adicionando um Corte ao Politopo Ω	58
5	Um Algoritmo para Calcular o Centro Analítico de um Politopo Usando a Trajetória Central	65
5.1	A Direção de Newton Dual	65
5.2	O Problema Auxiliar	69
5.3	Introduzindo um Corte Pouco Profundo	73
5.4	Um Algoritmo para Introdução de um Corte em Ω	76
6	Um Algoritmo de Plano de Corte	82
6.1	O Problema de Viabilidade Convexa	82
6.2	O Potencial de um Politopo	84
6.3	O Algoritmo de Plano de Corte	86
6.4	A Análise de Convergência	90
7	Comentários Finais	99
	Bibliografia	101

Capítulo 1

Introdução

1.1 Apresentação do Problema

Consideramos o seguinte problema: Dado um politopo

$$\Omega = \{y \in \mathbb{R}^m; A^t y \leq c\},$$

onde $c \in \mathbb{R}^n$ e A é uma matriz $m \times n$ de posto máximo igual a m . Dados um vetor $a \in \mathbb{R}^m$ e um número K o novo politopo resultante da adição de um corte profundo $a^t y \leq K$ é

$$\Omega_K = \{y \in \mathbb{R}^m; a^t y \leq K, y \in \Omega\}.$$

Suporemos que a é tal que $\|a\| = 1$, então $K - a^t \hat{y}$ representa a distância real de \hat{y} ao hiperplano $\{y \in \mathbb{R}^m; a^t y = K\}$. Seja y^0 um centro analítico aproximado de Ω . Um corte é dito profundo se $K \leq a^t y^0$, ou seja, o corte é dado por $a^t y \leq a^t y^0 + \gamma$, onde $\gamma \leq 0$.

Suporemos que

$$\overset{\circ}{\Omega} = \{y; A^t y < c\}$$

e

$$\overset{\circ}{\Omega}_K = \{y; a^t y < K, y \in \overset{\circ}{\Omega}\}$$

são não-vazios.

Queremos calcular um centro analítico aproximado de Ω_K iniciando em y^0 . Observemos que como y^0 não pertence ao interior de Ω_K , então nenhum algoritmo de centralização pode ser iniciado em y^0 .

1.2 Aplicação do Problema

O problema de calcular o centro analítico de um politopo é o passo central em métodos de plano de corte para resolver problemas de admissibilidade convexa, otimização convexa não-diferenciável, programação estocástica linear, programação inteira, etc.

1.3 Os Métodos de Plano de Corte

Os métodos de plano de corte foram desenvolvidos para resolver problemas de otimização convexa. Eles geram uma seqüência de problemas de programação linear que refinam as aproximações poliedrais do problema original. As desigualdades lineares que definem as aproximações são geradas por um oráculo como hiperplanos separando um ponto teste do conjunto solução.

A questão crucial no desenvolvimento de algoritmos de plano de corte é a escolha do ponto teste na aproximação poliedral. Na literatura corrente en-

contramos vários métodos de plano de corte com pontos testes diferentes. Um dos primeiros foi o método clássico de KELLEY[21], onde o ponto teste é uma solução ótima da relaxação linear, isto é, a cada iteração é resolvido um problema de programação linear. Este método tem tido um bom desempenho em alguns problemas, mas péssimo em outros. E isto tem motivado o desenvolvimento de muitos outros métodos com o objetivo de melhorar o desempenho prático e o limite de complexidade. Entre eles estão os métodos de plano de corte central, isto é, os métodos que usam um “centro” de um politopo como o ponto teste. O primeiro a aparecer foi o método do centro de gravidade de LEVIN[24], seguido pelo método de máxima esfera de ELZINGA e MOORE[5], método elipsóide de YUDIN e NEMIROVSKI[50], SHOR[41] e KHACHIYAN[22], método elipsóide de volume máximo de TARASOV *et al.*[45] e KHACHIYAN e TODD[23], método de centro volumétrico de VAIDYA[46] e método de plano de corte no centro analítico. Este é que tem feito um grande sucesso, devido à facilidade de cálculo do centro analítico e também pelo bom desempenho computacional obtido até aqui. O conceito de centro analítico de um politopo foi desenvolvido por SONNEVEND[42], que também sugeriu o seu uso em métodos de plano de corte.

Desde então, vários autores têm desenvolvido algoritmos de plano de corte usando o centro analítico, dentre eles, GOFFIN e VIAL[8] e YE[49]. Estes métodos têm sido usados e testados para uma variedade de problemas de grande escala, onde eles têm tido um desempenho encorajador, como podemos ver pelos resultados obtidos por GOFFIN *et al.*[12], que usaram um muito bem sucedido método de plano de corte no centro analítico para resolver problemas de otimização convexa não-diferenciável, BAHN *et al.*[3] aplicaram em problemas de programação geométrica, BAHN *et al.*[2] em problemas de programação estocástica linear, GOFFIN *et al.*[11] em “nonlinear multicommodity flow problems” e MITCHELL[27], [28], [29] na resolução de problemas de programação inteira.

Na análise de complexidade é fácil mostrar que um algoritmo de plano de corte no centro analítico é polinomial no número de cortes, se o oráculo gera um número finito de cortes. Mas isto não é sustentável para o caso infinito. Então, um problema difícil é mostrar que um tal algoritmo tem limite de complexidade que depende da dimensão do espaço no qual o conjunto solução está e não do número de cortes gerados. ATKINSON e VAIDYA[1] desenvolveram um algoritmo de plano de corte no centro analítico polinomial na dimensão do espaço. O algoritmo deles descarta restrições que se tornam insignificantes e isto é essencial na sua análise de complexidade. Existem outros métodos que não eliminam restrições e têm limite de complexidade dependendo da dimensão do espaço. Como por exemplo, NESTEROV[35] que usou um método híbrido envolvendo um termo adicional quadrático. GOFFIN *et al.*[13] que usaram a mesma argumentação de NESTEROV[35] junto com alguns resultados de YE[49], e sem o termo quadrático, para dar uma estimativa de complexidade para o problema de viabilidade convexa de $O^*(\frac{n^2}{\epsilon^2})$, onde n é a dimensão do espaço e ϵ é o parâmetro de precisão. MOKHTARIAN e GOFFIN[33] e GOFFIN e VIAL[9] relaxaram a exigência dos cortes serem centrais que é feita em [13], eles usaram cortes rasos, profundos e muito profundos, e obtiveram o mesmo limite de complexidade. Muito recentemente, GOFFIN e VIAL[10] consideraram a abordagem de dois cortes no seu método de plano de corte no centro analítico, obtendo ainda o mesmo limite de complexidade. Um algoritmo de plano de corte com complexidade independente da dimensão do espaço foi apresentado por NESTEROV e VIAL[37] e NESTEROV *et al.*[36], que consideraram um método de plano de corte no centro analítico num espaço projetivo e mostraram uma estimativa de complexidade que só depende do parâmetro da barreira autoconcordante para o conjunto viável. Para o caso de n muito grande, o resultado deles é ótimo, conforme NEMIROVSKY e YUDIN[34].

A maioria dos métodos de plano de corte no centro analítico dos autores

mencionados acima, em geral, não aceitam o corte fornecido pelo oráculo. Eles fazem restrição quanto a profundidade do corte, tomando-o ou muito raso, ou central, ou interceptando o elipsóide de Dikin inscrito no politopo.

1.4 O Objetivo do Trabalho

Neste trabalho desenvolveremos um algoritmo para calcular um centro analítico aproximado do novo politopo gerado por um corte tal como foi dado pelo oráculo, ou seja, um corte que pode ser muito profundo. Quando consideramos cortes muito profundos, a grande dificuldade é a recuperação da viabilidade para iniciar um algoritmo para calcular o centro analítico do novo politopo. GOFFIN e VIAL[9] já consideraram este caso e a estratégia deles, para contornar tal dificuldade, foi considerar o problema num formato primal, o dual do espaço onde está a aproximação poliedral, e então com o auxílio da direção de MITCHELL e TODD[30] recuperar a viabilidade num único passo. Os passos iterativos são passos de Newton com busca com respeito a um problema que está relacionado ao problema primal canônico de KARMAKAR[20]. Por esta razão, o algoritmo deles é chamado projetivo primal.

O nosso trabalho é totalmente diferente do de GOFFIN e VIAL[9], pois, para recuperar a viabilidade, usaremos um procedimento para seguir a trajetória central do seguinte problema de programação linear auxiliar:

$$\begin{array}{ll}
 & \text{minimizar} \quad a^t y \\
 (D_A) & \text{sujeito a:} \\
 & A^t y + s = c \\
 & s \geq 0,
 \end{array}$$

até obter um centro analítico aproximado de Ω_K . O procedimento para seguir a trajetória central pode basear-se em qualquer algoritmo de trajetória central. Na

literatura encontramos bons algoritmos de trajetória central do tipo primal-dual com passos longos, como por exemplo, MIZUNO- TODD-YE[32] e GONZAGA[18]. O algoritmo de trajetória central primal- dual mais eficiente em implementações atuais é o preditor-corretor de MEHROTRA[26]. Com uma particularização deste procedimento, onde seguimos a trajetória usando um algoritmo dual de passos curtos, desenvolveremos um algoritmo de plano de corte para resolver o problema de viabilidade convexa. A finalidade desta particularização é para estudar a complexidade do método de plano de corte e não é eficiente do ponto de vista prático.

O trabalho está organizado da seguinte forma: No capítulo 2 introduzimos alguns conceitos e resultados de álgebra linear, algumas técnicas básicas em programação não-linear e o problema de programação linear. No capítulo 3 apresentamos o conceito de trajetória central e centro analítico junto com algumas propriedades e procedimentos de centralização. Fazemos também um estudo da variação da função barreira sobre a trajetória central. No capítulo 4 descrevemos o algoritmo de Goffin-Vial para calcular o centro analítico de um politopo. No capítulo 5, está nossa contribuição, desenvolvemos um algoritmo para calcular o centro analítico de um politopo usando a trajetória central. No capítulo 6 apresentaremos um método de plano de corte, que usa uma particularização do algoritmo desenvolvido no capítulo 5, para resolver o problema de viabilidade convexa. No capítulo 7 encerramos o trabalho com a apresentação de algumas conclusões.

1.5 Notação

No que segue usaremos as seguintes notações: \mathbb{R}^n é o espaço euclidiano, \mathbb{R}_+^n é o ortante não-negativo e \mathbb{R}_{++}^n é o ortante positivo. Usaremos vetor coluna e matriz denotados por letras minúsculas e maiúsculas respectivamente. Os diferentes

vetores serão denotados por índices superiores, os índices inferiores denotarão as componentes do vetor. A letra e denotará o vetor dado por $e = [11\dots 1]^t$ com dimensão explicitada pelo contexto. A letra e com índice inferior i , e_i , denotará o vetor unitário com a i -ésima componente igual a 1. Para um vetor x a letra maiúscula X denotará a matriz diagonal formada pelas componentes do vetor x . Dados os vetores $x, s \in \mathbb{R}^n$ as notações x^{-1} , xs , $\frac{x}{s}$ serão usadas para os vetores com componentes x_i^{-1} , $x_i s_i$ e $\frac{x_i}{s_i}$, $i = 1, \dots, n$. A notação $\|x\|$ será usada para a norma euclidiana do vetor x e $x^t s$ denotará o produto interno dos vetores x e s .

Capítulo 2

Preliminares

Todo o material desenvolvido neste capítulo é fundamental. Seu objetivo é introduzir algumas ferramentas e conceitos de álgebra linear, otimização não-linear e programação linear que estarão presentes em praticamente todos os capítulos posteriores.

2.1 Um Pouco de Álgebra Linear

Um hiperplano é um tipo de conjunto convexo muito importante em otimização. Em \mathbb{R}^n , um hiperplano é um subespaço afim de dimensão $n - 1$. Sejam dados o vetor $a \in \mathbb{R}^n, a \neq 0$, e o escalar $b \in \mathbb{R}$. O conjunto

$$H(a, b) = \{x \in \mathbb{R}^n; a^t x = b\}$$

é dito um hiperplano em \mathbb{R}^n . Um hiperplano $H(a, b)$ divide o espaço em dois semi-espacos fechados dados por

$$H_+(a, b) = \{x \in \mathbb{R}^n; a^t x \geq b\},$$

$$H_-(a, b) = \{x \in \mathbb{R}^n; a^t x \leq b\}.$$

Um conjunto $P \subset \mathbb{R}^n$ que pode ser expresso como a interseção finita de semi-espacos fechados,

$$P = \{x \in \mathbb{R}^n; Ax \leq b\},$$

é chamado de poliedro convexo. Um poliedro convexo limitado é um politopo.

Considere $\Gamma \subset \mathbb{R}^n$ e $y \in \mathbb{R}^n$. Um hiperplano $H(a, b)$ é dito separador de y e Γ se, e só se, $a^t y \geq b$ e $\Gamma \subset H_-(a, b)$.

Seja $A \in \mathbb{R}^{n \times n}$ uma matriz. O polinômio $f(\lambda) = \det(A - \lambda I)$ é chamado o polinômio característico e a equação $f(\lambda) = 0$ de equação característica da matriz A . Os autovalores de A são os $\lambda \in \mathbb{R}$ tais que $Ax = \lambda x$ possui soluções não-nulas. As soluções não-nulas correspondentes a x são os autovetores de A .

Proposição 2.1.1 *O polinômio característico de uma matriz $A \in \mathbb{R}^{n \times n}$ é um polinômio de grau n com coeficiente líder $(-1)^n$ e termo constante $\det(A)$. O coeficiente de λ^{n-1} é $(-1)^{n-1} \text{tr}A$. Existem n autovalores de A , e se eles forem representados por $\lambda_1, \lambda_2, \dots, \lambda_n$, então $\sum_{i=1}^n \lambda_i = \sum_{i=1}^n a_{ii} = \text{tr}A$.*

Demonstração. (NOBLE e DANIEL[38]) □

Considere os vetores $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$. Uma matriz da forma xy^t é chamada matriz de posto-um. Quando x e y têm a mesma dimensão, então a matriz xy^t é quadrada. Uma matriz da forma

$$A = I - \alpha uv^t,$$

onde $\alpha \in \mathbb{R}$ e $u, v \in \mathbb{R}^n$, é chamada uma modificação posto-um da identidade.

Proposição 2.1.2 *Sejam $u, v \in \mathbb{R}^n$ e $A = I - \alpha uv^t$. Então A tem pelo menos $n - 1$ autovalores iguais a 1 e o outro é igual a $1 + \alpha u^t v$.*

Demonstração. (GILL et al.[7]) □

Seja $A \in \mathbb{R}^{m \times n}$ com $m < n$. Como a matriz A representa uma transformação linear de \mathbb{R}^n em \mathbb{R}^m , então podemos associar à matriz A quatro subespaços fundamentais (para mais detalhes, ver STRANG[44]):

i) O espaço coluna (ou espaço imagem) de A ,

$$\mathcal{R}(A) = \{y \in \mathbb{R}^m; y = Ax, x \in \mathbb{R}^n\}.$$

ii) O espaço nulo de A ,

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n; Ax = 0\}.$$

iii) O espaço linha de A (ou espaço coluna de A^t),

$$\mathcal{R}(A^t) = \{x \in \mathbb{R}^n; A^t y = x, y \in \mathbb{R}^m\}.$$

iv) O espaço nulo de A^t ,

$$\mathcal{N}(A^t) = \{y \in \mathbb{R}^m; A^t y = 0\}.$$

Uma propriedade geométrica muito interessante relaciona esses subespaços:

$$\mathcal{N}(A) \perp \mathcal{R}(A^t) \quad \text{e} \quad \mathcal{N}(A^t) \perp \mathcal{R}(A),$$

ou seja, $\mathcal{N}(A)$ e $\mathcal{R}(A^t)$ são subespaços ortogonais em \mathbb{R}^n e $\mathcal{N}(A^t)$ e $\mathcal{R}(A)$ são subespaços ortogonais em \mathbb{R}^m .

Os subespaços $\mathcal{N}(A)$ e $\mathcal{R}(A^t)$ desempenharão um papel muito importante no trabalho que vamos desenvolver. Sabemos que o \mathbb{R}^n é a soma direta desses dois subespaços, $\mathbb{R}^n = \mathcal{N}(A) \oplus \mathcal{R}(A^t)$. Assim, todo vetor $d \in \mathbb{R}^n$ pode ser decomposto de modo único como

$$d = d_p + \tilde{d}_p,$$

onde $d_p \in \mathcal{N}(A)$ e $\tilde{d}_p \in \mathcal{R}(A^t)$. O vetor d_p é a projeção do vetor d sobre $\mathcal{N}(A)$ e \tilde{d}_p é o complemento de d em relação a $\mathcal{N}(A)$. Como a aplicação projeção é um operador linear, ela pode ser representada por uma matriz. Então, a aplicação projeção sobre o espaço nulo de A será representada pela matriz P_A e a projeção sobre o complemento ortogonal de $\mathcal{N}(A)$ será $\tilde{P}_A = I - P_A$. Portanto, podemos escrever

$$d_p = P_A d \quad \text{e} \quad \tilde{d}_p = \tilde{P}_A d.$$

Se A é de posto máximo, então existe uma expressão para P_A que é dada por

$$P_A = I - A^t(AA^t)^{-1}A.$$

A projeção de d em $\mathcal{N}(A)$ é o ponto em $\mathcal{N}(A)$ cuja distância ao ponto d é a menor possível, ou seja,

$$d_p = \arg \min\{\|x - d\|; x \in \mathcal{N}(A)\}.$$

Também podemos definir a projeção de d em $\mathcal{N}(A)$ como o vetor de norma mínima em $\mathcal{N}(A)$ entre todos os vetores diferenças do tipo $d - A^t y$, $y \in \mathbb{R}^m$, ou seja,

$$d_p = \arg \min\{\|z\|; z = d - A^t y, y \in \mathbb{R}^m\}$$

donde

$$\|d_p\| = \min\{\|d - A^t y\|; y \in \mathbb{R}^m\}. \quad (2.1.1)$$

Proposição 2.1.3 (GONZAGA [14]) *Seja A uma matriz $m \times n$. Então*

$$\mathcal{R}(A^t) = \mathcal{N}(P_A).$$

Demonstração. Como o espaço linha de A e o espaço nulo de A são subespaços ortogonais, então dado um vetor $z \in \mathbb{R}^n$, $z = z_N + z_R$ onde $z_N \in \mathcal{N}(A)$ e $z_R \in \mathcal{R}(A^t)$. Portanto, temos que $z = z_R$ se, e só se, $z_N = 0$ o que implica $P_A z = 0$, ou seja, $z \in \mathcal{N}(P_A)$. \square

2.2 Otimização Não-Linear

Seja $Q \subset \mathbb{R}^n$ um conjunto aberto e seja $f : Q \rightarrow \mathbb{R}$ uma função diferenciável em Q com derivadas contínuas. O problema de programação não-linear sem restrição é escrito como

$$\begin{aligned} &\text{minimizar } f(x) \\ &x \in Q \end{aligned} \quad (2.2.1)$$

Observemos que Q é um aberto, em nosso caso será, geralmente, o ortante positivo. Os seguintes fatos são bastante conhecidos para este problema, ver LUENBERGER[25]:

- Se \bar{x} é solução ótima do problema sem restrição, então $\nabla f(\bar{x}) = 0$.
- Se Q é convexo e f é convexa, então a condição $\nabla f(\bar{x}) = 0$ é necessária e suficiente para a otimalidade.

O problema de programação não-linear com restrição é construído impondo ao problema (2.2.1) restrições de igualdade ou desigualdade:

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a:} && \\ &&& h_i(x) = 0 \quad i = 1, \dots, m \\ &&& h_i(x) \leq 0 \quad i = m + 1, \dots, m + r \\ &&& x \in Q \end{aligned}$$

onde as funções $h_i : Q \rightarrow \mathbb{R}, i = 1, \dots, m + r$ são diferenciáveis.

As condições de otimalidade para este problema são as condições de Karush-Kuhn-Tucker. Essas condições só podem ser aplicadas se o problema satisfizer certas condições de regularidade que não detalharemos aqui. Essas condições são satisfeitas por todos os problemas que estudaremos neste trabalho.

Os problemas gerais de programação não-linear são resolvidos por algoritmos que resolvem a cada iteração um problema fácil. A seqüência de pontos gerada pelas soluções desses problemas fáceis deve convergir para uma solução ótima do problema original.

Um problema é fácil se a sua região viável é simples e se a função objetivo é simples. Veremos a seguir alguns exemplos:

1) Minimização de uma função linear numa bola.

$$\begin{aligned} &\text{minimizar} && c^t h \\ &\text{sujeito a:} && \\ &&& \|h\| \leq 1. \end{aligned}$$

A solução é obtida usando a desigualdade de Cauchy-Schwarz: $|c^t h| \leq \|c\| \|h\|$.

Para todo h tal que $\|h\| = 1$, temos

$$|c^t h| \leq \|c\|$$

e a igualdade acontece para

$$h = -\frac{c}{\|c\|}.$$

2) Minimização de uma função linear numa bola e com restrição de igualdade.

minimizar $c^t h$

sujeito a:

$$\|h\| \leq 1$$

$$Ah = 0.$$

Considere $c = c_p + \tilde{c}_p$ onde $c_p = P_A c \in \mathcal{N}(A)$ e $\tilde{c}_p = \tilde{P}_A c \in \mathcal{R}(A^t)$. Então, $\forall h \in \mathcal{N}(A)$ tal que $\|h\| = 1$, temos que $c^t h = c_p^t h$, pois $\tilde{c}_p^t h = 0$. Portanto, $|c^t h| = |c_p^t h| \leq \|c_p\|$ e a igualdade é verificada para

$$h = \frac{-c_p}{\|c_p\|} = -\frac{P_A c}{\|P_A c\|}.$$

3) Minimização de uma função linear num elipsóide simples.

Seja D uma matriz diagonal definida positiva e consideremos o problema

minimizar $c^t h$

sujeito a:

$$h^t D^{-2} h \leq 1$$

$$Ah = 0.$$

A região viável é a interseção do espaço nulo de A com a região elipsoidal dada por $h^t D^{-2} h \leq 1$. Este elipsóide tem os eixos paralelos aos eixos coordenados e os comprimentos dados pelos elementos da diagonal de D . A região viável definida no problema é chamada de **elipsóide simples** ou **elipsóide de Dikin**. Se D é a matriz identidade então este problema coincide com o problema 2). Podemos escalar o problema de modo que o elipsóide se torne uma bola, com a mudança de variável $\bar{h} = D^{-1} h$:

$$\begin{aligned} &\text{minimizar } (Dc)^t \bar{h} \\ &\text{sujeito a:} \\ &\quad \|\bar{h}\| \leq 1 \\ &\quad AD\bar{h} = 0, \end{aligned}$$

cuja solução é dada como em 2):

$$\bar{h} = -\frac{P_{AD} Dc}{\|P_{AD} Dc\|}.$$

Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função duas vezes continuamente diferenciável. Consideremos o problema

$$\begin{aligned} &\text{minimizar } f(x) \\ &\text{sujeito a:} \\ &\quad Ax = b \\ &\quad x \geq 0 \end{aligned}$$

Vamos aplicar as técnicas dos exemplos 1), 2) e 3) à expansão de Taylor de f . A expansão de Taylor de f em torno de um ponto $x \in \mathbb{R}^n$ é

$$f(x+h) = f(x) + \nabla f(x)^t h + \frac{1}{2} h^t \nabla^2 f(x) h + o_2(x, h).$$

O gradiente da expansão de Taylor de f é

$$\nabla f(x+h) = \nabla f(x) + \nabla^2 f(x)h + o_1(x,h).$$

a) **Minimização da aproximação linear de f numa bola:**

$$\begin{aligned} & \underset{h}{\text{minimizar}} && f(x) + \nabla f(x)^t h \\ & \text{sujeito a:} && \end{aligned}$$

$$\|h\| \leq 1,$$

como em 1) a solução é dada por

$$h = -\frac{\nabla f(x)}{\|\nabla f(x)\|}.$$

Este é um resultado fundamental em otimização: **a direção de máxima descida de uma função é a direção oposta ao gradiente.**

b) **Minimização da aproximação linear de f numa bola e restrita a um hiperplano:**

$$\begin{aligned} & \underset{h}{\text{minimizar}} && f(x) + \nabla f(x)^t h \\ & \text{sujeito a:} && \end{aligned}$$

$$\begin{aligned} A(x+h) &= b \\ \|h\| &\leq 1, \end{aligned}$$

que é equivalente ao problema

$$\begin{aligned} & \underset{h}{\text{minimizar}} && \nabla f(x)^t h \\ & \text{sujeito a:} && \end{aligned}$$

$$\begin{aligned} Ah &= 0 \\ \|h\| &\leq 1, \end{aligned}$$

cuja solução é dada como em 2)

$$h = -\frac{P_A \nabla f(x)}{\|P_A \nabla f(x)\|}.$$

c) **Minimização da aproximação linear de f no elipsóide de Dikin centrado em x interior viável.**

$$\underset{h}{\text{minimizar}} \quad \nabla f(x)^t h$$

sujeito a:

$$h^t D^{-2} h \leq 1$$

$$Ah = 0,$$

cujas soluções são dadas como em 3),

$$h = -\frac{DP_{AD} D \nabla f(x)}{\|P_{AD} D \nabla f(x)\|}.$$

Esta direção é chamada de **direção de máxima descida** ou **direção afim-escala**.

d) **Minimização da aproximação quadrática de f sem restrição.**

$$\text{minimizar} \quad Q(h)$$

$$h \in \mathbb{R}^n$$

onde $Q(h) = \nabla f(x)^t h + \frac{1}{2} h^t \nabla^2 f(x) h$. Já vimos do problema (2.2.1) que h é solução ótima do problema acima se, e só se, $\nabla Q(h) = 0$. Se $\nabla^2 f(x)$ é definida positiva, então a solução do sistema linear

$$\nabla^2 f(x) h = -\nabla f(x) \tag{2.2.2}$$

é única. A direção solução

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

é chamada de **direção de Newton**.

e) **Minimização da aproximação quadrática de f restrita a um hiperplano.**

minimizar $Q(h)$

sujeito a:

$$Ah = 0,$$

onde $Q(h) = \nabla f(x)h + \frac{1}{2}h^t \nabla^2 f(x)h$. Agora usaremos o seguinte fato: Se h é uma solução ótima, então $\nabla Q(h)$ é ortogonal a $\mathcal{N}(A)$. Se a matriz Hessiana é definida positiva e A tem posto completo, então esta condição é necessária e suficiente e a solução ótima é única. O gradiente da função quadrática Q corresponde à aproximação linear do gradiente de f , dado por

$$\nabla Q(h) = \nabla f(x) + \nabla^2 f(x)h.$$

Então, a **direção de Newton projetada** para $\nabla^2 f(x)$ definida positiva é a única solução do sistema linear

$$P_A \nabla Q(h) = P_A (\nabla f(x) + \nabla^2 f(x)h) = 0,$$

que é equivalente a

$$P_A \nabla^2 f(x)h = -P_A \nabla f(x). \tag{2.2.3}$$

2.3 O Problema de Programação Linear

Consideraremos as três formulações do problema de programação linear: os problemas primal, dual e primal-dual.

O problema primal:

$$\begin{aligned}
 & \text{minimizar } c^t x \\
 (P) \quad & \text{sujeito a:} \\
 & Ax = b \\
 & x \geq 0,
 \end{aligned}$$

onde $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ e $A \in \mathbb{R}^{m \times n}$ ($m < n$) é de posto máximo.

O conjunto primal viável é dado por

$$\mathcal{X} = \{x \in \mathbb{R}^n; Ax = b, x \geq 0\}$$

e o conjunto dos pontos interiores primais

$$\overset{\circ}{\mathcal{X}} = \{x \in \mathcal{X}; x > 0\}.$$

O problema dual: é um problema associado com o problema (P)

$$\begin{aligned}
 & \text{maximizar } b^t y \\
 (D) \quad & \text{sujeito a:} \\
 & A^t y + s = c \\
 & s \geq 0,
 \end{aligned}$$

onde as variáveis $s \in \mathbb{R}^n$ são chamadas de folgas duais. Como A tem posto máximo, então o vetor y é unicamente determinado por s nas equações de viabilidade. Assim, podemos chamar uma solução dual como s , em vez de (y, s) . O conjunto dual viável é dado por

$$\mathcal{S} = \{s \in \mathbb{R}^n; A^t y + s = c \text{ para algum } y \in \mathbb{R}^m, s \geq 0\}$$

e o conjunto dos pontos interiores duais

$$\overset{\circ}{\mathcal{S}} = \{s \in \mathcal{S}; s > 0\}.$$

Relação entre os problemas (P) e (D): Dada uma solução primal viável $x \in \mathcal{X}$ e uma solução dual viável $s \in \mathcal{S}$, o **gap de dualidade** entre estas soluções é definido por

$$c^t x - b^t y.$$

O problema primal-dual: Quando os dois problemas (P) e (D) são considerados juntos eles são chamados de par primal-dual. O par $(x, s) \in \mathcal{X} \times \mathcal{S}$ é dito par primal-dual viável.

O conjunto primal-dual viável será denotado por

$$\mathcal{F} = \mathcal{X} \times \mathcal{S}$$

e o conjunto de pontos interiores primais-duais

$$\overset{\circ}{\mathcal{F}} = \overset{\circ}{\mathcal{X}} \times \overset{\circ}{\mathcal{S}}.$$

Hipótese: Suporemos $\overset{\circ}{\mathcal{F}} \neq \emptyset$.

Condições de otimalidade de Karush-Kuhn-Tucker (K-K-T): O vetor $x \in \mathbb{R}^n$ é uma solução ótima de (P) se, e só se, existem $y \in \mathbb{R}^m$ e $s \in \mathbb{R}^n$ tais que

$$\begin{aligned} x s &= 0 \\ Ax &= b \\ A^t y + s &= c \\ x, s &\geq 0 \end{aligned} \tag{2.3.1}$$

Observemos que os multiplicadores de Lagrange são denotados como as variáveis duais (y, s) . Esta escolha não é por acaso, pois se aplicarmos as condições de otimalidade ao problema dual (D), encontraremos as mesmas condições (2.3.1).

Isto define uma relação entre os problemas primal e dual. Formalmente, a condição de otimalidade dual é a seguinte:

O vetor $(y, s) \in \mathbb{R}^m \times \mathbb{R}^n$ é uma solução ótima do problema (D) se, e só se, existe um vetor $x \in \mathbb{R}^n$ tal que as condições (2.3.1) são satisfeitas para (x, y, s) .

Olhando para as condições (2.3.1) sob o ponto de vista primal-dual, podemos concluir que o vetor (x, y, s) satisfaz (2.3.1) se, e só se, x é solução ótima do problema primal (P) e (y, s) é solução ótima do problema dual (D) . Então, o vetor (x, y, s) ou (x, s) é dito solução ótima primal-dual.

Uma consequência imediata da condição de complementaridade é que $x^t s = 0$. Assim temos que se $(x, s) \in \mathcal{F}$, então

$$0 \leq x^t s = x^t (c - A^t y) = c^t x - y^t A x = c^t x - b^t y,$$

donde

$$x^t s = c^t x - b^t y \geq 0,$$

que é conhecida como a relação de dualidade fraca.

O problema primal-dual é definido como o sistema não-linear de equações e inequações (2.3.1), o qual pode ser visto como um sistema linear, bastando substituir $x s = 0$ por $x^t s = c^t x - b^t y = 0$.

Os problemas escalados

Consideremos os problemas de programação linear definidos acima. Seja $D \in \mathbb{R}^{n \times n}$ uma matriz diagonal definida positiva. Uma mudança de escala é uma mudança de variáveis definida pelas transformações

$$\bar{x} = D^{-1} x \quad \text{e} \quad \bar{s} = D s.$$

Aplicando esta mudança de variáveis em (P) e (D) e definindo

$$\bar{A} = AD \quad \text{e} \quad \bar{c} = Dc,$$

temos

$$\begin{aligned} & \text{minimizar} \quad \bar{c}^t \bar{x} \\ (\bar{P}) \quad & \text{sujeito a:} \\ & \bar{A} \bar{x} = b \\ & \bar{x} \geq 0, \end{aligned}$$

$$\begin{aligned} & \text{maximizar} \quad b^t y \\ (\bar{D}) \quad & \text{sujeito a:} \\ & \bar{A}^t y + \bar{s} = \bar{c} \\ & \bar{s} \geq 0, \end{aligned}$$

$$\begin{aligned} & \bar{x} \bar{s} = 0 \\ (\overline{PD}) \quad & \bar{A} \bar{x} = b \\ & \bar{A}^t y + \bar{s} = \bar{c} \\ & \bar{x}, \bar{s} \geq 0. \end{aligned}$$

Capítulo 3

Centro Analítico e Trajetória Central

O conceito de centro analítico de um politopo de SONNEVEND[42], vai desempenhar um papel fundamental no nosso trabalho.

Vamos dedicar este capítulo ao estudo da trajetória central associada ao problema (P) e das propriedades dos pontos centrais, destacando estas propriedades para o centro analítico de um politopo junto com alguns procedimentos de centralização. Estudaremos também a variação da função barreira ao longo da trajetória central. Todos os resultados apresentados nas seções 3.1, 3.2 e 3.3 foram demonstrados em GONZAGA[16].

3.1 A Trajetória Central Associada ao Problema (P)

A função barreira

Começamos com a função barreira que foi utilizada pela primeira vez em otimização em 1955 por FRISCH[6]. Esta função possui uma propriedade especial, a função cresce indefinidamente perto da fronteira do conjunto viável, e pode ser usada como uma penalidade para os pontos da fronteira. Definimos a função barreira por

$$x \in \mathbb{R}_{++}^n \mapsto p(x) = - \sum_{i=1}^n \ln x_i.$$

Existe a derivada de p para todo $x \in \mathbb{R}_{++}^n$,

$$\nabla p(x) = -x^{-1} \quad \text{e} \quad \nabla^2 p(x) = X^{-2}.$$

A função penalizada do problema (P).

A função penalizada associada com o problema (P) é definida por

$$\alpha \geq 0, x \in \overset{\circ}{\mathcal{X}} \mapsto f_\alpha(x) = \alpha c^t x + p(x),$$

onde α é chamado de parâmetro de penalidade.

A função f_α é estritamente convexa em $\overset{\circ}{\mathcal{X}}$, pois $\nabla^2 f_\alpha(x) = X^{-2}$ é definida positiva, e cresce indefinidamente próximo da fronteira do conjunto limitado \mathcal{X} . Então, podemos associar a cada $\alpha \geq 0$ o ponto central

$$x(\alpha) = \arg \min_{x \in \overset{\circ}{\mathcal{X}}} f_\alpha(x).$$

A curva definida por

$$\alpha \geq 0 \mapsto x(\alpha)$$

é chamada de **trajetória central primal de** (P) . Observemos que se $\alpha = 0$, então o ponto central é definido como o centro analítico do politopo \mathcal{X} . Na seção 3.2 estudaremos com mais detalhes este conceito e suas propriedades.

Um algoritmo de trajetória central para resolver o problema (P) é aquele que gera uma seqüência $\alpha^k \rightarrow +\infty$ e uma seqüência (x^k) em $\overset{\circ}{\mathcal{X}}$ tal que cada x^k está próximo de $x(\alpha^k)$. Quando $\alpha^k \mapsto +\infty$, o custo se torna cada vez menor em f_α e será minimizado.

Propriedades dos pontos centrais primais.

O ponto central $x(\alpha)$, para $\alpha \geq 0$, é a única solução ótima do problema

$$(P)_\alpha \quad \begin{array}{l} \text{minimizar } f_\alpha(x) \\ \text{sujeito a:} \\ Ax = b \\ x \geq 0. \end{array}$$

A condição de otimalidade: $x = x(\alpha) \in \overset{\circ}{\mathcal{X}}$ é a solução ótima de (P_α) se, e só se, $P_A \nabla f_\alpha(x) = 0$. Mas, $P_A \nabla f_\alpha(x) = 0$ é equivalente a

$$\nabla f_\alpha(x) = A^t y,$$

para algum $y \in \mathbb{R}^m$.

A expansão de Taylor de f_α em x é dada por:

$$f_\alpha(x+h) = f_\alpha(x) + \nabla f_\alpha(x)^t h + \frac{1}{2} h^t \nabla^2 f_\alpha(x) h + o_2(x, h).$$

Como $\nabla f_\alpha(x)^t h = 0, \forall h \in \mathcal{N}(A)$, e $\nabla^2 f_\alpha(x) = X^{-2}$, então

$$f_\alpha(x+h) - f_\alpha(x) = \frac{1}{2}h^t X^{-2}h + o_2(x,h), \forall h \in \mathcal{N}(A).$$

Isto leva a um resultado muito interessante: **a aproximação quadrática de f_α em $x = x(\alpha)$ tem curvas de nível definidas por $h^t X^{-2}h = \text{constante}$.** Estas curvas de nível são precisamente os elipsóides de Dikin definidos no capítulo anterior.

Mudança de escala.

Seja $x^0 \in \overset{\circ}{\mathcal{X}}$ um ponto dado e seja $D = \text{diag } x^0$. A mudança de escala $\bar{x} = D^{-1}x$ transforma x^0 no ponto e e a função penalizada no espaço escalado é dada por:

$$\bar{f}_\alpha(\bar{x}) = \alpha \bar{c}^t \bar{x} - \sum_{i=1}^n \ln x_i + \sum_{i=1}^n \ln x_i^0.$$

O último termo desta expressão é uma constante, então, dados dois pontos $x^1, x^2 \in \overset{\circ}{\mathcal{X}}$, temos

$$f_\alpha(x^2) - f_\alpha(x^1) = \bar{f}_\alpha(\bar{x}_2) - \bar{f}_\alpha(\bar{x}_1).$$

Conclusão: uma mudança de escala não afeta variações da função penalizada. O ponto x^0 é transformado em e e temos que

$$\nabla \bar{f}_\alpha(e) = \alpha \bar{c} - e \quad \text{e} \quad \nabla^2 \bar{f}_\alpha(e) = I,$$

que é muito simples! E mais ainda, a aproximação quadrática da função penalizada escalada é

$$\bar{f}_\alpha(e+h) - \bar{f}_\alpha(e) \approx \frac{\|h\|^2}{2}.$$

Portanto, a variação da função penalizada (antes ou depois da mudança de escala) fornece uma excelente medida da distância de um ponto x a $x(\alpha)$.

Infelizmente, esta distância só pode ser medida se o ponto central é conhecido, o que raramente é possível. Veremos a seguir aproximações para esta distância que são mais fáceis de calcular.

Seja $\bar{h}(x, \alpha) = -P_{AX}X\nabla f_\alpha(x)$ a direção de máxima descida escalada para minimizar f_α iniciando em x . Seja $h(x, \alpha) = X\bar{h}(x, \alpha)$ a direção de máxima descida no espaço original.

Medida de proximidade primal.

Dados $x \in \overset{\circ}{\mathcal{X}}$ e $\alpha > 0$, a medida de proximidade primal de x ao ponto central $x(\alpha)$ é dada por:

$$\delta(x, \alpha) = \|\bar{h}(x, \alpha)\|.$$

Considere $0 < \delta < 0.5$. Um ponto $x \in \overset{\circ}{\mathcal{X}}$ é dito central δ -aproximado em relação ao problema $(P)_\alpha$ se $\delta(x, \alpha) \leq \delta$.

Proposição 3.1.1 $h(x, \alpha)$ coincide com a direção de Newton projetada para minimizar $f_\alpha(\cdot)$ a partir de x .

Demonstração. De (2.2.3) temos que

$$P_A(\nabla^2 f_\alpha(x)h + \nabla f_\alpha(x)) = 0,$$

então a direção de Newton deve satisfazer para algum $y \in \mathbb{R}^m$,

$$\nabla f_\alpha(x) + \nabla^2 f_\alpha(x)h = A^t y, \quad h \in \mathcal{N}(A).$$

Assim,

$$\nabla f_\alpha(x) + X^{-2}h = A^t y.$$

Multiplicando a igualdade por X e como $\bar{h} = X^{-1}h$,

$$X\nabla f_\alpha(x) + \bar{h} = XA^t y, \quad \bar{h} \in \mathcal{N}(AX),$$

logo,

$$X\nabla f_\alpha(x) = (-\bar{h}) + XA^t y.$$

Portanto, $\bar{h} = -P_{AX} X\nabla f_\alpha(x) = \bar{h}(x, \alpha)$. □

Propriedades primais-duais dos pontos centrais.

Se um ponto central é conhecido, então uma solução dual pode ser associada a ele.

Proposição 3.1.2 *O vetor x é o ponto central primal associado com $\alpha > 0$ se, e só se, $s = \frac{1}{\alpha}x^{-1}$ é o ponto central dual associado com α . Neste caso, o gap de dualidade entre as soluções dual e primal é dado por $x^t s = \frac{n}{\alpha}$.*

Demonstração. (\Rightarrow) Seja $x = x(\alpha)$ para um dado $\alpha > 0$. Então, $x \in \overset{\circ}{\mathcal{X}}$ e para algum $y \in \mathbb{R}^m$, temos que

$$\nabla f_\alpha(x) = A^t y.$$

Como $\nabla f_\alpha(x) = \alpha c - x^{-1}$, então

$$c - \frac{x^{-1}}{\alpha} = A^t \frac{y}{\alpha}$$

fazendo

$$s = \frac{x^{-1}}{\alpha} \quad \text{e} \quad \bar{y} = \frac{y}{\alpha},$$

obtemos,

$$s > 0 \quad \text{e} \quad A^t \bar{y} + s = c,$$

donde $s \in \mathcal{S}$.

Agora vamos provar que s é o ponto central dual associado a $\alpha > 0$. Da simetria dos problemas primal e dual, temos que o problema dual é equivalente ao problema

$$\begin{aligned} &\text{minimizar} && \tilde{x}^t s \\ &\text{sujeito a:} && \\ &&& P_A s = P_A c \\ &&& s \geq 0 \end{aligned}$$

onde $\tilde{x} \in \mathcal{X}$ e $\tilde{x}^t s = c^t \tilde{x} - b^t y$.

Temos que mostrar que $\tilde{P}_A \nabla f_\alpha(s) = 0$ (lembrando que $P_{P_A} = \tilde{P}_A$, ver proposição 2.1.3), que é equivalente a

$$\nabla f_\alpha(s) = \alpha \tilde{x} - s^{-1} \in \mathcal{N}(A).$$

Como $s = \frac{x^{-1}}{\alpha}$, então $\frac{s^{-1}}{\alpha} = x$, donde segue que

$$A\left(\tilde{x} - \frac{s^{-1}}{\alpha}\right) = A \tilde{x} - Ax = b - b = 0.$$

Portanto, s é um ponto central dual associado com $\alpha > 0$.

(\Leftarrow) é provada de modo análogo.

$$\text{O gap de dualidade é } x^t s = \frac{x^t x^{-1}}{\alpha} = \frac{n}{\alpha}. \quad \square$$

Um ponto $(x, s) \in \mathbb{R}^n \times \mathbb{R}^n$ será dito ponto central primal-dual associado com $\alpha > 0$ se, e só se, (x, s) é uma solução primal-dual viável e $xs = \frac{e}{\alpha}$. Em outras palavras, (x, s) é um ponto central primal-dual se as seguintes condições

são satisfeitas:

$$(Q)_\alpha \quad \begin{aligned} xs &= \frac{e}{\alpha} \\ Ax &= b \\ A^t y &= c \\ x, s &> 0. \end{aligned}$$

Como A é de posto máximo então a solução de $(Q)_\alpha$ é única.

Medida de proximidade primal-dual.

Podemos definir uma proximidade primal-dual para um dado par primal-dual. Já sabemos que (x, s) é central se $\alpha xs = e$. Então a proximidade primal-dual mede o erro dessa condição.

Dados $x \in \overset{\circ}{\mathcal{X}}$, $s \in \overset{\circ}{\mathcal{S}}$ e $\alpha > 0$, a proximidade primal-dual de (x, s) a $(x(\alpha), s(\alpha))$ é dada por

$$\delta(x, s, \alpha) = \|\alpha x s - e\|.$$

A próxima proposição relaciona as proximidades primal e primal-dual.

Proposição 3.1.3 *Sejam dados $x \in \overset{\circ}{\mathcal{X}}$ e $\alpha > 0$. Então*

$$\delta(x, \alpha) = \min\{\delta(x, s, \alpha); s \text{ é folga dual}\}$$

ou, equivalentemente,

$$\delta(x, \alpha) = \min\{\|\alpha xs - e\|; A^t y + s = c, y \in \mathbb{R}^m\},$$

(observe que não é exigido $s \geq 0$).

Demonstração. Como $\bar{h}(x, \alpha) = -P_{AX} X \nabla f_\alpha(x)$ e de (2.1.1) temos

$$\delta(x, \alpha) = \|P_{AX} X \nabla f_\alpha(x)\| = \min\{\|X \nabla f_\alpha(x) - X A^t y\|; y \in \mathbb{R}^m\}.$$

Como $X\nabla f_\alpha(x) = \alpha Xc - e$ e fazendo $\bar{y} = \frac{y}{\alpha}$, temos

$$\begin{aligned}\delta(x, \alpha) &= \min\{\|\alpha Xc - \alpha XA^t\bar{y} - e\|; \bar{y} \in \mathbb{R}^m\} \\ &= \min\{\|\alpha X(c - A^t\bar{y}) - e\|; \bar{y} \in \mathbb{R}^m\} \\ &= \min\{\|\alpha Xs - c\|; A^t\bar{y} + s = c, \bar{y} \in \mathbb{R}^m\}\end{aligned}$$

□

Proposição 3.1.4 *Sejam $x \in \overset{\circ}{\mathcal{X}}$ e $\alpha > 0$ tais que $\delta(x, \alpha) < 1$. Defina*

$$s(x, \alpha) = \frac{x^{-1}}{\alpha}(e - \bar{h}(x, \alpha)).$$

i) $s(x, \alpha)$ é uma folga dual viável.

ii) $\delta(x, s(x, \alpha), \alpha) = \delta(x, \alpha)$.

Demonstração.

i) Da hipótese que $\delta(x, \alpha) = \|\bar{h}(x, \alpha)\| < 1$, segue imediato que $s(x, \alpha) > 0$. Por definição,

$$\bar{h}(x, \alpha) = -P_{AX}X\nabla f_\alpha(x),$$

o que implica

$$\bar{h}(x, \alpha) = -X\nabla f_\alpha(x) + XA^ty.$$

Assim,

$$s(x, \alpha) = \frac{x^{-1}}{\alpha}(e + X\nabla f_\alpha(x) - XA^ty).$$

Fazendo $\nabla f_\alpha(x) = \alpha c - x^{-1}$, temos

$$s(x, \alpha) = \frac{x^{-1}}{\alpha}(e + X(\alpha c - x^{-1}) - XA^t y) = c - A^t \frac{y}{\alpha}.$$

Portanto, $A^t \frac{y}{\alpha} + s = c$.

ii) Substituindo $s(x, \alpha)$ em $\delta(x, s, \alpha)$, temos

$$\begin{aligned} \delta(x, s(x, \alpha), \alpha) &= \|\alpha x s(x, \alpha) - e\| \\ &= \|\alpha x \left(\frac{x^{-1}}{\alpha}(e - \bar{h}(x, \alpha)) \right) - e\| \\ &= \|\bar{h}(x, \alpha)\| \\ &= \delta(x, s). \end{aligned}$$

□

A eficiência do passo de Newton primal.

A seguir, um resultado que é conhecido como um dos mais importantes em métodos de ponto interior, é a chamada propriedade de convergência quadrática.

Proposição 3.1.5 *Sejam $x \in \overset{\circ}{\mathcal{X}}$, $\alpha > 0$ e $x^+ = x + h(x, \alpha)$. Se $\delta(x, \alpha) = \delta < 1$, então $x^+ \in \overset{\circ}{\mathcal{X}}$ e*

$$\delta(x^+, \alpha) \leq \delta^2.$$

Demonstração. Sejam $x \in \overset{\circ}{\mathcal{X}}$ e $\alpha > 0$. Vamos primeiro provar que

$$x^+ = x + h(x, \alpha) \in \overset{\circ}{\mathcal{X}}.$$

Como $x > 0$ e $\|\bar{h}(x, \alpha)\| < 1$, segue imediato que

$$x^+ = x(e + \bar{h}(x, \alpha)) > 0.$$

Agora, $Ax^+ = Ax + Ah(x, \alpha)$, como $h(x, \alpha) \in \mathcal{N}(A)$, então segue que $Ax^+ = b$. Já vimos na proposição 3.1.4 que podemos construir uma folga dual viável

$$s = s(x, \alpha) = \frac{x^{-1}}{\alpha}(e - \bar{h}(x, \alpha)).$$

Usando a proposição 3.1.3, temos

$$\begin{aligned} \delta(x^+, \alpha) &\leq \|\alpha x^+ s - e\| \\ &= \|\alpha x(e + \bar{h}(x, \alpha))\frac{x^{-1}}{\alpha}(e - \bar{h}(x, \alpha))\| \\ &= \|\bar{h}^2(x, \alpha)\| \\ &\leq \|\bar{h}(x, \alpha)\|^2 \\ &= \delta^2. \end{aligned}$$

A última desigualdade vem do fato que

$$\|\bar{h}^2(x, \alpha)\|^2 = \sum_{i=1}^n \bar{h}_i^4(x, \alpha) \leq \left(\sum_{i=1}^n h_i^2(x, \alpha)\right)^2 = \|\bar{h}(x, \alpha)\|^4.$$

□

Direção de Newton primal-dual.

Sejam $(x, s) \in \overset{\circ}{\mathcal{F}}$ e $\alpha > 0$. Queremos calcular a direção de Newton primal-dual $h = (h_x, h_s)$ tal que

$$x^+ = x + h_x \quad \text{e} \quad s^+ = s + h_s$$

satisfaçam as condições de otimalidade de $(Q)_\alpha$ para o problema $(P)_\alpha$:

$$(Q)_\alpha \quad \begin{aligned} xs &= \frac{e}{\alpha} \\ Ax &= b \\ A^t y + s &= c \\ s, x &\geq 0. \end{aligned}$$

Como já vimos, este sistema tem uma solução única $(x(\alpha), s(\alpha))$, para cada $\alpha > 0$, e define uma curva

$$\alpha > 0 \mapsto (x(\alpha), s(\alpha))$$

chamada de **trajetória central primal-dual**.

A direção (h_x, h_s) precisa ser viável, ou seja,

$$A(x + h_x) = b,$$

donde $Ah_x = 0$, e

$$A^t y^+ + s^+ = c, \text{ para algum } y^+ = y + h_y,$$

donde $A^t h_y = -h_s$. Isto é equivalente a dizer que $h_x \in \mathcal{N}(A)$ e $h_s \in \mathcal{R}(A^t)$.

Substituindo x^+ e s^+ em $(Q)_\alpha$, a solução exata deve satisfazer

$$Sh_x + Xh_s = \frac{e}{\alpha} - xs - h_x h_s$$

$$h_x \in \mathcal{N}(A)$$

$$h_s \in \mathcal{R}(A^t).$$

Estas equações são não-lineares. O passo de Newton resolve este sistema ignorando o termo não-linear:

$$Sh_x + Xh_s = \frac{e}{\alpha} - xs$$

$$h_x \in \mathcal{N}(A)$$

$$h_s \in \mathcal{R}(A^t).$$

(3.1.1)

A solução de (3.1.1) é obtida fazendo uma mudança de escala. Façamos

$$(x, s) = (D\bar{x}, D^{-1}\bar{s}), \text{ onde } D = X^{\frac{1}{2}}S^{-\frac{1}{2}}.$$

Então,

$$\begin{aligned}\bar{s}\bar{h}_x + \bar{x}\bar{h}_s &= \frac{e}{\alpha} - \bar{x}\bar{s} \\ \bar{h}_x &\in \mathcal{N}(AD) \\ \bar{h}_s &\in \mathcal{R}(DA^t).\end{aligned}$$

Como $\bar{x} > 0$, então podemos multiplicar a primeira equação por \bar{x}^{-1} ,

$$\bar{x}^{-1}\bar{s}\bar{h}_x + \bar{h}_s = \frac{\bar{x}^{-1}}{\alpha} - \bar{s},$$

mas $\bar{x} = D^{-1}x$ e $\bar{s} = Ds$, então $\bar{x}^{-1}\bar{s} = e$. Assim,

$$\begin{aligned}\bar{h}_x + \bar{h}_s &= \frac{\bar{x}^{-1}}{\alpha} - \bar{s} \\ \bar{h}_x &\in \mathcal{N}(AD) \\ \bar{h}_s &\in \mathcal{R}(DA^t).\end{aligned}\tag{3.1.2}$$

e a solução é a projeção ortogonal do vetor $\frac{\bar{x}^{-1}}{\alpha} - \bar{s}$ em $\mathcal{N}(AD)$ e seu complemento ortogonal, ou seja,

$$\bar{h}_x = P_{AD}\left(\frac{\bar{x}^{-1}}{\alpha} - \bar{s}\right) \quad \text{e} \quad \bar{h}_s = \tilde{P}_{AD}\left(\frac{\bar{x}^{-1}}{\alpha} - \bar{s}\right).$$

A direção de Newton no espaço original é dada por:

$$h_x = D\bar{h}_x \quad \text{e} \quad h_s = D^{-1}\bar{h}_s.$$

A eficiência do passo de Newton primal-dual.

O resultado a seguir é análogo à proposição 3.1.5, ou seja, numa vizinhança da trajetória central o método de Newton primal-dual é muito eficiente, mas antes vamos enunciar um resultado de Mizuno, que é importante para a demonstração de convergência quadrática..

Proposição 3.1.6 (MIZUNO[31]) *Se $h_x, h_s \in \mathbb{R}^n$ são tais que $h_x^t h_s \geq 0$, então*

$$\|h_x h_s\| \leq \frac{1}{\sqrt{8}} \|h_x + h_s\|^2.$$

Proposição 3.1.7 *Sejam $(x, s) \in \mathcal{F}$ e $\alpha > 0$ tais que $\delta(x, s, \alpha) = \delta < 1$, e sejam $x^+ = x + h_x$ e $s^+ = s + h_s$, onde h_x e h_s são soluções de (3.1.1). Então*

$$\delta(x^+, s^+, \alpha) \leq \frac{\delta^2}{\sqrt{8}(1 - \delta)}.$$

Demonstração. Considerando o sistema (3.1.2) temos

$$\bar{h}_x + \bar{h}_s = \frac{\bar{x}^{-1}}{\alpha} - \bar{s}.$$

Como $\bar{x} = \sqrt{xs} = \bar{s}$, então

$$\bar{h}_x + \bar{h}_s = \frac{(XS)^{-\frac{1}{2}}}{\alpha} (e - \alpha Xs),$$

donde

$$\|\bar{h}_x + \bar{h}_s\|^2 \leq \frac{1}{\alpha^2 \|xs\|_\infty} \|e - \alpha Xs\|^2.$$

Mas, $\|e - \alpha xs\| = \delta < 1$, então

$$1 - \alpha x_i s_i \leq \delta, \quad \forall i = 1, \dots, n$$

o que implica

$$x_i s_i \geq \frac{1 - \delta}{\alpha}, \quad \forall i = 1, \dots, n.$$

Assim,

$$\|\bar{h}_x + \bar{h}_s\|^2 \leq \frac{\alpha}{\alpha^2(1-\delta)} \|e - \alpha Xs\|^2 \leq \frac{\delta^2}{\alpha(1-\delta)}.$$

Agora, usando a definição de (h_x, h_s) , temos

$$\delta(x^+, s^+, \alpha) = \|\alpha x^+ s^+ - e\| = \|\alpha(x + h_x)(s + h_s) - e\| = \|\alpha h_x h_s\|.$$

Como $h_x \perp h_s$, então $h_x^t h_s = 0$. Assim, podemos usar a proposição anterior e donde vem que

$$\delta(x^+, s^+, \alpha) = \|\alpha h_x h_s\| = \|\alpha \bar{h}_x \bar{h}_s\| \leq \frac{\alpha}{\sqrt{8}} \|\bar{h}_x + \bar{h}_s\|^2.$$

Portanto,

$$\delta(x^+, s^+, \alpha) \leq \frac{\delta^2}{\sqrt{8}(1-\delta)}.$$

□

Então, temos que o método de Newton primal-dual também goza da propriedade de convergência quadrática, para pontos próximos da trajetória central, ou seja, para $\delta \leq 1 - \frac{1}{2\sqrt{2}}$.

O Algoritmo de Trajetória Central de Passos Curtos

Agora descrevemos o algoritmo de trajetória central primal de passos curtos, de GONZAGA[15], é o que utilizaremos no que vamos desenvolver. Este algoritmo gera uma seqüência de valores para o parâmetro de penalidade (α^k) e uma seqüência de pontos (x^k) . À seqüência (α^k) está associada uma seqüência de pontos desconhecidos $(x(\alpha^k))$ da trajetória central. Cada ponto x^k satisfaz o seguinte:

i) x^k está muito próximo de $x(\alpha^k)$, pois $\delta(x^k, \alpha^k) \leq \delta^2$.

ii) x^k está próximo de $x(\alpha^{k+1})$, pois $\delta(x^k, \alpha^k) \leq \delta$.

Algoritmo 3.1.1 Dados $x^\circ \in \overset{\circ}{\mathcal{X}}$ e $\alpha^\circ > 0$ tais que $\delta(x^\circ, \alpha^\circ) \leq \delta^2$.

$k = 0$

REPITA

Encontrar α^{k+1} tal que $\delta(x^k, \alpha^{k+1}) = \delta$

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

$k = k + 1$

ATÉ convergência.

Comentário: A variação de $\alpha^{k+1} - \alpha^k$ é a maior possível mantendo

$$\delta(x^k, \alpha^{k+1}) = \delta$$

e necessita para o seu cálculo do conhecimento explícito de $\bar{c}_p = P_{AX}Xc$ e $e_p = P_{AX}e$.

3.2 Centro Analítico

Consideremos o politopo

$$\mathcal{X} = \{x \in \mathbb{R}^n; Ax = b, x \geq 0\},$$

com $b \in \mathbb{R}^m$ e $A \in \mathbb{R}^{m \times n}$, $m < n$. Suporemos o conjunto de pontos interiores

$$\overset{\circ}{\mathcal{X}} = \{x \in \mathcal{X}; x > 0\}$$

não-vazio.

Ao politopo \mathcal{X} temos um problema de programação linear associado

$$\begin{aligned}
&\text{maximizar } b^t y \\
&\text{sujeito a:} \\
&\quad A^t y + s = 0 \\
&\quad s \geq 0
\end{aligned}$$

o qual chamaremos problema dual.

O problema de centralização será dado por

$$\begin{aligned}
&\text{minimizar } f_0(x), \\
&\quad x \in \overset{\circ}{\mathcal{X}}
\end{aligned}$$

onde f_0 é a função de penalidade com $\alpha = 0$, ou seja, $f_0(x) = -\sum_{i=1}^n \ln x_i = p(x)$. O centro analítico de \mathcal{X} é a única solução ótima do problema de centralização

$$x_0 = x(0) = \arg \min_{x \in \overset{\circ}{\mathcal{X}}} p(x).$$

Então, a condição de otimalidade é dada por

$$P_A(-x_0^{-1}) = 0.$$

Dado um ponto $x \in \overset{\circ}{\mathcal{X}}$, a **direção de centralização de Newton primal** é dada por

$$h(x) = x\bar{h}(x),$$

onde $\bar{h}(x) = P_{AX}e$ é a direção de centralização de Newton primal após mudança de escala para o problema de centralização, de modo que o ponto x é transformado em e .

A proximidade primal de um ponto $x \in \overset{\circ}{\mathcal{X}}$ ao centro analítico x_0 é dada por

$$\delta(x) = \|\bar{h}(x)\| = \|P_{AX}e\|.$$

Considere $0 < \delta < 0.5$. Um ponto $x \in \overset{\circ}{\mathcal{X}}$ será chamado de **centro analítico δ -aproximado de \mathcal{X}** se $\delta(x) \leq \delta$ e δ será chamado de parâmetro de centralização.

A condição de otimalidade para o centro analítico é equivalente a

$$-x^{-1} \in \mathcal{R}(A^t)$$

que é equivalente a

$$-x^{-1} = A^t y,$$

para algum $y \in \mathbb{R}^m$. Fazendo $x^{-1} = s$, teremos as condições de otimalidade primal-dual:

$$\begin{aligned} x s &= e \\ A^t y + s &= 0 \\ A x &= b \\ x, s &\geq 0 \end{aligned} \tag{3.2.1}$$

A primeira equação de (3.2.1) fornece um critério simples de proximidade

$$\delta(x, s) = \|x s - e\|$$

a qual chamaremos de **proximidade primal-dual**.

Agora, destacamos algumas propriedades que são parecidas com as propriedades de pontos centrais apresentadas na seção anterior. As demonstrações são análogas.

Proposição 3.2.1 *Seja $x \in \overset{\circ}{\mathcal{X}}$. Então,*

$$\delta(x) = \min\{\|X s - e\|; A^t y + s = 0, y \in \mathbb{R}^m\}.$$

Proposição 3.2.2 *Seja $x \in \overset{\circ}{\mathcal{X}}$ tal que $\delta(x) < 1$. Defina*

$$s(x) = X^{-1}(e - \bar{h}(x)).$$

- i) $s(x)$ é uma folga dual viável para o problema de centralização.*
- ii) $\delta(x, s(x)) = \delta(x)$.*

3.3 Procedimentos de centralização

Procedimento de Newton primal

Dado $x \in \overset{\circ}{\mathcal{X}}$ satisfazendo

$$\delta(x) = \delta < 1,$$

podemos aplicar o método de Newton primal para gerar um novo ponto

$$x^+ = x + x\bar{h}(x),$$

onde $\bar{h}(x) = P_{AX}e$ é a direção de centralização de Newton primal.

Proposição 3.3.1 (convergência quadrática) *Seja $x \in \overset{\circ}{\mathcal{X}}$ e $x^+ = x + x\bar{h}(x)$.*

Se $\delta(x) = \delta < 1$, então $\delta(x^+) \leq \delta^2$.

Como consequência desta proposição faremos a seguinte definição: Se $x \in \overset{\circ}{\mathcal{X}}$ é tal que $\delta(x) = \eta < 1$, então diremos que x está na região de convergência quadrática do método de Newton primal.

Procedimento de Newton primal-dual.

Dado (x, s) primal-dual interior viável, queremos calcular a direção de Newton primal-dual $h = (h_x, h_s)$ para gerar um novo ponto

$$(x^+, s^+) = (x + h_x, s + h_s)$$

até obter uma solução do sistema (3.2.1).

Como já vimos na seção 3.1, fazendo a substituição obtemos o sistema

$$\begin{aligned} Sh_x + Xh_s &= e - xs \\ h_x &\in \mathcal{N}(A) \\ h_s &\in \mathcal{R}(A^t), \end{aligned} \tag{3.3.1}$$

fazendo a mudança de escala

$$x = D\bar{x} \quad \text{e} \quad s = D^{-1}\bar{s}$$

onde $D = X^{\frac{1}{2}}S^{-\frac{1}{2}}$, obtemos o sistema

$$\begin{aligned} \bar{h}_x + \bar{h}_s &= \bar{x}^{-1} - \bar{s} \\ \bar{h}_x &\in \mathcal{N}(AD) \\ \bar{h}_s &\in \mathcal{R}(DA^t) \end{aligned} \tag{3.3.2}$$

Proposição 3.3.2 *Seja (x, s) primal-dual viável tal que $\delta(x, s) = \delta < 1$. Se $x^+ = x + h_x$ e $s^+ = s + h_s$, onde (h_x, h_s) é solução de (3.3.1), então*

$$\delta(x^+, s^+) \leq \frac{\delta^2}{\sqrt{8}(1 - \delta)}.$$

3.4 Variação da Função Barreira ao longo da Trajetória Central

Nesta seção vamos considerar o problema de programação linear em formato primal (P) , como dado na seção 2.3,

$$(P) \quad \begin{array}{ll} \text{minimizar} & c^t x \\ \text{sujeito a:} & \\ & Ax = b \\ & x \geq 0, \end{array}$$

e a função penalizada associada a (P) ,

$$f_\alpha(x) = \alpha c^t x + p(x),$$

onde $p(x) = -\sum_{i=1}^n \ln x_i$ é a função barreira de (P) .

Vamos estudar algumas propriedades de pontos sobre a trajetória central, definida por

$$\alpha \geq 0 \mapsto x(\alpha) \text{ tal que } P_A \nabla f_\alpha(x(\alpha)) = 0.$$

Iniciamos por propriedades das medidas de proximidade. Dado um ponto $z \in \mathbb{R}_{++}^n$, definimos a norma relativa a z por

$$x \in \mathbb{R}^n \mapsto \|x\|_z = \|Z^{-1}x\|,$$

então podemos associar uma norma a cada ponto do ortante positivo. Fixando-se um ponto $z \in \mathbb{R}_{++}^n$, observemos que

$$\|x - z\|_z = \|Z^{-1}x - e\|,$$

para qualquer ponto $x \in \mathbb{R}_{++}^n$. Isto é, $\|x - z\|_z$ corresponde à distância Euclidiana entre esses dois pontos após uma mudança de escala que leva z ao ponto e .

Proposição 3.4.1 *Considere o ponto central $x(\alpha)$ para $\alpha > 0$ e $x \in \overset{\circ}{\mathcal{X}}$. Então, $\delta(x, \alpha) \leq \|x - x(\alpha)\|_{x(\alpha)}$.*

Demonstração. Consideremos $x(\alpha)$ o ponto central para $\alpha > 0$ e $x \in \overset{\circ}{\mathcal{X}}$. Da proposição 3.1.3, temos

$$\delta(x, \alpha) = \min\{\|\alpha x s - e\|; s \text{ é folga dual}\}.$$

Em particular, $s = (\alpha x(\alpha))^{-1}$ é folga dual, pela proposição 3.1.2, então

$$\delta(x, \alpha) \leq \left\| \frac{x}{x(\alpha)} - e \right\| = \|x - x(\alpha)\|_{x(\alpha)}.$$

□

De ROOS *et.al.*[40], para $\delta(x, \alpha) < \frac{\sqrt{2}}{2}$,

$$\left\| \frac{x}{x(\alpha)} - e \right\| \leq \sqrt{1 - \sqrt{1 - \sqrt{1 - 2\delta(x, \alpha)^2}}} \quad (3.4.1)$$

Da proposição acima e do resultado de ROOS, temos:

i) Se $\delta(x, \alpha) = 0.1$ então $0.1 \leq \|x - x(\alpha)\|_{x(\alpha)} \leq 0.1004$.

ii) Se $\delta(x, \alpha) = 0.4$ então $0.4 \leq \|x - x(\alpha)\|_{x(\alpha)} \leq 0.42$.

É fácil verificar que para $\delta(x, \alpha) \leq 0.4$,

$$\delta(x, \alpha) \leq \|x - x(\alpha)\|_{x(\alpha)} \leq 1.05\delta(x, \alpha). \quad (3.4.2)$$

A trajetória central inicia no centro analítico $x(0)$, e afasta-se dele quando α cresce. Nos próximos resultados, bastante técnicos, mostramos que após afastar-se do centro analítico, a trajetória central não volta a aproximar-se muito dele, o mesmo ocorrendo com pontos aproximadamente centrais.

A proximidade no centro analítico é dada por

$$x \in \overset{\circ}{\mathcal{X}} \mapsto \delta(x) = \delta(x, 0) = \|P_{AX}e\|.$$

Proposição 3.4.2 *Considere $x^1 = x(\alpha^1)$, $\alpha^1 > 0$, com $\delta(x^1) \geq 0.25$. Então, para todo $\alpha > \alpha^1$, $\delta(x(\alpha)) > 0.159$.*

Demonstração. Suponhamos que o centro analítico de \mathcal{X} é $x(0) = e$. Suponhamos também que $x^1 = x(\alpha^1)$ e $\delta(x^1) \geq 0.25$. Usaremos os seguintes fatos com $h = x - e$:

i) $0.25 \leq \|h\| \leq 0.26$

ii) $p(x^1) \geq \frac{\|h\|^2}{2} - \frac{\|h\|^3}{3}$, de GONZAGA[17].

iii) $\alpha > 0 \mapsto p(x(\alpha))$ é crescente, pela proposição 3.4.4 adiante.

De i) e ii) concluímos que $p(x^1) \geq 0.026$ e de iii) temos que para todo $\alpha > \alpha^1$,

$$p(x(\alpha)) \geq p(x(\alpha^1)) \geq 0.026.$$

De GONZAGA[15], segue que

$$p(x(\alpha)) \leq \frac{\delta(x(\alpha))^2}{1 - \delta(x(\alpha))^2},$$

para $\delta(x(\alpha)) < 1$. Desta desigualdade temos facilmente que $\delta(x(\alpha)) \geq 0.159$. \square

Proposição 3.4.3 *Considere os pontos $x(\alpha)$, para $\alpha > 0$ dado, e x tais que $\delta(x, \alpha) \leq \delta$. Então:*

$$\text{Se } \delta \leq 0.1 \text{ implica que } |\delta(x) - \delta(x(\alpha))| \leq 0.302\delta(x(\alpha)) + \delta.$$

$$\text{Se } \delta \leq 0.01 \text{ implica que } |\delta(x) - \delta(x(\alpha))| \leq 0.0301\delta(x(\alpha)) + \delta$$

Demonstração. Suponhamos, sem perda de generalidade, que $x(\alpha) = e$, para $\alpha > 0$, e seja x tal que $\delta(x, \alpha) \leq \delta$. Queremos encontrar um limite para $\|P_{AX}e - P_Ae\|$.

Definamos $e_p = P_Ae$ e $c_p = P_Ac$. Como $\delta(e, \alpha) = 0$, temos $P_A(\alpha c - e) = 0$. Dado um vetor $v \in \mathcal{R}(A^t)$, temos $P_Av = 0$ e também $P_{AX}Xv = 0$, para todo $x > 0$. Portanto, $P_{AX}X(\alpha c - e) = 0$, donde

$$\alpha P_{AX}Xc = P_{AX}Xe \tag{3.4.3}$$

Mas $\delta(x, \alpha) = \|P_{AX}(\alpha Xc - e)\| \leq \delta$, e portanto,

$$\alpha P_{AX}Xc = P_{AX}e + \rho, \quad (3.4.4)$$

onde $\|\rho\| \leq \delta$. De (3.4.3) e (3.4.4), temos

$$P_{AX}Xe = P_{AX}e + \rho. \quad (3.4.5)$$

Vamos estabelecer limites para $d = P_{AX}Xe$. Como $e = e_p + \tilde{e}_p$, $\tilde{e}_p \in \mathcal{R}(A^t)$,

$$d = P_{AX}Xe = P_{AX}Xe_p,$$

pois $P_{AX}X\tilde{e}_p = 0$, como vimos acima. Logo, por definição de projeção,

$$Xe_p = d + z,$$

onde $\|z\| = \min\{\|Xe_p - w\|; AXw = 0\}$. Como $AXX^{-1}e_p = 0$, temos

$$\begin{aligned} \|z\| &\leq \|Xe_p - X^{-1}e_p\| \\ &\leq \|x - x^{-1}\|_{\infty} \|e_p\|. \end{aligned}$$

Obtemos então, usando (3.4.5),

$$\begin{aligned} P_{AX}e - e_p &= d - \rho - e_p \\ &= Xe_p - z - \rho - e_p, \\ &= (X - I)e_p - z - \rho \end{aligned}$$

donde

$$\begin{aligned} \|P_{AX}e - e_p\| &\leq \|x - e\|_{\infty} \|e_p\| + \|z\| + \|\rho\| \\ &\leq \|x - e\|_{\infty} \|e_p\| + \|x - x^{-1}\|_{\infty} \|e_p\| + \delta. \\ &\leq (\|x - e\|_{\infty} + \|x - x^{-1}\|_{\infty}) \|e_p\| + \delta \end{aligned}$$

Usando (3.4.1) com $\delta = 0.1$, temos

$$\|x - e\| \leq 0.1004,$$

donde $\|x - e\|_\infty + \|x - x^{-1}\|_\infty \leq 0.302$, com cálculos imediatos. Portanto, como $\|e_p\| = \delta(e)$, temos

$$\|P_{AX}e - e_p\| \leq 0.302\delta(e) + \delta.$$

Repetindo as contas para $\delta = 0.01$, completamos a demonstração. \square

Propriedades Diferenciais da Trajetória Central

Proposição 3.4.4 *Seja $x = x(\alpha)$ o ponto central associado a $\alpha > 0$ e considere as funções*

$$\alpha > 0 \mapsto x(\alpha)$$

e

$$\alpha > 0 \mapsto p(x(\alpha)).$$

Então,

$$\text{i) } \frac{dx(\alpha)}{d\alpha} = -\frac{1}{\alpha}XP_{AX}e$$

$$\text{ii) } \frac{dp(x(\alpha))}{d\alpha} = \frac{1}{\alpha}\|P_{AX}e\|^2 = \frac{1}{\alpha}\delta(x(\alpha))^2.$$

Demonstração. Suponhamos $x(\alpha) = e$. Como

$$\nabla f_\alpha(x) = \alpha c - x^{-1} \quad \text{e} \quad \nabla^2 f_\alpha(x) = X^{-2},$$

então projetando $\nabla f_\alpha(x)$ em $\mathcal{N}(A)$, temos

$$P_A \nabla f_\alpha(x) = \alpha P_A c - P_A x^{-1},$$

derivando em relação a α ,

$$\frac{dP_A \nabla f_\alpha(x)}{d\alpha} = P_A c + P_A X^{-2} \frac{dx}{d\alpha},$$

como $x(\alpha) = e$, então

$$\frac{dP_A \nabla f_\alpha(e)}{d\alpha} = P_A c + P_A \frac{dx}{d\alpha},$$

donde

$$0 = P_A c + P_A \frac{dx}{d\alpha},$$

pois $\frac{dx}{d\alpha} \in \mathcal{N}(A)$ e $P_A \nabla f_\alpha(x(\alpha)) = 0, \forall \alpha > 0$. Assim, $\frac{dx}{d\alpha} = -P_A c$ em $x(\alpha) = e$ e como $P_A \nabla f_\alpha(e) = \alpha c_p - e_p = 0$, então

$$\frac{dx}{d\alpha} = -\frac{e_p}{\alpha}.$$

Em geral, com $X = \text{diag}(x(\alpha))$ fixo e a mudança de escala $x = X\bar{x}$ e $x(\alpha) = Xe$, então

$$\frac{dx}{d\alpha} = X \frac{d\bar{x}}{d\alpha} = -X P_{AX} X c = -\frac{X}{\alpha} P_{AX} e,$$

Agora, vamos derivar $p(x(\alpha)) = -\sum_{i=1}^n \ln x_i(\alpha)$,

$$\begin{aligned} \frac{dp(x(\alpha))}{d\alpha} &= \frac{dp(x(\alpha))}{dx} \frac{dx}{d\alpha} \\ &= (-x^{-1})^t \left(-\frac{X}{\alpha} P_{AX} e\right) \\ &= \frac{1}{\alpha} e^t P_{AX} e \\ &= \frac{1}{\alpha} \|P_{AX} e\|^2. \end{aligned}$$

□

Variação de $p(\cdot)$ ao longo da Trajetória Central

Vamos estudar a variação de $p(\cdot)$ ao longo da trajetória devido à variação de α correspondente a um algoritmo de passos curtos. Essa variação aparece na literatura como $\alpha^+ = (1 + \frac{\nu}{\sqrt{n}})\alpha$, onde usualmente $\nu = 0.2$. Mostraremos que essa estimativa é geralmente muito conservativa.

Proposição 3.4.5 *Considere o algoritmo de passos curtos definido na seção 3.1. Sejam x e $\alpha > 0$ tais que $\delta(x, \alpha) \leq \delta < \beta < 1$ e suponha que $\delta(x) = \|P_{AX} e\| > \delta$. Então*

$$\alpha^+ \geq \alpha \left(1 + \frac{\nu}{\delta(x)}\right),$$

onde $\nu \geq \frac{\beta - \delta}{2}$.

Demonstração. Sem perda de generalidade, analisamos o passo curto a partir de e . Sejam dados $\|h(e, \alpha)\| = \|\alpha c_p - e_p\|$, $\alpha^+ = \alpha + \Delta\alpha$ com $\Delta\alpha = \frac{\nu\alpha}{\|e_p\|}$ e $\|e_p\| > \delta$. Vamos calcular o valor de ν tal que $\|\alpha^+ c_p - e_p\| = \beta < 1$. Então,

$$\begin{aligned}\beta &= \|(\alpha + \frac{\nu\alpha}{\|e_p\|})c_p - e_p\| \\ &= \|\alpha c_p - e_p + \frac{\nu\alpha c_p}{\|e_p\|}\| \\ &\leq \delta + \nu\alpha \frac{\|c_p\|}{\|e_p\|} \\ &\leq \delta + \nu \frac{(\delta + \|e_p\|)}{\|e_p\|} \\ &\leq \delta + 2\nu\end{aligned}$$

Portanto, $\nu \geq \frac{\beta - \delta}{2}$. □

Podemos, finalmente, estudar a variação da função barreira devido a um passo curto ao longo da trajetória central. Estudamos a variação de $p(x(\cdot))$ entre dois pontos α e α^+ tais que $\alpha^+ = (1 + \frac{\nu}{\delta(x(\alpha))})\alpha$.

Proposição 3.4.6 *Considere $\alpha^\circ > \alpha^1$ tal que $\delta(x(\alpha^1)) \geq 0.25$ e x tal que $\delta(x, \alpha^\circ) \leq 0.01$. Seja α^+ o valor de α gerado por um algoritmo de passos curtos a partir de x , com $\delta(x, \alpha^+) = 0.1$. Então,*

$$p(x(\alpha^+)) \geq p(x(\alpha^\circ)) + 0.007\delta(x(\alpha^\circ)).$$

Demonstração. Seja $\alpha^\circ > \alpha^1$ tal que $\delta(x(\alpha^1)) \geq 0.25$. Então, pela proposição 3.4.2, $\delta(x(\alpha)) \geq 0.159$ para todo $\alpha \geq \alpha^\circ$. Suponhamos sem perda de generalidade que $x = e$. Então temos, $\delta(x, \alpha) \leq 0.1$ para todo $\alpha \in [\alpha^\circ, \alpha^+]$, por definição de α . Pela proposição 3.4.3, com $\delta(x, \alpha^\circ) \leq 0.01$, temos

$$\delta(x) \geq \delta(x(\alpha^\circ))(1 - 0.0301) - 0.01.$$

Como $\delta(x(\alpha^\circ)) \geq 0.159$, temos $0.01 < 0.1\delta(x(\alpha^\circ))$ e $\delta(x) \geq 0.87\delta(x(\alpha^\circ))$. Por substituição direta de $\delta(x(\alpha^\circ)) \geq 0.159$, obtemos também $\delta(x) \geq 0.144$. Pela mesma proposição, para $\alpha \in [\alpha^\circ, \alpha^+]$, $\delta(x, \alpha) \leq 0.1$,

$$\delta(x(\alpha)) \geq \frac{\delta(x) - 0.1}{1.302}.$$

Se $\delta(x) \leq 0.3$ então trivialmente $\delta(x(\alpha)) \geq \frac{\delta(x)}{2}$, pois $\delta(x(\alpha)) \geq 0.159$. Se $\delta(x) > 0.3$ então $\delta(x) - 0.1 \geq \frac{2}{3}\delta(x)$ e obtemos $\delta(x(\alpha)) \geq \frac{\delta(x)}{2}$. Dessas relações concluímos que para todo $\alpha \in [\alpha^\circ, \alpha^+]$,

$$\delta(x(\alpha)) \geq \frac{\delta(x)}{2} \quad (3.4.6)$$

Então, usando o Teorema Fundamental do Cálculo, a proposição 3.4.4 e (3.4.6), temos

$$\begin{aligned} p(x(\alpha^+)) &= p(x(\alpha^\circ)) + \int_{\alpha^\circ}^{\alpha^+} \frac{dp(x(\alpha))}{d\alpha} d\alpha \\ &= p(x(\alpha^\circ)) + \int_{\alpha^\circ}^{\alpha^+} \frac{\delta(x(\alpha))^2}{\alpha} d\alpha \\ &\geq p(x(\alpha^\circ)) + \frac{\delta(x)^2}{4\alpha^+} (\alpha^+ - \alpha^\circ) \end{aligned}$$

Mas $\alpha^+ = (1 + \frac{\nu}{\delta(x)})\alpha^\circ$, com $\nu \geq \frac{0.1-0.01}{2} > 0.04$, entao

$$\frac{\alpha^+ - \alpha^\circ}{\alpha^+} = \frac{\frac{\nu}{\delta(x)}\alpha^\circ}{(1 + \frac{\nu}{\delta(x)})\alpha^\circ} \geq \frac{1}{\delta(x)} \frac{0.04}{(1 + \frac{0.04}{0.144})} \geq \frac{0.0344}{\delta(x)}$$

Finalmente,

$$\begin{aligned} p(x(\alpha^+)) &\geq p(x(\alpha^\circ)) + \frac{0.0344}{4}\delta(x) \\ &\geq p(x(\alpha^\circ)) + 0.0086\delta(x) \quad . \\ &\geq p(x(\alpha^\circ)) + 0.007\delta(x(\alpha^\circ)) \end{aligned}$$

□

Capítulo 4

O Algoritmo de Goffin-Vial para Calcular o Centro Analítico de um Politopo

Neste capítulo apresentaremos o procedimento de GOFFIN-VIAL[9] para calcular o centro analítico de um politopo depois de adicionado um corte muito profundo e que, essencialmente, consiste no seguinte. O problema é considerado num formato primal, o dual do espaço onde está a aproximação poliedral. Com o auxílio da direção de MITCHELL e TODD[30] a viabilidade é recuperada num único passo. A centralização é obtida de passos de Newton com busca com respeito a um problema primal, que está relacionado ao problema primal canônico de KARMAKAR[20], e que chamaremos de algoritmo de centralização projetivo primal.

4.1 Os Problemas Dual e Primal

Considere o politopo dado no formato dual

$$\Omega = \{y \in \mathbb{R}^m; A^t y \leq c\}$$

onde $c \in \mathbb{R}^n$ e A é uma matriz $m \times n$ de posto máximo $m < n$. Suporemos que $\overset{\circ}{\Omega}$ é não-vazio. Como A tem posto máximo, então podemos representar o politopo Ω em termo das folgas

$$\mathcal{S}_\Omega = \{s \in \mathbb{R}^n; A^t y + s = c, s \geq 0\}.$$

O centro analítico dual é o único ponto que minimiza a função barreira dual que denotaremos por

$$s \in \overset{\circ}{\mathcal{S}}_\Omega \mapsto \varphi_D(s) = - \sum_{i=1}^n \ln s_i.$$

As condições de otimalidade de primeira ordem para o problema de centralização dual são:

$$\begin{aligned} xs &= e \\ A^t y + s &= c \\ Ax &= 0 \\ x, s &> 0. \end{aligned} \tag{4.1.1}$$

Do teorema de dualidade, temos um problema de programação linear relacionado ao politopo $\Omega = \{y \in \mathbb{R}^m; A^t y \leq c\}$

minimizar $c^t x$
sujeito a:

$$\begin{aligned} Ax &= 0 \\ x &\geq 0, \end{aligned}$$

o qual chamaremos de problema primal. Como Ω é não-vazio, então o valor mínimo deste problema é 0. Mas Ω também é limitado e como tem um conjunto interior não-vazio, então o conjunto dos pontos interiores primais viáveis

$$\bar{\mathcal{X}}_{\Omega} = \{x \in \mathbb{R}^n; Ax = 0, x \geq 0\}$$

é não-vazio, ver VIAL[48], e $x = 0$ é uma solução ótima primal.

Assim, podemos definir uma função potencial primal para $\bar{\mathcal{X}}_{\Omega}$ como,

$$P(x) = n \ln c^t x - \sum_{i=1}^n \ln x_i, \forall x \in \overset{\circ}{\bar{\mathcal{X}}}_{\Omega}.$$

Agora consideremos o problema

$$\text{minimizar } P(x)$$

sujeito a:

$$Ax = 0$$

$$x > 0.$$

Como $P(\cdot)$ é uma função homogênea de grau zero, ou seja, para qualquer $x > 0$, $\lambda > 0$, $P(\lambda x) = P(x)$, então podemos fixar $c^t x = n$ como uma restrição de normalização. Denotando como polítopo primal o polítopo $\bar{\mathcal{X}}_{\Omega}$ acrescido da restrição $c^t x = n$,

$$\mathcal{X}_{\Omega} = \{x \in \mathbb{R}^n; Ax = 0, c^t x = n, x \geq 0\},$$

teremos que dado qualquer ponto $x > 0$ tal que $Ax = 0$, mas $c^t x \neq n$ então, o ponto $\frac{n}{c^t x} x \in \overset{\circ}{\mathcal{X}}_{\Omega}$ e tem o mesmo valor potencial. Então vamos definir o centro analítico primal como a solução ótima do problema de centralização primal

minimizar $\varphi_P(x)$

sujeito a:

$$Ax = 0$$

$$c^t x = n$$

$$x > 0,$$

onde $\varphi_P(x) = -\sum_{i=1}^n \ln x_i$. Em VIAL[48], podemos ver que as condições de otimalidade para o problema de centralização primal são iguais a (4.1.1).

Seja (\hat{x}, \hat{s}) um par primal-dual interior tal que $A\hat{x} = 0$ e $\hat{s} = c - A^t\hat{y}$, para algum \hat{y} . O par (\hat{x}, \hat{s}) é dito um centro analítico δ -aproximado se

$$\|e - \hat{x}\hat{s}\| \leq \delta$$

$$A^t\hat{y} + \hat{s} = c$$

$$A\hat{x} = 0$$

$$\hat{x}, \hat{s} > 0.$$

Observemos que nesta definição não estamos exigindo que $c^t\hat{x} = n$.

Proposição 4.1.1 (desigualdade de dualidade, VIAL[48]) *Sejam $x \in \overset{\circ}{\mathcal{X}}_\Omega$ e $s \in \overset{\circ}{\mathcal{S}}_\Omega$. Então*

$$\varphi_P(x) + \varphi_D(s) \geq 0,$$

com igualdade se, e só se, $xs = e$, ou seja, o par (x, s) é centro analítico.

4.2 O Algoritmo de Centralização Projetivo Primal

Começamos apresentando alguns resultados já conhecidos sobre o método de Newton primal para o problema de centralização primal reformulado

$$\begin{aligned} &\text{minimizar} && \varphi_P(x) \\ &\text{sujeito a:} && \\ &&& (c A^t)^t x = n e_1 \\ &&& x > 0, \end{aligned}$$

onde $e_1 \in \mathbb{R}^{m+1}$.

Dado um ponto viável interior primal x , a direção de Newton primal em x , como já vimos na seção 3.2, é

$$h(x) = x \bar{h}(x),$$

onde $\bar{h}(x) = P_{(c A^t)^t X} e$. O vetor $\bar{h}(x)$ também pode ser dado pela formulação explícita

$$\bar{h}(x) = e - \rho x s,$$

com

$$y = (AX^2A^t)^{-1}AX^2c,$$

$$s = c - A^t y,$$

$$\rho = \frac{c^t x}{\|xs\|^2}.$$

No que vamos desenvolver a seguir, usaremos esta definição para $\bar{h}(x)$, ou seja,

$$\bar{h}(x) = e - \rho(x)xs(x)$$

com

$$(\rho(x), s(x)) = \arg \min_{\rho, s} \{\|e - \rho x s\|; s = c - A^t y\} \quad (4.2.1)$$

Para maiores detalhes, ver VIAL[48].

Já sabemos que a norma de $\bar{h}(x)$ mede a proximidade ao centro analítico.

Proposição 4.2.1 *Seja x tal que $\|\bar{h}(x)\| \leq \delta < 1$ e seja $(\rho(x), s(x))$ definido como em (4.2.1). Defina $\hat{x} = \rho x$. Então $\bar{h}(\hat{x}) = \bar{h}(x)$. Além disso, se $\hat{s} = s(\hat{x})$, onde $s(\hat{x})$ é dado como em (4.2.1), então $\hat{s} = s(x)$, $\bar{h}(\hat{x}) = e - \hat{x}\hat{s} = \bar{h}(x)$ e (\hat{x}, \hat{s}) é um centro analítico δ -aproximado.*

Proposição 4.2.2 (convergência quadrática) *Seja $\|\bar{h}(x)\| \leq \delta < 1$ e seja s definido como em (4.2.1). Então s é viável interior dual. Além disso, $x^+ = x + x\bar{h}(x) > 0$ é viável interior primal e*

$$\|\bar{h}(x^+)\| \leq \|\bar{h}(x)\|^2.$$

Demonstração. (VIAL[48]). □

Corolário 4.2.1 *Seja $\|\bar{h}(x)\| \leq \delta < 1$ e $c^t x = n$. Seja $\hat{\varphi}_P$ o valor mínimo de $\varphi_P(x)$, $\forall x \in \overset{\circ}{\mathcal{X}}_\Omega$. Então,*

$$\varphi_P(x) - \hat{\varphi}_P \leq \frac{\delta^2}{1 - \delta^2}.$$

Proposição 4.2.3 *Seja $\|\bar{h}(x)\| > \eta > 0$. Seja $\alpha = \frac{1}{1 + \eta}$. O passo de Newton com busca*

$$x(\alpha) = x + \alpha x \bar{h}(x) > 0$$

é viável primal e

$$\varphi_P(x(\alpha)) \leq \varphi_P(x) - (\eta - \ln(1 + \eta)).$$

Agora, vamos apresentar um procedimento para calcular um centro analítico δ -aproximado de um politopo.

O Algoritmo de Centralização Projetivo Primal

Dados dois parâmetros δ e η , tais que $0 < \delta \leq \eta < 1$, onde δ é o parâmetro de centralização e η é o parâmetro que garante a convergência quadrática do método de Newton. Dado o ponto inicial $\bar{x} > 0$ tal que $A\bar{x} = 0$.

Inicialização:

Faça $x^1 = \bar{x}$. Calcule $\|\bar{h}(x^1)\|$.

Faça $k = 1$.

Enquanto $\|\bar{h}(x^k)\| > \delta$ faça

Se $\|\bar{h}(x^k)\| > \eta$, então $x^{k+1} = x^k + \alpha x^k \bar{h}(x^k)$, onde $\alpha = \frac{1}{1 + \eta}$

Senão

$x^{k+1} = x^k + x^k \bar{h}(x^k)$

$k = k + 1$

Fim

O problema de calcular o centro analítico aproximado de um politopo usando o método de Newton com busca foi resolvido por VAIDYA[47]. A análise de complexidade é muito simples e podemos encontrá-la em GONZAGA[15].

Proposição 4.2.4 *Sejam $\bar{s} \in \mathring{S}_\Omega$, $0 < \delta \leq \eta < 1$ e $\sigma_1 = \eta - \ln(1 + \eta)$. A seqüência*

(x^k) gerada pelo algoritmo projetivo primal iniciado em \bar{x} tem as seguintes propriedades:

i) x^k é centro analítico η -aproximado, ou seja, $\|\bar{h}(x^k)\| \leq \eta$ para

$$k \geq k_1 \geq \frac{\varphi_P(\bar{x}) + \varphi_D(\bar{s})}{\sigma_1}$$

ii) $\|\bar{h}(x^k)\| \leq \delta$ para $k \geq k_1 + \frac{1}{\log 2} \ln \left[\frac{\log \delta}{\log \eta} \right]$.

4.3 Adicionando um Corte ao Politopo Ω

Na seção anterior vimos como calcular o centro analítico δ -aproximado de um politopo $\Omega = \{y \in \mathbb{R}^m : A^t y \leq c\}$. Agora, sendo conhecido o centro analítico δ -aproximado de Ω , queremos calcular um novo centro analítico δ -aproximado do politopo obtido de Ω acrescido de um corte, para tal usaremos como ponto inicial o centro analítico antigo.

Seja \bar{x} centro analítico δ -aproximado de Ω . Então,

$$A^t \bar{y} + \bar{s} = c, \quad \bar{s} > 0$$

$$A \bar{x} = 0, \quad \bar{x} > 0$$

$$\|e - \bar{x} \bar{s}\| \leq \delta < 1$$

com $\bar{y} = (A \bar{X}^2 A^t)^{-1} A \bar{X}^2 c$. O ponto \bar{x} é gerado de uma iteração de Newton primal com a condição

$$\bar{x} = \frac{c^t \hat{x}}{\|\hat{x} \bar{s}\|^2} \hat{x},$$

onde \hat{x} é o primeiro iterado para o qual $\|\bar{h}(\hat{x})\| \leq \delta$.

Dado o corte

$$a^t y \leq a^t \bar{y} - \bar{\gamma} = K.$$

Suporemos que

$$\Omega \cap \{y \in \mathbb{R}^m; a^t y \leq K\}$$

tem conjunto interior não-vazio.

Após adicionar a nova restrição, denotaremos o novo politopo por

$$\Omega_K = \{y \in \mathbb{R}^m; \tilde{A}^t y \leq \tilde{c}\},$$

onde $\tilde{A} = (A \ a)$ e $\tilde{c} = \begin{pmatrix} c \\ K \end{pmatrix}$.

No que segue, usaremos o til para denotar os elementos do sistema aumentado.

Ao adicionar a nova restrição ao politopo, precisamos deslocar o ponto \bar{y} no conjunto viável Ω . O elipsóide de Dikin com a escala primal,

$$\{y; \|\bar{X} A^t (y - \bar{y})\| \leq 1 - \delta\},$$

define uma região de confiança em torno do ponto \bar{y} , para mais detalhes ver DIKIN[4] e VIAL[48]. Para recuperar a viabilidade, consideramos os pontos que estão ao longo de uma certa direção. Tomaremos a direção que maximiza a folga correspondente à nova restrição dentro da região de confiança definida pelo elipsóide de Dikin. Formalmente, o problema é

$$\begin{aligned} &\text{minimizar} && a^t \Delta y \\ &\text{sujeito a:} && \\ &&& \|\bar{X} A^t \Delta y\| \leq 1 - \delta, \end{aligned} \tag{4.3.1}$$

Proposição 4.3.1 A solução do problema (4.3.1) é a direção afim dual, usando uma escala primal,

$$\Delta y = -\frac{(1-\delta)}{r}(A\bar{X}^2A^t)^{-1}a,$$

onde $r = \sqrt{a^t(A\bar{X}^2A^t)^{-1}a}$

Um corte é dito pouco profundo se ele intercepta o elipsóide de Dikin, ou seja, se $\bar{\gamma} = \gamma r(1-\delta)$, com $-1 \leq \gamma \leq 1$.

Um passo

$$y = \bar{y} + \alpha \Delta y$$

com $-1 \leq \alpha \leq 1$ está no elipsóide de Dikin e, portanto, em Ω . Isto implica que

$$s(\alpha) = \bar{s} + \alpha \Delta s = \bar{s} - \alpha A^t \Delta y \geq 0,$$

onde

$$\Delta s = \frac{1-\delta}{r} A^t (A\bar{X}^2A^t)^{-1} a.$$

Seguindo este passo, a folga correspondente à restrição adicionada é

$$\begin{aligned} s_{n+1}(\alpha) &= K - a^t y = a^t \bar{y} - \bar{\gamma} - a^t (\bar{y} + \alpha \Delta y) \\ &= -\alpha a^t \Delta y - \gamma r(1-\delta) \\ &= \alpha r(1-\delta) - \gamma r(1-\delta) \\ &= r(1-\delta)(\alpha - \gamma). \end{aligned}$$

Assim, o novo ponto é

$$\tilde{s}(\alpha) = \begin{pmatrix} \bar{s} - \alpha A^t \Delta y \\ r(1-\delta)(\alpha - \gamma) \end{pmatrix}.$$

Portanto, se $-1 < \alpha < 1$ e $\alpha > \gamma$, temos a viabilidade dual estrita para a variável dual atualizada.

A atualização da variável primal é feita seguindo a direção de MITCHELL e TODD[30], que é obtida da seguinte forma. O vetor $\begin{pmatrix} \bar{x} \\ 0 \end{pmatrix}$ é primal viável, pois

$(A \ a) \begin{pmatrix} \bar{x} \\ 0 \end{pmatrix} = 0$, mas não é um ponto interior. Então, queremos uma direção

tal que partindo de $\begin{pmatrix} \bar{x} \\ 0 \end{pmatrix}$ obtemos um ponto estritamente positivo. Tomaremos uma direção que é proporcional à direção de máxima descida do problema

$$\begin{aligned} & \text{minimizar} && x_{n+1} \\ & \text{sujeito a:} && \\ & && Ax + x_{n+1}a = 0 \\ & && x \geq 0 \\ & && x_{n+1} \geq 0 \end{aligned} \tag{4.3.2}.$$

Proposição 4.3.2 *A direção de máxima descida do problema (4.3.2), $P_{(A\bar{X} \ a)} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ é proporcional ao vetor $\begin{pmatrix} -\bar{X}A^t(A\bar{X}^2A^t)^{-1}a \\ 1 \end{pmatrix}$.*

Demonstração. Ver MITCHELL e TODD[30] □

Assim, o ponto

$$\tilde{x}(\alpha) = \begin{pmatrix} \bar{x} \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} \Delta x \\ \frac{1-\delta}{r} \end{pmatrix},$$

onde $\Delta x = -\bar{X}^2 \Delta s = \frac{(1-\delta)}{r} \bar{X}^2 A^t (A\bar{X}^2 A^t)^{-1} a$, é tal que $\tilde{A}\tilde{x}(\alpha) = 0$ e $\tilde{x}(\alpha) > 0$.

Recuperação da centralidade

Proposição 4.3.3 *Se $\|\tilde{e} - \tilde{x}(\alpha)\tilde{s}(\alpha)\| \leq \delta^{\frac{1}{2^k}}$, então é preciso de no máximo k iterações do método de Newton primal para obter um centro analítico δ -aproximado para o politopo Ω_K .*

Seja

$$\tilde{e} - \tilde{x}(\alpha)\tilde{s}(\alpha) = \begin{pmatrix} e - (\bar{x} + \alpha\Delta x)(\bar{s} + \alpha\Delta s) \\ 1 - \alpha(\alpha - \gamma)(1 - \delta)^2 \end{pmatrix}.$$

Como

$$\begin{aligned} (\bar{x} + \alpha\Delta x)(\bar{s} + \alpha\Delta s) &= \bar{x}\bar{s} + \alpha(\bar{x}\Delta s + \bar{s}\Delta x) + \alpha^2\Delta x\Delta s \\ &= \bar{x}\bar{s} + \alpha(e - \bar{x}\bar{s})\bar{X}\Delta s - \alpha^2\bar{X}^2(\Delta s)^2. \end{aligned}$$

Então,

$$\begin{aligned} \|e - (\bar{x} + \alpha\Delta x)(\bar{s} + \alpha\Delta s)\| &= \|e - \bar{x}\bar{s} - \alpha(e - \bar{x}\bar{s})\bar{X}\Delta s - \alpha^2\bar{X}^2(\Delta s)^2\| \\ &\leq \|e - \bar{x}\bar{s}\| + \alpha\|e - \bar{x}\bar{s}\| \|\bar{X}\Delta s\| + \alpha^2\|\bar{X}^2(\Delta s)^2\| \\ &\leq \delta + \alpha\delta\|\bar{X}\Delta s\| + \alpha^2\|\bar{X}^2(\Delta s)^2\|. \end{aligned}$$

Mas,

$$\begin{aligned} \|\bar{X}\Delta s\|^2 &= (\Delta s)^t \bar{X}^2 \Delta s \\ &= \frac{(1-\delta)^2}{r^2} a^t (A\bar{X}^2 A^t)^{-1} A\bar{X}^2 A^t (A\bar{X}^2 A^t)^{-1} a \\ &= (1 - \delta)^2, \end{aligned}$$

donde $\|\bar{X}\Delta s\| = 1 - \delta$.

Temos também que

$$\|\bar{X}^2(\Delta s)^2\| \leq \|\bar{X}\Delta s\|^2.$$

Assim,

$$\begin{aligned} \|\tilde{e} - \tilde{x}(\alpha)\tilde{s}(\alpha)\|^2 &\leq \|e - (\bar{x} + \alpha\Delta x)(\bar{s} + \alpha\Delta s)\|^2 + |1 - \alpha(\alpha - \gamma)(1 - \delta)^2|^2 \\ &\leq (\delta + \alpha\delta\|\bar{X}\Delta s\| + \alpha^2\|\bar{X}^2(\Delta s)^2\|)^2 + (1 - \alpha(\alpha - \gamma)(1 - \delta)^2)^2 \\ &\leq (\delta + \alpha\delta(1 - \delta) + \alpha^2(1 - \delta)^2)^2 + (1 - \alpha(\alpha - \gamma)(1 - \delta)^2)^2. \end{aligned}$$

Portanto, obteremos

$$\sqrt{(\alpha + \alpha\delta(1 - \delta) + \alpha^2(1 - \delta)^2)^2 + (1 - \alpha(\alpha - \gamma)(1 - \delta)^2)^2} \leq \delta^{\frac{1}{2k}}$$

para uma escolha conveniente de δ , α e γ .

Agora, queremos calcular um centro analítico aproximado do polítopo Ω_K obtido da adição de um corte muito profundo, ou seja, um corte que está além do elipsóide de Dikin. Então, $\tilde{s}(\alpha)$ pode não ser positivo, ou seja, não é mais possível recuperar a viabilidade dual ao longo de Δs . Mas $\tilde{x}(\alpha)$ continua viável interior primal, então podemos usá-lo como ponto inicial para o algoritmo de centralização projetivo primal.

O algoritmo de Goffin-Vial:

Seja $0 < \delta \leq \eta < 1$. Dados o polítopo

$$\Omega = \{y \in \mathbb{R}^m; A^t y \leq c\}$$

e o ponto $(\bar{x}, \bar{y}, \bar{s})$ tal que

$$A^t \bar{y} + \bar{s} = c, \quad \bar{s} > 0$$

$$A \bar{x} = 0, \quad \bar{x} > 0$$

$$\|e - \bar{x} \bar{s}\| \leq \delta,$$

com $\bar{y} = (A \bar{X}^2 A^t)^{-1} A \bar{X}^2 c$. Sejam ainda dados o vetor $a \in \mathbb{R}^m$ tal que $\|a\| = 1$ e o número $K = a^t \bar{y} - \gamma$. Encontrar um centro analítico δ -aproximado de

$$\Omega_K = \{y \in \mathbb{R}^m; \tilde{A}^t y \leq \tilde{c}\},$$

onde $\tilde{A} = (A \ a)$ e $\tilde{c} = \begin{pmatrix} c \\ K \end{pmatrix}$.

Inicialização

Faça

$$x^1 = \tilde{x}(\alpha) = \begin{pmatrix} \bar{x} \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} \Delta x \\ \frac{1-\delta}{r} \end{pmatrix}$$

onde $-1 \leq \alpha \leq 1$ e $\Delta x = -\bar{X}^2 \Delta s$.

Faça $k = 1$

Enquanto $\|\bar{h}(x^k)\| > \delta$ faça

Se $\|\bar{h}(x^k)\| > \eta$ então $x^{k+1} = x^k + \alpha x^k \bar{h}(x^k)$, onde $\alpha = \frac{1}{1 + \eta}$

Senão

$$x^{k+1} = x^k + x^k \bar{h}(x^k)$$

$$k = k + 1$$

Fim.

Capítulo 5

Um Algoritmo para Calcular o Centro Analítico de um Politopo Usando a Trajetória Central

Neste capítulo desenvolveremos um algoritmo para calcular um centro analítico aproximado de um politopo após a adição de um corte tal como foi dado pelo oráculo, ou seja, um corte que pode ser muito profundo. A estratégia que usaremos para recuperar a viabilidade é a de seguir a trajetória central de um problema de programação linear auxiliar no formato dual.

5.1 A Direção de Newton Dual

Consideremos o problema dual (D) como dado na seção 2.3,

$$\begin{aligned}
& \text{maximizar} && b^t y \\
& \text{sujeito a:} \\
(D) & && A^t y + s = c \\
& && s \geq 0.
\end{aligned}$$

O problema penalizado dual associado a (D) é

$$\begin{aligned}
& \text{minimizar} && -\alpha b^T y + p(s) \\
(D)_\alpha & \text{sujeito a:} \\
& && A^T y + s = c \\
& && s > 0,
\end{aligned}$$

onde $p(s) = -\sum_{i=1}^n \ln s_i = -\sum_{i=1}^n \ln (c_i - a_i^t y)$.

Dados y°, s° tais que $s^\circ > 0$ e $A^t y^\circ + s^\circ = c$, vamos descrever a direção de Newton dual a partir de (y°, s°) para a resolução do problema $(D)_\alpha$. Denotaremos a função penalizada dual por

$$f_\alpha(y) = -\alpha b^t y - \sum_{i=1}^n \ln (c_i - a_i^t y),$$

então

$$\nabla f_\alpha(y) = -\alpha b + \sum_{i=1}^n \frac{a_i}{c_i - a_i^t y} = -\alpha b + AS^{-1}e$$

e

$$\nabla^2 f_\alpha(y) = \sum_{i=1}^n \frac{a_i a_i^t}{(c_i - a_i^t y)^2} = AS^{-2}A^t.$$

A direção de Newton dual $(h(y^\circ, \alpha), h(s^\circ, \alpha))$ a partir de (y°, s°) deve satisfazer

$$\nabla^2 f_\alpha(y^\circ)h(y^\circ, \alpha) + \nabla f_\alpha(y^\circ) = 0,$$

conforme (2.2.2).

Definindo $\bar{A} = A(S^\circ)^{-1}$, temos

$$(\bar{A}\bar{A}^t)h(y^\circ, \alpha) - \alpha b + \bar{A}e = 0,$$

donde

$$h(y^\circ, \alpha) = (\bar{A}\bar{A}^t)^{-1}(\alpha b - \bar{A}e)$$

e

$$h(s^\circ, \alpha) = -A^t(\bar{A}\bar{A}^t)^{-1}(\alpha b - \bar{A}e),$$

pois $A^t h(y^\circ, \alpha) + h(s^\circ, \alpha) = 0$.

A direção de Newton dual pode também ser obtida a partir da formulação simétrica do problema dual

minimizar $\hat{x}^t s$

sujeito a:

$$P_A s = P_A c$$

$$s \geq 0,$$

onde \hat{x} é uma solução primal viável arbitrária, para mais detalhes ver GONZAGA[14].

Dados s° viável e $\alpha > 0$, a direção de Newton a partir de s° é dada por

$$h(s^\circ, \alpha) = S^\circ \bar{h}(s^\circ, \alpha),$$

com

$$\bar{h}(s^\circ, \alpha) = -P_{P_A S^\circ}(\alpha S^\circ \hat{x} - e).$$

Mas, usando a proposição 2.1.3, é fácil provar que

$$P_{P_A S^\circ} = \tilde{P}_{A(S^\circ)^{-1}},$$

então

$$\begin{aligned} \bar{h}(s^\circ, \alpha) &= -\tilde{P}_{A(S^\circ)^{-1}}(\alpha S^\circ \hat{x} - e) \\ &= -\bar{A}^t(\bar{A}\bar{A}^t)^{-1}\bar{A}(\alpha S^\circ \hat{x} - e) \end{aligned}$$

e

$$\begin{aligned} h(s^\circ, \alpha) &= -A^t(\bar{A}\bar{A}^t)^{-1}(\alpha A \hat{x} - \bar{A}e) \\ &= -A^t(\bar{A}\bar{A}^t)^{-1}(\alpha b - \bar{A}e) \end{aligned}$$

Da análise feita para o problema primal no capítulo 3, podemos concluir que:

1. $\delta(y, \alpha) = \delta(s, \alpha) = \|\bar{h}(s, \alpha)\|$.
2. Se $\delta(s, \alpha) < 1$, então $\delta(s + h(s, \alpha), \alpha) \leq \delta(s, \alpha)^2$.
3. Se $\delta(s, \alpha) < 1$, então $x(s, \alpha) = \frac{s^{-1}}{\alpha}(e - \bar{h}(s, \alpha))$ é solução primal viável e

$$\delta(s, x(s, \alpha), \alpha) = \delta(s, \alpha).$$

Observemos que se $\alpha = 0$ então a solução ótima do problema $(D)_\alpha$ é o centro analítico do politopo definido por

$$\Omega = \{y \in \mathbb{R}^m; A^t y \leq c\}$$

que também podemos escrever em termo das folgas, pois A é de posto máximo,

$$\mathcal{S}_\Omega = \{s; A^t y + s = c, s \geq 0\}.$$

Denotemos a proximidade de um ponto y ao centro analítico de Ω ou de s ao centro analítico de \mathcal{S}_Ω por

$$\delta(y) = \delta(s) = \|\bar{h}(s)\| = \|\tilde{P}_{AS^{-1}}e\|.$$

A condição de otimalidade para o problema $(D)_{\alpha=0}$ é $\tilde{P}_{AS^{-1}}e = 0$, o que é equivalente a $e \in \mathcal{N}(AS^{-1})$ donde $AS^{-1}e = 0$, fazendo $S^{-1}e = x$ temos as condições de otimalidade primal-dual para o centro analítico

$$\begin{aligned} xs &= e \\ Ax &= 0 \\ A^t y + s &= c \\ x, s &> 0. \end{aligned} \tag{5.1.1}$$

Como já vimos na seção 3.2, a primeira equação de (5.1.1) fornece uma medida de proximidade primal-dual que denotaremos por

$$\delta(x, s) = \|xs - e\|.$$

Uma versão para o caso dual da proposição 3.2.1 nos dá uma relação entre as medidas de proximidade dual e primal-dual, a saber,

$$\delta(y) = \delta(s) \leq \delta(x, s). \quad (5.1.2)$$

5.2 O Problema Auxiliar

Seja dado o politopo no formato dual

$$\Omega = \{y \in \mathbb{R}^m; A^t y \leq c\},$$

onde $c \in \mathbb{R}^n$ e $A \in \mathbb{R}^{m \times n}$ é de posto máximo igual a m . Suporemos que o conjunto de pontos interiores

$$\overset{\circ}{\Omega} = \{y \in \mathbb{R}^m; A^t y < c\}$$

é não-vazio. Seja y° um centro analítico δ -aproximado de Ω tal que $\delta(y^\circ) \leq \delta < \beta < 1$. Dados $a \in \mathbb{R}^m$ tal que $\|a\| = 1$ e $K \in \mathbb{R}$, queremos encontrar um centro analítico β -aproximado do politopo

$$\Omega_K = \{y \in \mathbb{R}^m; A^t y \leq c, a^t y \leq K\},$$

que podemos escrever em termo das folgas como

$$\mathcal{S}_{\Omega_K} = \{(s, s_{n+1}); A^t y + s = c, a^t y + s_{n+1} = K, s, s_{n+1} \geq 0\}.$$

Denotaremos a proximidade de y ao centro analítico de Ω_K por $\tilde{\delta}(y)$.

Com os dados acima, construímos o seguinte problema de programação linear auxiliar no formato dual

$$\begin{aligned}
& \text{maximizar} && -a^t y \\
& \text{sujeito a:} && \\
(D_A) & && A^t y + s = c \\
& && s \geq 0,
\end{aligned}$$

que tem como problema primal associado

$$\begin{aligned}
& \text{minimizar} && c^t x \\
& \text{sujeito a:} && \\
(P_A) & && A^t y = -a \\
& && x \geq 0.
\end{aligned}$$

A este par de problemas associamos a sua trajetória central primal-dual

$$\alpha > 0 \longmapsto (x(\alpha), s(\alpha)).$$

O próximo resultado vai nos garantir que quando temos um ponto (\bar{y}, \bar{s}) central β -aproximado de (D_A) tal que $a^t \bar{y} \leq K$ e $\alpha = \frac{1}{K - a^t \bar{y}}$, então \bar{y} é centro analítico β -aproximado de Ω_K , ou seja, $\tilde{\delta}(\bar{y}) \leq \beta$. Este resultado é consequência dos trabalhos de HUARD[19] e RENEGAR[39].

Proposição 5.2.1 *Considere uma solução $(\bar{x}, \bar{y}, \bar{s})$ primal-dual viável para os problemas (P_A) e (D_A) tal que $a^t \bar{y} \leq K$, $K \in \mathbb{R}$, e $\alpha > 0$. Então,*

$$\tilde{\delta}(\bar{y})^2 \leq \delta(\bar{x}, \bar{s}, \alpha)^2 + (\alpha(K - a^t \bar{y}) - 1)^2.$$

Em particular, se $\alpha = \frac{1}{K - a^t \bar{y}}$ então

$$\tilde{\delta}(\bar{y}) \leq \delta(\bar{x}, \bar{s}, \alpha).$$

Demonstração. Seja $(\bar{x}, \bar{y}, \bar{s})$ uma solução primal-dual viável para (P_A) e (D_A) tal que $a^t \bar{y} \leq K$ e seja $\alpha > 0$. Vamos construir os vetores:

$$\tilde{x} = \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} \alpha \bar{x} \\ \alpha \end{pmatrix}, \quad \tilde{s} = \begin{pmatrix} s \\ s_{n+1} \end{pmatrix} = \begin{pmatrix} \bar{s} \\ K - a^t \bar{y} \end{pmatrix} \quad \text{e} \quad \tilde{y} = \bar{y}.$$

Então, temos por substituição direta que

$$\begin{pmatrix} A^t \\ a^t \end{pmatrix} \bar{y} + \begin{pmatrix} \bar{s} \\ K - a^t \bar{y} \end{pmatrix} = \begin{pmatrix} A^t \bar{y} + \bar{s} \\ a^t \bar{y} + K - a^t \bar{y} \end{pmatrix} = \begin{pmatrix} c \\ K \end{pmatrix},$$

da hipótese temos que $\bar{s} \geq 0$ e $s_{n+1} = K - a^t \bar{y} \geq 0$, logo $\tilde{s} \in \mathcal{S}_{\Omega_K}$. Além disso, usando a viabilidade de (P_A) , temos

$$(A \ a) \begin{pmatrix} \alpha \bar{x} \\ \alpha \end{pmatrix} = \alpha A \bar{x} + \alpha a = -\alpha a + \alpha a = 0.$$

Portanto, os vetores $(\tilde{x}, \tilde{y}, \tilde{s})$ formam uma solução primal-dual para o problema de centro analítico de Ω_K e, de (5.1.2), temos

$$\begin{aligned} \tilde{\delta}(y)^2 &\leq \delta(\tilde{x}, \tilde{s})^2 \\ &= \|\tilde{x} \tilde{s} - e\|^2 \\ &= \left\| \begin{array}{c} \alpha \bar{x} \bar{s} - e \\ \alpha(K - a^t \bar{y}) - 1 \end{array} \right\|^2 \\ &= \|\alpha \bar{x} \bar{s} - e\|^2 + (\alpha(K - a^t \bar{y}) - 1)^2. \end{aligned}$$

□

Dado $y^\circ \in \overset{\circ}{\Omega}$ com $\delta(y^\circ) \leq \delta$, vamos gerar uma solução primal-dual (x, y, s) próxima à trajetória central do par de problemas auxiliares (D_A) e (P_A) , para algum $\alpha > 0$.

Proposição 5.2.2 *Considere um centro analítico (y°, s°) δ -aproximado de Ω . Dado $\alpha > 0$ com $\delta(s^\circ, \alpha) < 1$, considere o par (x_α, s_α) definido por:*

$$s_\alpha = S^\circ(e + \bar{h}(s^\circ, \alpha)) \tag{5.2.1}$$

e

$$x_\alpha = \frac{(S^\circ)^{-1}}{\alpha}(e - \bar{h}(s^\circ, \alpha)) \quad (5.2.2).$$

Então (x_α, s_α) é viável para (D_A) e (P_A) e tal que

$$\delta(x_\alpha, s_\alpha, \alpha) \leq (\alpha r + \delta)^2,$$

onde $r = \|\bar{A}^t(\bar{A}\bar{A}^t)^{-1}a\|$ com $\bar{A} = A(S^\circ)^{-1}$.

Demonstração. Dados (y°, s°) centro analítico δ -aproximado de Ω , $\alpha > 0$ com $\delta(s^\circ, \alpha) < 1$ e os vetores

$$s_\alpha = S^\circ(e + \bar{h}(s^\circ, \alpha)) \quad (5.2.1)$$

e

$$x_\alpha = \frac{(S^\circ)^{-1}}{\alpha}(e - \bar{h}(s^\circ, \alpha)) \quad (5.2.2).$$

onde $\bar{h}(s^\circ, \alpha) = \bar{A}^t(\bar{A}\bar{A}^t)^{-1}(\alpha a + \bar{A}e)$. Como $\|\bar{h}(s^\circ, \alpha)\| < 1$, segue imediato que (x_α, s_α) é primal-dual viável e de 3) acima

$$\delta(s^\circ, x_\alpha, \alpha) = \delta(s^\circ, \alpha).$$

Vamos calcular $\delta(x_\alpha, s_\alpha, \alpha)$.

$$\delta(x_\alpha, s_\alpha, \alpha) = \|\alpha x_\alpha s_\alpha - e\| = \|- \bar{h}(s^\circ, \alpha)^2\| \leq \|\bar{h}(s^\circ, \alpha)\|^2.$$

Mas

$$\begin{aligned} \|\bar{h}(s^\circ, \alpha)\| &\leq \alpha\|\bar{A}^t(\bar{A}\bar{A}^t)^{-1}a\| + \|\bar{A}^t(\bar{A}\bar{A}^t)^{-1}\bar{A}e\| \\ &\leq \alpha r + \delta(y^\circ) \\ &\leq \alpha r + \delta, \end{aligned}$$

lembrando que $\delta(y^\circ) = \|\tilde{P}_{\bar{A}}e\|$ e fazendo $r = \|\bar{A}^t(\bar{A}\bar{A}^t)^{-1}a\|$, completamos a demonstração. \square

5.3 Introduzindo um Corte Pouco Profundo

Agora pretendemos responder à seguinte questão: Dado $y^\circ \in \Omega$ com $\delta(y^\circ) < \delta$, qual é o menor valor possível de K para que o ponto y_α associado a s_α contruído na proposição 5.2.2 satisfaça $\tilde{\delta}(y_\alpha) \leq \beta < 1$, com $\tilde{\delta}(\cdot)$ proximidade em Ω_K .

Em outras palavras, queremos introduzir um corte o mais profundo possível para o qual temos um centro analítico β -aproximado.

Procedimento

Dado $y^\circ \in \Omega$ com $\delta(y^\circ) \leq \delta$ e $s^\circ = c - A^t y^\circ$. Construimos, para $\alpha > 0$ tal que $\delta(y^\circ, \alpha) < 1$, os pontos (x_α, s_α) , como na proposição 5.2.2,

$$s_\alpha = S^\circ(e + \bar{h}(s^\circ, \alpha))$$

$$x_\alpha = \frac{S^{\circ-1}}{\alpha}(e - \bar{h}(s^\circ, \alpha))$$

Proposição 5.3.1 *Sejam dados (x_α, s_α) nas condições acima, com y_α tal que $A^t y_\alpha + s_\alpha = c$, e $a^t y_\alpha < K$, com $K = a^t y^\circ + \gamma$, $\gamma \in \mathbb{R}$. Então,*

$$\tilde{\delta}(y_\alpha)^2 \leq (\alpha r + \delta)^4 + (1 + \alpha r \delta - \alpha \gamma - \alpha^2 r^2)^2.$$

Demonstração. Da proposição 5.2.1, temos para Ω_K que

$$\begin{aligned} \tilde{\delta}(y_\alpha)^2 &\leq \delta(x_\alpha, s_\alpha, \alpha)^2 + (\alpha(K - a^t y_\alpha) - 1)^2 \\ &\leq (\alpha r + \delta)^4 + (\alpha(K - a^t y_\alpha) - 1)^2. \end{aligned}$$

Como $K = a^t y^\circ + \gamma$, então

$$\begin{aligned} K - a^t y_\alpha &= \gamma - a^t(y_\alpha - y^\circ) \\ &= \gamma - a^t h(y^\circ, \alpha). \end{aligned}$$

Mas $\bar{A}^t h(y^\circ, \alpha) = -\bar{h}(s^\circ, \alpha)$, então

$$h(y^\circ, \alpha) = -(\bar{A}\bar{A}^t)^{-1}\bar{A}\bar{h}(s^\circ, \alpha),$$

logo,

$$a^t h(y^\circ, \alpha) = a^t (\bar{A}\bar{A}^t)^{-1} \bar{A}\bar{h}(s^\circ, \alpha),$$

fazendo, $\tilde{x}^t = -\bar{A}^t(\bar{A}\bar{A}^t)^{-1}a$, temos

$$a^t h(y^\circ, \alpha) = \tilde{x}^t \bar{h}(s^\circ, \alpha) = \hat{x}^t h(s^\circ, \alpha),$$

pois $\hat{x} = S^\circ \tilde{x}$ e $h(s^\circ, \alpha) = (S^\circ)^{-1} \bar{h}(s^\circ, \alpha)$. Então

$$\begin{aligned} \bar{h}(s^\circ, \alpha) &= -\bar{A}^t(\bar{A}\bar{A}^t)^{-1}(-\alpha a - \bar{A}e) \\ &= -\alpha \tilde{x} + \tilde{P}_{\bar{A}}e, \end{aligned}$$

donde segue,

$$\begin{aligned} a^t h(y^\circ, \alpha) &= -\alpha \|\tilde{x}\|^2 + \tilde{x}^t \tilde{P}_{\bar{A}}e \\ &\leq -\alpha r^2 + \delta r. \end{aligned}$$

Portanto, temos o pior caso quando $a^t h(y^\circ, \alpha) = -\alpha r^2 + \delta r$, logo

$$K - a^t y_\alpha = \gamma + \alpha r^2 - \delta r$$

e

$$\tilde{\delta}(y_\alpha)^2 \leq (\alpha r + \delta)^4 + (1 + \alpha r \delta - \alpha \gamma - \alpha^2 r^2)^2.$$

□

Podemos concluir do resultado acima que $\tilde{\delta}(y_\alpha) = \beta < 1$ depende de uma escolha conveniente de $\hat{\alpha} = \alpha r$, $\hat{\gamma} = \frac{\gamma}{r}$ e δ . Isto pode ser observado pela tabela abaixo:

δ	β	$\hat{\alpha}$	$\hat{\gamma}$
0	1	0.8	-0.52
0.01	0.99	0.9	-0.48
0.01	0.9	0.77	-0.33
0.1	0.9	0.6	-0.1
0.2	0.8	0.56	0.47

Adição de um Corte Pouco Profundo em Ω

Aqui estudaremos o que se pode fazer em uma única iteração, usando os resultados anteriores. Dados $y^\circ \in \Omega$ com $\delta(y^\circ) \leq \delta$, definimos $s^\circ = c - A^t y^\circ$ e $\bar{A} = A(S^\circ)^{-1}$. Vamos construir os seguintes mapeamentos:

$$\alpha > 0 \text{ tal que } \delta(s^\circ, \alpha) < 1 \mapsto \left\{ \begin{array}{l} h(y^\circ, \alpha) = -(\bar{A}\bar{A}^t)^{-1}(\alpha a + \bar{A}e) \\ \bar{h}(s^\circ, \alpha) = -\bar{A}^t h(y^\circ, \alpha) \\ s_\alpha = s^\circ + S^\circ \bar{h}(s^\circ, \alpha) \\ x_\alpha = \frac{(S^\circ)^{-1}}{\alpha}(e - \bar{h}(s^\circ, \alpha)) \\ y_\alpha = y^\circ + h(y^\circ, \alpha) \\ \delta(x_\alpha, s_\alpha, \alpha) = \|\alpha x_\alpha s_\alpha - e\|. \end{array} \right.$$

Observemos que a única operação trabalhosa nesta construção é o cálculo de $(\bar{A}\bar{A}^t)^{-1}a$ e $(\bar{A}\bar{A}^t)^{-1}\bar{A}e$, que necessita uma decomposição e duas retro-substituições.

O Problema Direto: Dado $K \in \mathbb{R}$, encontrar $\bar{y} \in \Omega_K$ tal que $\delta(\bar{y}) \leq \beta < 1$.

Se K for suficientemente grande, corte pouco profundo, o problema poderá ser resolvido em uma única iteração, usando a expressão da proposição 5.2.1, basta encontrar $\alpha > 0$ tal que

$$\delta(x_\alpha, s_\alpha, \alpha)^2 + (\alpha(K - a^t y_\alpha) - 1)^2 \leq \beta^2,$$

com $a^t y_\alpha < K$. Este é um problema unidimensional (α é a única variável) e pode ser resolvido como

$$\begin{array}{l} \text{minimizar } \delta(x_\alpha, s_\alpha, \alpha)^2 + (\alpha(K - a^t y_\alpha) - 1)^2 \\ \text{sujeito a:} \end{array}$$

$$\begin{array}{l} K - a^t y_\alpha \geq 0 \\ \alpha > 0. \end{array}$$

A resolução é simples, podemos calcular α_{max} tal que $\delta(x_\alpha, s_\alpha, \alpha) = \beta$, por tentativa ou a partir da definição, o que produz uma equação quártica, e α_{min} tal

que $K - a^t y_{\alpha_{min}} = 0$. Se $\alpha_{min} < \alpha_{max}$, então a minimização pode ser feita entre esses dois valores. Se resultar em um valor inferior a β , então temos uma solução para o problema. Senão, o corte é muito profundo e necessita mais iterações, descritas na próxima seção.

O Problema Inverso: Buscar o valor do menor K tal que é possível introduzir o corte $a^t y \leq K$ pela nossa técnica.

Este problema pode ser escrito como

$$\begin{aligned} & \text{minimizar} && K \\ & \text{sujeito a:} && \\ & && \delta(x_\alpha, s_\alpha, \alpha)^2 + (\alpha(K - a^t y_\alpha) - 1)^2 \leq \beta^2 \\ & && \alpha > 0, K \in \mathbb{R}, \end{aligned}$$

e é também um problema fácil, pois na realidade tem apenas uma dimensão para cada valor de $\alpha > 0$ tal que $\delta(x_\alpha, s_\alpha, \alpha) \leq \beta$, então basta determinar $K > a^t y_\alpha$ satisfazendo a restrição com igualdade.

5.4 Um Algoritmo para Introdução de um Corte em Ω

Nesta seção vamos desenvolver um algoritmo para introdução de um corte em um dado politopo. Se o corte é pouco profundo, já vimos na seção anterior que, basta uma única iteração para encontrar um centro analítico β -aproximado de Ω_K . Se o corte é muito profundo necessitamos de mais iterações. Essas iterações são obtidas seguindo a trajetória central do problema auxiliar (D_A).

Inicialmente, vamos apresentar um algoritmo geral, onde seguiremos a trajetória utilizando qualquer algoritmo de trajetória central. Em seguida, para fins de estudo da complexidade de um algoritmo de plano de corte para resolver o

problema de viabilidade convexa, apresentaremos um algoritmo onde seguimos a trajetória usando o algoritmo de trajetória central de passos curtos dado na seção 3.1.

Algoritmo de Trajetória Central em uma Vizinhança

Os algoritmos comuns de trajetória central geram pontos (x^k, s^k, α^k) em uma β -vizinhança da trajetória. Os pontos satisfazem $\delta(x^k, s^k, \alpha^k) \leq \beta$ e, além disso, para $\lambda \in [0, 1]$,

$$\delta(\lambda x^k + (1 - \lambda)x^{k+1}, \lambda s^k + (1 - \lambda)s^{k+1}, \lambda \alpha^k + (1 - \lambda)\alpha^{k+1}) \leq \beta,$$

isto é, não somente os pontos estão na vizinhança, mas o mesmo ocorre com os segmentos que unem cada par de iterados consecutivos.

Algoritmo 5.4.1 Dados $y^\circ \in \Omega$ tal que $\delta(y^\circ) \leq \delta \leq 0.1$, $K \in \mathbb{R}$ e $0 < \beta < 1$.

Inicialização

Executar o procedimento para introdução de um corte pouco profundo, como foi visto na seção 5.3. Se for possível introduzir o corte em uma iteração, PARE.

Senão, faça α^1 tal que $(\alpha^1 r + \delta)^2 = \beta < 1$ e utilize (5.2.1) e (5.2.2) para calcular (x^1, s^1) tal que $\delta(x^1, s^1, \alpha^1) \leq (\alpha^1 r + \delta)^2$.

Seguir a Trajetória

Utilize qualquer algoritmo de trajetória central na vizinhança β da trajetória partindo de $(x^1, y^1, s^1, \alpha^1)$ para obter uma seqüência de pontos $(x^k, y^k, s^k, \alpha^k)$ tal que $\delta(x^k, s^k, \alpha^k) \leq \beta$, até que se satisfaça $K - a^t y^k \geq \frac{1}{\alpha^k}$.

Finalização (ver o Problema Direto acima)

Definir

$$(x_\lambda, y_\lambda, s_\lambda, \alpha_\lambda) = (1 - \lambda)(x^{k-1}, y^{k-1}, s^{k-1}, \alpha^{k-1}) + \lambda(x^k, y^k, s^k, \alpha^k),$$

para todo $\lambda \in [0, 1]$.

Resolver

$$\begin{aligned} &\text{minimizar } \delta(x_\lambda, s_\lambda, \alpha_\lambda)^2 + (\alpha_\lambda(K - a^t y_\lambda) - 1)^2 \\ &\text{sujeito a:} \end{aligned}$$

$$K - a^t y_\lambda \geq 0$$

$$\lambda \in [0, 1].$$

Centralizar usando um dos procedimentos dados na seção 3.3.

Comentários:

Pode-se escolher simplesmente λ tal que $\alpha_\lambda(K - a^t y_\lambda) = 1$. Neste caso, $\delta(x_\lambda, s_\lambda, \alpha_\lambda) \leq \beta$ e, pela proposição 5.2.1, temos $\tilde{\delta}(y_\lambda) \leq \beta$.

O algoritmo pode basear-se em qualquer algoritmo de trajetória central. O algoritmo mais eficiente em implementações atuais é o preditor-corretor de Mehrotra[26]. A seguir vamos apresentar uma particularização do algoritmo 5.4.1, onde usaremos um algoritmo dual de passos curtos para seguir a trajetória central. A finalidade deste algoritmo é estudar a complexidade do método de plano de corte resultante e não é eficiente do ponto de vista prático. Observemos que métodos como Mizuno-Todd-Ye[32] e o algoritmo de passos mais longos, ver [18], dão passos pelo menos iguais ao de passos curtos e levam aos mesmos resultados de complexidade.

Algoritmo 5.4.2 Dados $y^\circ \in \Omega$ com $\delta(y^\circ) \leq \delta \leq 0.1$, $K \in \mathbb{R}$.

Inicialização

Proceda como na inicialização do algoritmo 5.4.1

Centralização

$$l = 1$$

Enquanto $\delta(x^l, s^l, \alpha^l) > 0.01$ faça

$$s^{l+1} = s^l + h(s^l, \alpha^l)$$

$$x^{l+1} = \frac{(S^l)^{-1}}{\alpha^l} (e - \bar{h}(s^l, \alpha^l))$$

$$l = l + 1$$

Fim

$$(x^1, s^1, \alpha^1) = (x^l, s^l, \alpha^l)$$

$$k = 1$$

Seguir a trajetória

Repita

Calcule

$$\alpha^{k+1} \text{ tal que } \delta(x^k, s^k, \alpha^{k+1}) = \delta$$

$$s^{k+1} = s^k + h(s^k, \alpha^{k+1})$$

$$x^{k+1} = \frac{(S^k)^{-1}}{\alpha^{k+1}} (e - \bar{h}(s^k, \alpha^{k+1}))$$

$$k = k + 1$$

Até que $K - a^t y^k \geq \frac{1}{\alpha^k}$.

Finalização

Como no algoritmo anterior.

O resultado a seguir vai nos assegurar que a função barreira cresce de uma constante, no primeiro passo do algoritmo 5.4.2.

Proposição 5.4.1 *Considere $s^\circ \in \Omega$ tal que $\delta(s^\circ) \leq 0.01$ e $\alpha > 0$ tal que $\|s(\alpha) - s^\circ\|_{s(\alpha)} = 0.4$. Então, $\delta(s(\alpha)) \geq 0.25$ e $p(s(\alpha)) \geq p(s(0)) + 0.026$.*

Demonstração. Sem perda de generalidade, consideremos que o centro analítico de Ω é $s(0) = e$. Suponhamos que $\|s^\circ - s(\alpha)\|_{s(\alpha)} = 0.4$ e $\delta(s^\circ) \leq 0.01$.

Vamos inicialmente mostrar que $s_i(\alpha) \geq 0.7$ para $i = 1, \dots, n$. Temos

$$\left| \frac{s_i^\circ}{s_i(\alpha)} - 1 \right| \leq 0.4$$

e portanto, $\frac{s_i(\alpha)}{s_i^\circ} \geq \frac{1}{1.4}$. De ROOS *et. al.*[40], como $\delta(s^\circ) \leq 0.01$, temos $|s_i^\circ - 1| \leq \|s^\circ - e\| \leq 0.011$ e então $s_i^\circ \geq 0.98$. Conclui-se que $s_i(\alpha) \geq \frac{0.98}{1.4} = 0.7$. Temos agora,

$$0.4 = \|s^\circ - s(\alpha)\|_{s(\alpha)} \leq \left\| \frac{1}{s(\alpha)} \right\|_\infty \|s^\circ - s(\alpha)\| \leq \frac{1}{0.7} \|s^\circ - s(\alpha)\|,$$

donde $\|s^\circ - s(\alpha)\| \geq 0.28$ o que implica

$$\|e - s(\alpha)\| \geq 0.28 - \|s^\circ - e\| \geq 0.28 - 0.011.$$

Usando novamente ROOS *et. al.*[40],

$$\delta(s(\alpha)) \geq \frac{0.28 - 0.011}{1.05} \geq 0.25.$$

A função barreira satisfaz para $h = s(\alpha) - e$, ver GONZAGA[15],

$$p(e + h) \geq \frac{\|h\|^2}{2} - \frac{\|h\|^3}{3} \geq 0.026.$$

□

Corolário 5.4.1 *O ponto resultante do primeiro passo do algoritmo 5.4.2 satisfaz*

$$p(s(\alpha^1)) \geq p(s(0)) + 0.026.$$

Demonstração. Da tabela da seção 5.3, $\alpha^1 > 0.4$ e $\delta(s^\circ, \alpha^1) > 0.4$, portanto, $\|s^\circ - s(\alpha^1)\|_{s(\alpha^1)} \geq 0.4$. Então, existe $\alpha \in (0, \alpha^1)$ tal que $\|s^\circ - s(\alpha)\| = 0.4$ e $p(s(\alpha^1)) \geq p(s(\alpha))$, pois $p(x(\cdot))$ é crescente. Portanto, temos que

$$p(s(\alpha^1)) \geq p(s(\alpha)) \geq p(s(0)) + 0.026.$$

□

Como consequência dos resultados acima e das proposições 3.4.2 e 3.4.6. temos que se na iteração k o critério de parada do algoritmo 5.4.2 é satisfeito,

então a função barreira terá uma variação ao longo da trajetória central de pelo menos

$$p(s(\alpha^k)) \geq p(s(0)) + j(0.007\delta(s(\alpha^1))) + 0.026 \geq p(s(0)) + j(0.001) + 0.026,$$

onde j é o número de passos sobre a trajetória.

Capítulo 6

Um Algoritmo de Plano de Corte

Consideramos os problemas de viabilidade convexa finito e infinito e para resolvê-los desenvolveremos um método de plano de corte no centro analítico onde os cortes são adicionados ao politopo tal como foram geradas pelo oráculo. A análise de convergência e complexidade segue as mesmas idéias de GOFFIN *et al.*[13].

6.1 O Problema de Viabilidade Convexa

O problema de viabilidade convexa é apresentado na seguinte forma: Seja $\Gamma \subset \mathbb{R}^m$ convexo e de interior não vazio. Achar um ponto $\bar{y} \in \Gamma$. Este é um problema de programação convexa que pode incluir o problema de programação linear, onde Γ é um poliedro, e o problema de otimização convexa que é dado como segue: Seja $f : \mathbb{R}^m \rightarrow \mathbb{R}$ uma função convexa sobre uma caixa $B = \{y \in \mathbb{R}^m; 0 \leq y \leq e\}$. Resolver o problema de otimização convexa é: Dado $\varepsilon > 0$,

achar $y^* \in B$ tal que

$$f(y^*) \leq \min_{y \in B} f(y) + \varepsilon.$$

Como o conjunto $\{y \in B \subset \mathbb{R}^m; f(y) - \min_{y \in B} f(y) \leq \varepsilon\}$ é convexo, então o problema de otimização convexa é equivalente a um problema de viabilidade.

Suporemos que Γ está contido num politopo pré-definido Ω° . O conjunto Γ pode ser definido por um sistema de desigualdades lineares finito ou infinito. No caso infinito, Γ é definido implicitamente por um oráculo o qual, para $\forall \bar{y} \in \Omega^\circ$ dado, responde que $\bar{y} \in \Gamma$ ou gera um hiperplano separador $a^t y = K$, onde $a \in \mathbb{R}^m$ e $K \in \mathbb{R}$, tal que

$$\Gamma \subset \{y \in \mathbb{R}^m; a^t y \leq K\}.$$

Um algoritmo de plano de corte no centro analítico constrói uma seqüência de centros analíticos aproximados $\{y^k\}$ em Ω . As respostas do oráculo nos centros analíticos aproximados definem com o politopo inicial Ω° uma aproximação poliedral

$$\Omega = \{y \in \mathbb{R}^m; A^t y \leq c\}$$

de Γ . Suporemos que a matriz $A \in \mathbb{R}^{m \times n}$ tem posto máximo m . Então, como já foi visto na seção 5.2, podemos representar Ω em termo das folgas

$$\mathcal{S}_\Omega = \{s \in \mathbb{R}^n; A^t y + s = c, s \geq 0\}$$

que é equivalente a

$$\mathcal{S}_\Omega = \{s \in \mathbb{R}^n; P_A s = P_A c, s \geq 0\},$$

onde P_A é a matriz de projeção sobre o espaço nulo de A .

6.2 O Potencial de um Politopo

A função barreira de Ω é definida por

$$p(s) = - \sum_{i=1}^n \ln s_i,$$

com $s_i = c_i - a_i^t y > 0$, $\forall i = 1, \dots, n$

O potencial de Ω é dado pelo valor de $p(\cdot)$ no centro analítico de Ω , o qual denotaremos por

$$P(\Omega) = \min_{s \in S_\Omega} p(s).$$

Seja y^a o centro analítico de Ω . Considere

$$\Omega_K = \{y \in \mathbb{R}^m; A^t y \leq c, a^t y \leq K\},$$

onde $K = a^t y^a + \gamma r$, $\gamma \leq 0$ e $r = \sqrt{a^t (A(S_a)^{-2} A^t)^{-1} a}$ é a folga correspondendo ao ponto mais distante na direção de a no elipsóide de Dikin, ou seja, é a folga da solução y^E do problema

$$\text{maximizar } -a^t(y - y^a)$$

sujeito a:

$$\|(S_a)^{-1} A^t (y - y^a)\| \leq 1,$$

donde

$$y^E = y^a - \frac{(A(S_a)^{-2} A^t)^{-1} a}{\sqrt{a^t (A(S_a)^{-2} A^t)^{-1} a}}.$$

Então,

$$s^E = a^t y^a - a^t y^E = -a^t (y^E - y^a) = r.$$

Observemos que $r = \|\hat{x}\|$, onde $\hat{x} = -\bar{A}^t (\bar{A} \bar{A}^t)^{-1} a$ e $\bar{A} = A(S^\circ)^{-1}$

Proposição 6.2.1 *Considere Ω, Ω_K e r como dados acima. Então,*

$$P(\Omega_K) \geq P(\Omega) - \ln r + j\sigma + \xi,$$

para $\sigma = 0.001 + \ln(1 + \frac{\nu}{\sqrt{n}})$, $\xi = \ln \alpha^1 + 0.026$ e j um inteiro não-negativo.

Demonstração. Considerando Ω e Ω_K em termo das folgas

$$\mathcal{S}_\Omega = \{s \in \mathbb{R}^n; A^t y + s = c, s \geq 0\}$$

e

$$\mathcal{S}_{\Omega_K} = \{s \in \mathbb{R}^n; A^t y + s = c, a^t y + s_{n+1} = K, s, s_{n+1} \geq 0\}.$$

Sejam s^a e \bar{s}^a os centros analíticos de \mathcal{S}_Ω e \mathcal{S}_{Ω_K} respectivamente. Como o centro analítico de Ω_K é calculado seguindo a trajetória central do problema de programação linear auxiliar

$$(P_A) \quad \begin{array}{ll} \text{minimizar} & a^t y \\ \text{sujeito a:} & \\ & A^t y \leq c \end{array}$$

que pode ser escrito em termos das folgas como

$$\begin{array}{ll} \text{minimizar} & \hat{x}^t s \\ \text{sujeito a:} & \\ & P_A s = P_A c \\ & s \geq 0 \end{array}$$

onde $a^t y = \hat{x}^t s - c^t \hat{x}$ e $\hat{x} = S_a^{-2} A^t (A S_a^{-2} A^t)^{-1} a$. Como o potencial de \mathcal{S}_Ω é dado por

$$P(\Omega) = -\sum_{i=1}^n \ln s_i^a = p(s^a)$$

e o potencial de \mathcal{S}_{Ω_K} é a função

$$P(\Omega_K) = -\ln \frac{1}{\alpha^k} - \sum_{i=1}^n \ln \bar{s}_i^a = \ln \alpha^k - \sum_{i=1}^n \ln \bar{s}_i^a,$$

onde $\alpha^k \geq \frac{\alpha^1}{r} (1 + \frac{\nu}{\sqrt{n}})^j$, com $\nu = 0.1$ e j igual ao número de passos ao longo da trajetória central do problema (P_A). Então, temos que

$$P(\Omega_K) \geq -\sum_{i=1}^n \ln \bar{s}_i^a - \ln r + j \ln(1 + \frac{\nu}{\sqrt{n}}) + \ln \alpha^1.$$

Agora, usando as proposições 3.4.2, 3.4.6 e 5.4.1, temos que

$$-\sum_{i=1}^n \ln \bar{s}_i^a \geq p(s^a) + j(0.007\delta(s(\alpha^1))) + 0.026 \geq p(s^a) + j(0.001) + 0.026$$

Portanto,

$$P(\Omega_K) \geq P(\Omega) - \ln r + j(0.001) + \ln(1 + \frac{\nu}{\sqrt{n}}) + \ln \alpha^1 + 0.026.$$

Fazendo $\sigma = 0.001 + \ln(1 + \frac{\nu}{\sqrt{n}})$ e $\xi = \ln \alpha^1 + 0.026$ que é negativo, pois $0.4 \leq \alpha^1 < 1$ conforme a tabela da seção 5.3, temos a demonstração completa. \square

6.3 O Algoritmo de Plano de Corte

a) Caso Finito

Dado $\Gamma = \{y \in \mathbb{R}^m; a_i^t y \leq c_i, i = 1, 2, \dots, n\}$. Suponhamos que os vetores a_i são normalizados de modo que

$$\|a_i\| = 1, i = 1, \dots, n$$

e que o conjunto Γ está contido numa caixa $\{y \in \mathbb{R}^m; 0 \leq y \leq e\}$.

Algoritmo 6.3.1

Inicialização

Sejam

$$A^0 = (I \ -I) \in \mathbb{R}^{m \times 2m}, c^0 = \begin{pmatrix} e \\ 0 \end{pmatrix} \in \mathbb{R}^{2m}$$

$$y^0 = \frac{e}{2} \in \mathbb{R}^m, s^0 = c^0 - (A^0)^t y^0 = \frac{e}{2} \in \mathbb{R}^{2m}$$

$$x^0 = (s^0)^{-1} = 2e \in \mathbb{R}^{2m}.$$

O ponto y^0 é o centro analítico de

$$\Omega^k = \{y \in \mathbb{R}^m; (A^k)^t y \leq c^k\}$$

para $k = 0$.

Se $y^k \notin \Gamma$, então

REPITA

Escolha o maior j tal que $a_j^t y^k > c_j$.

Faça $a_{2m+k+1} = a_j$ e

$$\Omega^{k+1} = \{y \in \mathbb{R}^m; (A^{k+1})^t y \leq c^{k+1}\}$$

$$\text{onde } A^{k+1} = (A^k \ a_{2m+k+1}) \text{ e } c^{k+1} = \begin{pmatrix} c^k \\ c_j \end{pmatrix}.$$

Calcule um centro analítico δ -aproximado $s^{k+1} \in \mathcal{S}_{\Omega^{k+1}}$ usando o algoritmo 5.4.2.

Calcule y^{k+1} .

Faça $k = k + 1$.

ATÉ $y^k \in \Gamma$.

Observações 6.3.1

1. O número total de restrições em Ω^k é no máximo $2m + n$, $\forall k \geq 0$.
2. $\Omega^k \supset \Omega^{k+1} \supset \dots \supset \Omega^n = \Gamma$, $\forall k \geq 0$.
3. Como $0 \leq y \leq e$ e $\|a_{2m+k}\| = 1$, $\forall k \geq 1$. Então, para todo $k \geq 1$,

$$\begin{aligned} A^k(S^k)^{-2}(A^k)^t &= (Y^k)^{-2} + (I - Y^k)^{-2} + \sum_{i=2m+1}^{2m+k} \frac{a_i a_i^t}{(s_i^k)^2} \\ &\geq (Y^k)^{-2} + (I - Y^k)^{-2} \\ &\geq 8I. \end{aligned}$$

De fato, pois cada elemento da diagonal de $(Y^k)^{-2} + (I - Y^k)^{-2}$ é do tipo $y_i^{-2} + (1 - y_i)^{-2}$, fazendo $f(y_i) = y_i^{-2} + (1 - y_i)^{-2}$, $\forall y_i \in (0, 1)$, como

$$f(y_i) \rightarrow \infty \text{ quando } y_i \rightarrow 0$$

e

$$f(y_i) \rightarrow \infty \text{ quando } y_i \rightarrow 1,$$

então f assume um valor mínimo em $(0, 1)$,

$$f'(y_i) = -2y_i^{-3} + 2(1 - y_i)^{-3} = 0$$

donde

$$\left(y_i - \frac{1}{2}\right) (2y_i^2 - 2y_i + 2) = 0$$

o que implica $y_i = \frac{1}{2}$ e $f(y_i) = 8$.

Portanto, obtemos

$$r^k \leq \sqrt{\frac{a_{2m+k+1}^t a_{2m+k+1}}{8}} = \sqrt{\frac{1}{8}} = \frac{\sqrt{2}}{4}, \forall k \geq 0.$$

b) Caso Infinito

Suporemos Γ convexo e contido numa caixa $\{y \in \mathbb{R}^m; 0 \leq y \leq e\}$.

Algoritmo 6.3.2

Inicialização

Sejam

$$A^0 = (I \quad -I) \in \mathbb{R}^{m \times 2m}, c^0 = \begin{pmatrix} e \\ 0 \end{pmatrix} \in \mathbb{R}^{2m},$$

$$y^0 = \frac{e}{2} \in \mathbb{R}^m, s^0 = \frac{e}{2} \in \mathbb{R}^{2m} \text{ e } x^0 = 2e \in \mathbb{R}^{2m}.$$

O ponto y^0 é o centro analítico de

$$\Omega^k = \{y \in \mathbb{R}^m; (A^k)^t y \leq c^k\}$$

para $k = 0$.

REPITA

Gerar um corte $a_{2m+k+1}^t y \leq a_{2m+k+1}^t y^k + \gamma$ com $\|a_{2m+k+1}\| = 1$ e $\gamma \leq 0$. Fazer

$$\Omega^{k+1} = \{y \in \mathbb{R}^m; (A^{k+1})^t y \leq c^{k+1}\}$$

$$\text{onde } A^{k+1} = (A^k \quad a_{2m+k+1}) \text{ e } c^{k+1} = \begin{pmatrix} c^k \\ a_{2m+k+1}^t y^k + \gamma \end{pmatrix}.$$

Calcular um centro analítico δ -aproximado $s^{k+1} \in \mathcal{S}_{\Omega_{k+1}}$ usando o algoritmo 5.4.2. Calcular y^{k+1} . Fazer $k = k + 1$

ATÉ $y^k \in \Gamma$.

Observações 6.3.2

A função potencial no centro analítico \bar{y}^k de Ω^k é dado por

$$P(\Omega^k) = - \sum_{i=1}^{2m+k} \ln(c_i^k - (a_i^k)^t \bar{y}^k) = p(\bar{s}^k).$$

Enquanto a regra de parada não for satisfeita, temos

1. O número total de restrições em Ω^k é de no máximo $2m + k$, $\forall k \geq 0$.
2. $\Gamma \subset \Omega^k$, $\forall k \geq 0$.
3. Da proposição 6.2.1, temos que

$$P(\Omega^{k+1}) \geq P(\Omega^k) - \frac{1}{2} \ln (\bar{r}^k)^2 + j^k \sigma + \xi, \forall k \geq 0,$$

$$\text{onde } (\bar{r}^k)^2 = a_{2m+k+1} (A^k(\bar{S}^k)^{-2} (A^k)^t)^{-1} a_{2m+k+1}.$$

6.4 A Análise de Convergência

a) Caso Finito

Proposição 6.4.1 *Seja Γ dado por n desigualdades lineares. Se Γ contém uma bola de raio $\varepsilon > 0$, então*

$$P(\Omega^k) \leq (2m + n) \ln \frac{1}{\varepsilon}, \forall k \geq 0$$

Demonstração. Como $\Gamma \subset \Omega^k$, $\forall k \geq 0$, então Ω^k contém uma bola de raio $\varepsilon > 0$. Seja \bar{y} o centro desta bola. Então,

$$\begin{aligned} P(\Omega^k) &= - \sum_{i=1}^{2m+k} \ln \bar{s}_i^k = - \sum_{i=1}^{2m+k} \ln(c_i^k - a_i^t \bar{y}^k) \leq \\ &\leq - \sum_{i=1}^{2m+k} \ln(c_i^k - a_i^t \bar{y}) \leq - \sum_{i=1}^{2m+k} \ln \varepsilon = (2m+k) \ln \frac{1}{\varepsilon} \end{aligned}$$

Portanto, como $k \leq n$, $\forall k \geq 1$, segue que

$$P(\Omega^k) \leq (2m+n) \ln \frac{1}{\varepsilon}, \forall k \geq 0$$

□

Proposição 6.4.2 *Seja Γ definido por n desigualdades lineares. Então o algoritmo 6.3.1 acha um ponto viável após no máximo n cortes.*

Demonstração. No algoritmo de plano de corte 6.3.1 cada restrição violada é colocada na sua verdadeira posição num único passo. Como Γ está definido por n desigualdades lineares, então após n cortes teremos que $\Omega^n = \Gamma$. □

Proposição 6.4.3 *Se Γ é definido por n desigualdades lineares e contém uma bola de raio $\varepsilon > 0$, então o número total de passos de Newton no algoritmo 6.3.1 é de no máximo $O(n \ln \frac{1}{\varepsilon})$.*

Demonstração. A proposição 6.2.1 nos dá uma variação do potencial quando um corte é adicionado a um politopo

$$P(\Omega^{k+1}) \geq P(\Omega^k) - \ln r^k + j^k \sigma + \xi, \forall k \geq 0.$$

Como $r^k \leq \frac{\sqrt{2}}{4}$, $\forall k \geq 0$, então

$$P(\Omega^{k+1}) \geq P(\Omega^k) + 1 + j^k \sigma + \xi$$

Usando a recursividade, temos

$$P(\Omega^{k+1}) \geq P(\Omega^0) + \sum_{i=0}^k 1 + \sigma \sum_{i=0}^k j^i,$$

mas da proposição 6.4.1, o potencial está limitado superiormente, logo temos

$$(2m + n) \ln \frac{1}{\varepsilon} \geq P(\Omega^{k+1}) \geq P(\Omega^0) + k + \sigma \sum_{i=0}^k j^i.$$

Portanto, o número de passos de Newton é de no máximo $O(n \ln \frac{1}{\varepsilon})$. \square

Caso Infinito

Proposição 6.4.4 *Seja $\Gamma \subset \mathbb{R}^m$ um convexo. Se Γ contém uma bola de raio $\varepsilon > 0$, então*

$$P(\Omega^k) \leq (2m + k) \ln \frac{1}{\varepsilon}, \forall k \geq 0.$$

Demonstração. é análoga à proposição 6.4.1. \square

Proposição 6.4.5 *Seja $s = c^k - (A^k)^t y, \forall y \in \Omega^k$. Então*

$$i) 0 \leq s_i \leq 1, i = 1, \dots, 2m$$

$$ii) 0 \leq s_i \leq \sqrt{m}; i = 2m + 1, \dots, 2m + k.$$

Demonstração. Seja $y \in \Omega^k$, então $0 \leq y \leq e$, donde $y + s = e$ e $-y + s = 0$.

Assim,

$$0 \leq s_i = 1 - y_i, \text{ para } i = 1, \dots, m,$$

$$0 \leq s_i = y_i \leq 1, \text{ para } i = m + 1, \dots, 2m,$$

e para $i = 2m + 1, \dots, 2m + k$, temos

$$0 \leq s_i \leq a_i^t y^i - a^t y \leq \|a_i\| \|y^i - y\| = \|y^i - y\| \leq \sqrt{m}.$$

□

Proposição 6.4.6 *Seja $s = c^k - (A^k)^t y$, $\forall y \in \Omega^k$ e sejam*

$$B^0 = 8I$$

$$B^{k+1} = B^k + \frac{a_{2m+k+1} a_{2m+k+1}^t}{m}.$$

Então,

$$A^k (S^k)^{-2} (A^k)^t - B^k$$

é semidefinida positiva.

Demonstração. Como $A^0 = (I - I)$ e

$$s_i = 1 - y_i \quad , \quad i = 1, \dots, m$$

$$s_i = y_i \quad , \quad i = m + 1, \dots, 2m.$$

Então,

$$\begin{aligned} A^k (S^k)^{-2} (A^k)^t &= (Y^k)^{-2} + (I - Y^k)^{-2} + \sum_{i=1}^k \frac{a_{2m+i} a_{2m+i}^t}{(s_{2m+i}^k)^2} \\ &\geq (Y^k)^{-2} + (I - Y^k)^{-2} + \frac{1}{m} \sum_{i=1}^k a_{2m+i} a_{2m+i}^t \\ &\geq 8I + \frac{1}{m} \sum_{i=1}^k a_{2m+i} a_{2m+i}^t \\ &= B^k. \end{aligned}$$

As duas últimas desigualdades seguem da proposição 6.4.5 e das observações 6.3.1.

□

Proposição 6.4.7 *Sejam $s^k = c^k - (A^k)^t y^k$ as folgas no centro analítico $y^k \in \Omega^k$ e $(\omega^k)^2 = a_{2m+k+1}^t B_k^{-1} a_{2m+k+1}$, então*

$$(\omega^k)^2 \geq a_{2m+k+1}^t (A^k (S^k)^{-2} (A^k)^t)^{-1} a_{2m+k+1} = (r^k)^2.$$

Demonstração. É imediata da proposição 6.4.6. □

A próxima proposição é uma variante de um resultado de NESTEROV[35] e foi demonstrada em GOFFIN *et al.*[13].

Proposição 6.4.8

$$\sum_{j=0}^k (\omega_j)^2 \leq 2m^2 \ln\left(1 + \frac{k+1}{8m^2}\right)$$

Demonstração. Como $B^{k+1} = B^k + \frac{1}{m} a_{2m+(k+1)} a_{2m+(k+1)}^t$, então

$$\begin{aligned} \det B^{k+1} &= \det\left(B^k + \frac{1}{m} a_{2m+(k+1)} a_{2m+(k+1)}^t\right) \\ &= \det B^k \det\left(I + \frac{B_k^{-1}}{m} a_{2m+(k+1)} a_{2m+(k+1)}^t\right) \\ &= \det B^k \det\left(I + \frac{B_k^{-1}}{m} a_{2m+(k+1)} a_{2m+(k+1)}^t\right) \\ &= \left(1 + \frac{a_{2m+(k+1)}^t B_k^{-1} a_{2m+(k+1)}}{m}\right) \det B^k \\ &= \left(1 + \frac{(\omega^k)^2}{m}\right) \det B^k. \end{aligned}$$

A penúltima igualdade segue da proposição 2.1.2. Logo,

$$\ln \det B^{k+1} = \ln \det B^k + \ln\left(1 + \frac{(\omega^k)^2}{m}\right).$$

Mas, da proposição 6.4.6, $B^k \geq 8I, \forall k \geq 0$, donde

$$\frac{(\omega^k)^2}{m} \leq a_{2m+(k+1)}^t (B^k)^{-1} a_{2m+(k+1)} \leq \frac{1}{8},$$

e, lembrando da desigualdade

$$\ln(1 + \lambda) \geq \lambda - \frac{\lambda^2}{2(1 - |\lambda|)}, \lambda \in (-1, 1),$$

que foi provada por KARMARKAR[20], temos

$$\begin{aligned} \ln\left(1 + \frac{(\omega^k)^2}{m}\right) &\geq \frac{(\omega^k)^2}{m} - \frac{\left(\frac{(\omega^k)^2}{m}\right)^2}{2\left(1 - \frac{(\omega^k)^2}{m}\right)} \\ &= \frac{(\omega^k)^2}{m} \left(\frac{2 - 3\frac{(\omega^k)^2}{m}}{2\left(1 - \frac{(\omega^k)^2}{m}\right)}\right) \\ &\geq \frac{(\omega^k)^2}{2m}. \end{aligned}$$

Então,

$$\ln \det B^{k+1} \geq \ln \det B^k + \frac{(\omega^k)^2}{2m}, \forall k \geq 0.$$

Agora, usando a recursividade, segue que

$$\begin{aligned} \ln \det B^{k+1} &\geq \ln \det B^0 + \sum_{j=0}^{(k+1)} \frac{(\omega^j)^2}{2m} \\ &= m \ln 8 + \sum_{j=0}^{(k+1)} \frac{(\omega^j)^2}{2m}. \end{aligned}$$

Como

$$\det B^{k+1} = \prod_{i=1}^m \lambda_{i, \text{MGA}} \leq \left(\frac{\sum_{i=1}^m \lambda_i}{m}\right)^m,$$

onde os $\lambda_{i,s}$ são os autovalores de B^{k+1} . Então, da proposição 2.1.1, segue que

$$\begin{aligned} \det B^{k+1} &\leq \left(\frac{\text{tr } B^{k+1}}{m}\right)^m \\ &= \left(\frac{8m + \sum_{i=1}^{(k+1)} a_{2m+i}^t a_{2m+i}}{m}\right)^m \\ &= \left(\frac{8m + \frac{(k+1)}{m}}{m}\right)^m \\ &= \left(8 + \frac{\binom{k+1}{m}}{m^2}\right)^m. \end{aligned}$$

Portanto,

$$m \ln 8 + \sum_{j=0}^{k+1} \frac{(\omega_j)^2}{2m} \leq \ln \det B^{k+1} \leq m \ln \left(8 + \frac{(k+1)}{m^2}\right),$$

donde

$$\sum_{j=0}^{(k+1)} \frac{(\omega_j)^2}{2m} \leq m \ln 8 \left(1 + \frac{(k+1)}{8m^2}\right) - m \ln 8,$$

o que implica

$$\sum_{j=0}^{(k+1)} (\omega_j)^2 \leq 2m^2 \ln \left(1 + \frac{(k+1)}{8m^2}\right).$$

□

Proposição 6.4.9 *Se Γ contém uma bola de raio $\varepsilon > 0$, então o algoritmo 6.3.2 pára quando k satisfaz*

$$\frac{\varepsilon^2}{m} \geq \frac{\frac{1}{2} + 2m \ln \left(1 + \frac{k+1}{8m^2}\right)}{2m + k + 1} e^{-\frac{2(k+1)\xi}{2m+k+1} - \frac{2\sigma}{2m+k+1} \sum_{i=0}^k j^i}.$$

Demonstração. Das observações 6.3.2 e da proposição 6.4.1, temos que

$$\begin{aligned} (2m + k + 1) \ln \frac{1}{\varepsilon} &\geq P(\Omega^{k+1}) \\ &\geq \dots \\ &\geq P(\Omega^0) - \frac{1}{2} \sum_{i=0}^k \ln (r^i)^2 + \sigma \sum_{i=0}^k j^i + (k+1)\xi, \end{aligned}$$

onde $\sigma = 0.001 + \ln(1 + \frac{\nu}{\sqrt{n}})$ e $\xi = \ln \alpha^1 + 0.026 < 0$. Então

$$(2m + k + 1) \ln \varepsilon + (k+1)\xi + \sigma \sum_{i=0}^k j^i \leq 2m \ln \frac{1}{2} + \frac{1}{2} \sum_{i=0}^k \ln (r^i)^2,$$

donde

$$\begin{aligned}
\ln \varepsilon + \frac{k+1}{2m+k+1} \xi + \frac{\sigma}{2(2m+k+1)} \sum_{i=0}^k j^i &\leq \frac{1}{2} \left(\frac{2m \ln \frac{1}{4} + \sum_{i=0}^k \ln(r^i)^2}{2m+k+1} \right) \\
&\leq \frac{1}{2} \ln \left[\left(\frac{1}{4} \right)^{2m} \prod_{i=0}^k r_i^2 \right]^{\frac{1}{2m+(k+1)}} \\
&\leq \frac{1}{2} \ln \left(\frac{\frac{2m}{4} + \sum_{i=0}^k r_i^2}{2m+k+1} \right)
\end{aligned}$$

o que implica

$$\ln(\varepsilon^2 \cdot e^{\frac{2(k+1)\xi}{2m+k+1} + \frac{2\sigma}{2m+k+1} \sum_{i=0}^k j^i}) \leq \ln \frac{\frac{m}{2} + \sum_{i=0}^k r_i^2}{2m+k+1},$$

eliminando os logaritmos, temos

$$\varepsilon^2 \leq \frac{\frac{m}{2} + \sum_{i=0}^k r_i^2}{2m+k+1} e^{\frac{-2(k+1)\xi}{2m+k+1} - \frac{2\sigma}{2m+k+1} \sum_{i=0}^k j^i},$$

usando as proposições 6.4.7 e 6.4.8, segue

$$\frac{\varepsilon^2}{m} \leq \frac{\frac{1}{2} + 2m \ln \left(1 + \frac{k+1}{8m^2} \right)}{2m+k+1} e^{\frac{-2(k+1)\xi}{2m+k+1} - \frac{2\sigma}{2m+k+1} \sum_{i=0}^k j^i}.$$

Como o termo exponencial é limitado e a seqüência $\frac{\ln \left(1 + \frac{k+1}{8m^2} \right)}{2m+k+1} \rightarrow 0$ quando $k \rightarrow \infty$, então existe um $\bar{k} \in \mathbb{N}$ tal que $\forall k \geq \bar{k}$, temos

$$\frac{\varepsilon^2}{m} \geq \frac{\frac{1}{2} + 2m \ln \left(1 + \frac{k+1}{8m^2} \right)}{2m+k+1} e^{\frac{-2(k+1)\xi}{2m+k+1} - \frac{2\sigma}{2m+k+1} \sum_{i=0}^k j^i}.$$

□

A proposição 6.4.9 implica que a complexidade do algoritmo 6.3.2 quanto ao número total de cortes é de $O^*\left(\frac{m^2}{\varepsilon^2}\right)$. A notação O^* significa que os termos de ordem inferior foram ignorados.

Proposição 6.4.10 *O número total de passos de Newton no algoritmo 6.3.2 é de no máximo*

$$O^*\left(\frac{m^2}{\varepsilon^2} \ln \frac{1}{\varepsilon}\right).$$

Demonstração. Da proposição anterior, temos que o algoritmo 6.3.2 pára quando um \bar{k} satisfaz

$$\frac{\varepsilon^2}{m} \geq \frac{\frac{1}{2} + 2m \ln\left(1 + \frac{\bar{k}+1}{8m^2}\right)}{2m + \bar{k} + 1} e^{\frac{-2(\bar{k}+1)\xi}{2m+\bar{k}+1} - \frac{2\sigma}{2m+\bar{k}+1} \sum_{i=0}^{\bar{k}} j^i}.$$

Como $\sigma > 0$, $\xi < 0$ e

$$\begin{aligned} \frac{\frac{1}{2} + 2m \ln\left(1 + \frac{\bar{k}+1}{8m^2}\right)}{2m + \bar{k} + 1} &\leq \frac{1}{2(2m + \bar{k} + 1)} + \frac{2m \left(\frac{\bar{k}+1}{8m^2}\right)}{2m + \bar{k} + 1} \\ &= \frac{1}{2(2m + \bar{k} + 1)} + \frac{\bar{k} + 1}{4m(2m + \bar{k} + 1)} \\ &= \frac{2m + \bar{k} + 1}{4m(2m + \bar{k} + 1)} \\ &< 1. \end{aligned}$$

Então, temos que a desigualdade também é satisfeita se

$$\frac{\varepsilon^2}{m} \geq e^{\frac{-2\xi(\bar{k}+1)}{2m+\bar{k}+1} - \frac{2\sigma}{2m+\bar{k}+1} \sum_{i=0}^{\bar{k}} j^i}.$$

Tomando o logaritmo em ambos os lados,

$$2 \ln \varepsilon - \ln m \geq \frac{-2\xi(\bar{k} + 1)}{2m + \bar{k} + 1} - \frac{2\sigma}{2m + \bar{k} + 1} \sum_{i=0}^{\bar{k}} j^i,$$

donde

$$\sum_{i=0}^{\bar{k}} j^i \geq \left(\frac{2m + \bar{k} + 1}{\sigma}\right) \ln \frac{1}{\varepsilon} + \frac{2m + \bar{k} + 1}{2\sigma} \ln m - \frac{\xi}{\sigma}(\bar{k} + 1).$$

Portanto, o número total de passos de Newton é de no máximo $O^*\left(\frac{m^2}{\varepsilon^2} \ln \frac{1}{\varepsilon}\right)$. \square

Capítulo 7

Comentários Finais

A nossa contribuição mais importante neste trabalho foi apresentar uma nova maneira de introduzir um corte em um politopo dado no formato dual e o desenvolvimento de um algoritmo para calcular um centro analítico aproximado do politopo obtido após a adição de um corte profundo. Com este procedimento desenvolvemos um algoritmo de plano de corte para o problema de viabilidade convexa.

GOFFIN e VIAL em [9] desenvolveram um algoritmo primal para calcular um centro analítico aproximado do politopo obtido após a adição de um corte profundo, onde eles fazem uso da direção de MITCHELL e TODD [30] para recuperar a viabilidade e passos de Newton primal com busca para centralizar.

A novidade em nosso trabalho é que partindo de um centro analítico aproximado, que é conhecido, usamos a trajetória central de um problema de programação linear auxiliar no formato dual para recuperação da viabilidade. O algoritmo usado para seguir a trajetória central pode ser qualquer algoritmo de trajetória central.

A análise de complexidade do nosso algoritmo de plano de corte, usando o procedimento dual que segue a trajetória central com passos curtos para obter um centro analítico aproximado, é a mesma de GOFFIN e VIAL[9]. Na prática os algoritmos de trajetória central primal-dual de passo longo têm apresentado um desempenho muito bom então, acreditamos que o nosso algoritmo de plano de corte usando um algoritmo de trajetória central primal-dual de passo longo terá um desempenho muito melhor que o resultado teórico obtido usando passos curtos.

Bibliografia

- [1] ATKINSON, D.S., VAIDYA, P.M. “A Cutting Plane Algorithm that uses Analytic Centers”, *Mathematical Programming*, n.69, pp.1-43, 1995.
- [2] BAHN, O., GOFFIN, J.-L., VIAL, J.-P. *et al.* “Experimental Behaviour of an Interior Point Cutting Plane Algorithm for Convex Programming: An Application to Geometric Programming”, *Discrete Applied Mathematics*, n.49, pp.3-23, 1994.
- [3] BAHN, O., MERLE, O. du, GOFFIN, J.-L. *et al.* “A Cutting Plane Method from Analytic Centers for Stochastic Programming”, *Mathematical Programming*, n.69, pp.45-73, 1995.
- [4] DIKIN, I.I., “Iterative Solution of Problems of Linear and Quadratic Programming”, *Doklady Akademii Nauk USSR*, 174. Translated in Soviet Mathematics Doklady, 8, pp. 674-675, 1967.
- [5] ELZINGA, J., MOORE, T. “A Central Cutting Plane Algorithm for Convex Programming”, *Mathematical Programming*, n.8, pp.134-145, 1975.
- [6] FRISCH, K.R. *The Logarithmic Potential Method of Convex Programming*. Memorandum, University Institute of Economics, Oslo, Norway, 1955.
- [7] GILL, .E., MURRAY, W., WRIGHT, M.H. *Numerical Linear Algebra and Optimization*. 1 ed. Redwood City, Ca, Addison-Wesley Publishing Company, vol. 1, 1991.

- [8] GOFFIN, J.-L., VIAL, J.P. “Cutting Planes and Column Generation Techniques with the Projective Algorithm”, *Journal of Optimization Theory and its Applications*, n.65, pp.409-429, 1989.
- [9] GOFFIN, J.-L., VIAL, J.P. *Shallow, Deep and very Deep Cuts in the Analytic Center Cutting Plane Method*. Logilab Technical Report 96-3, Department of Management Studies, University of Geneva, Switzerland, 1996.
- [10] GOFFIN, J.L., VIAL, J.P. *A Two-Cut Approach in the Analytic Center Cutting Plane Method*. Logilab Technical Report 97-6, Department of Management Studies, University of Geneva, Switzerland, 1997.
- [11] GOFFIN, J.L., GONDZIO, J., SARKISSIAM, R. *et al.* “Solving Nonlinear Multicommodity Flows Problems by the Analytic Centers Cutting Plane Method”, *Mathematical Programming*, v.76, n.1, pp.131-154, 1997.
- [12] GOFFIN, J.L., HAURIE, A., VIAL, J.P. “Decomposition and Nondifferentiable Optimization with the Projective Algorithm”, *Management Science*, n.38, pp.284-302, 1992.
- [13] GOFFIN, J.L., LUO, Z.Q., YE, Y. “Complexity Analysis of an Interior Cutting Plane for Convex Feasibility Problems”, *SIAM Journal of Optimization*, n.6, pp.638-652, 1996.
- [14] GONZAGA, C.C. “Interior Point Algorithms for Linear Programming with Inequality Constraints”, *Mathematical Programming*, n.52, pp.209-255, 1991.
- [15] GONZAGA, C.C. “Path-Following Methods for Linear Programming”, *SIAM Review*, v.34, n.2, pp.167-224, 1992.
- [16] GONZAGA, C.C. *Interior Point Methods: Algorithms and Complexity*. Minicurso, Depto. de Matemática, UFSC, 1997.

- [17] GONZAGA, C.C., *Algoritmo de Pontos Interiores para Programação Linear*, 17^o Colóquio Brasileiro de Matemática, Impa, 1989.
- [18] GONZAGA, C.C., “The Largest Step Path Following Algorithm for Monotone Linear Complementarity Problems”, *Mathematical Programming*, n. 76, pp. 309-332, 1997.
- [19] HUARD, P. “Resolution of Mathematical Programming with Nonlinear Constraints by the Method of Centers”. In: *Nonlinear Programming*, North-Holland, Amsterdam, J. Abadie, ed., 1967.
- [20] KARMAKAR, N. “A New Polynomial Time Algorithm for Linear Programming”, *Combinatoria*, n.4, pp.373-395, 1984.
- [21] KELLEY, J.E. “The Cutting-Plane Method for Solving Convex Programs”. *Journal of the SIAM*, n.8, 1960.
- [22] KHACHIYAN, L.G. “A Polynomial Algorithm for Linear Programming”, *Soviet. Math. Doklady*, n.20, pp.191-194, 1979.
- [23] KHACHIYAN, L.G., TODD, M.J. “On the Complexity of Approximating the Maximal Inscribed Ellipsoid for a Polytope”, *Mathematical Programming*, n.61, pp.137-160, 1993.
- [24] LEVIN, A. “An Algorithm of Minimization of Convex Functions”. *Soviet. Math. Doklady*, v.160, n.6, pp.1244-1247, 1965.
- [25] LUENBERGER, D.G. *Linear and Nonlinear Programming*, 2a. ed., Addison-Wesley, Reading, MA, 1984.
- [26] MEHROTRA, S., “On the Implementation of a Primal-Dual Interior Point Method”, *SIAM Journal on Optimization*, n. 2, 1992.
- [27] MITCHELL, J.E. “An Interior Point Column Generation Method for Linear Programming using Shifted Barriers”, *SIAM Journal of Optimization*, v.4, n.2, pp.423-440, 1994.

- [28] MITCHELL, J.E. “Interior Point Algorithm for Integer Programming”. In: J.E. Brasley, editor, *Advances in Linear and Integer Programming*, Oxford University Press, pp.223-248, 1996.
- [29] MITCHELL, J.E. “Interior Point Methods for Combinatorial Optimization”. In: Tamás Terlaky, editor, *Interior Point Methods in Mathematical Programming*, Kluwer Academic Publishers, pp.417-466, 1996.
- [30] MITCHELL, J.E., TODD, M.J. “Solving Combinatorial Optimization Problems using Karmarkar’s Algorithms”, *Mathematical Programming*, n.56, pp.245-284, 1992.
- [31] MIZUNO, S. “A New Polynomial Time Method for a Linear Complementarity Problem”. *Mathematical Programming*, n.56, 1992.
- [32] MIZUNO, S., TODD, M.J., YE, Y., “On Adaptive-Step Primal-Dual Interior-Point Algorithms for Linear Programming”, *Mathematics of Operations Research*, vol. 18, n.4, 1993.
- [33] MOKHTARIAN, F.S., GOFFIN, J.L. *Using the Primal-Dual Infeasible Newton Method in the Analytic Centers Method for Problems Defined by Deep Cutting Planes*. Cahier du GERAD G 94-41, ISSN: 0711-2440, 1994.
- [34] NEMIROVSKY, A.S., YUDIN, D.B. *Problem Complexity and Method Efficiency in Optimization*, Chichester, John Willey and Sons, 1983.
- [35] NESTEROV, Y. “Cutting Plane Algorithms from Analytic Centers: Efficiency Estimates”, *Mathematical Programming*, n.69, pp.149-176, 1995.
- [36] NESTEROV, Y., PÉTON, O., VIAL, J.P., *Homogeneous Analytic Center Cutting Plane Methods with Approximate Centers*. Logilab Technical Report 98, Department of Management Studies, University of Geneva, Switzerland, 1998.

- [37] NESTEROV, Y., VIAL, J.P. *Homogeneous Analytic Center Cutting Plane Methods for Convex Problems and Variational Inequalities*. Logilab Technical Report 97-4, Department of Management Studies, University of Geneva, Switzerland, 1997.
- [38] NOBLE, B., DANIEL, J.W. *Applied Linear Algebra*. Prentice-Hall, 1988.
- [39] RENEGAR, J. "A Polynomial-Time Algorithm based on Newton's Method for Linear Programming", *Mathematical Programming*, n.40, 1988.
- [40] ROOS, C., TERLAKY, T., VIAL, J.P. *Theory and Algorithms for Linear Optimization: An Interior Point Approach*, Chichester, John Wiley and Sons, 1997.
- [41] SHOR, N. "A Cutting Plane Method for Solving Convex Programming Problems" *Cybernetics*, n.1, pp.42-50, 1977.
- [42] SONNEVEND, G. "An Analytical Centre for Polyhedron and New Class of Global Algorithms for Linear (smooth convex) Programming". In: *Lectures Notes in Control and Information Sciences*, n.84, New York, Springer Verlag, 1985.
- [43] SONNEVEND, G. "New Algorithms in Convex Programming based on a notion of 'Centre' (for Systems of Analytic Inequalities) and on Rational Extrapolation". In: *Trends in Mathematical Optimization*, ISNM, 84, Basel, Birkuser Verlag, 1988.
- [44] STRANG, G. *Linear Algebra and its Applications*. 3 ed., USA, Harcourt Brace Jovanovich, Inc., 1988.
- [45] TARASOV, S., KHACHIYAN, L.G.; ERLICH, I. "The Method of Inscribed Ellipsoids", *Soviet. Math. Doklady*, n.37, 1988.
- [46] VAIDYA, P.M. "A New Algorithm for Minimizing a Convex Function over Convex Sets". In: *Proceedings of the 30th Annual Symposium of Foun-*

- dations of Computer Science*, Research Triangle Park, NC, USA, pp.338-343, IEEE Computer Society Press, Los Alamitos, CA, 1990.
- [47] VAIDYA, P.M. “A Locally Well-Behaved Potential Function and a simple Newton-type Method for Finding the Center of a Polytope”. In: *Progress in Mathematical Programming - Interior Point and Related Methods*, N. Megiddo, ed., Springer-Verlag, Berlin, Chap. 5, 1989.
- [48] VIAL, J.P. *A Generic Path-Following Algorithm with a Sliding Constraint and its Applications to Linear Programming and the Computation of Analytic Centers*. Logilab Technical Report 96-8, Department of Management Studies, University of Geneva, Switzerland, 1996.
- [49] YE, Y. “A Potential Reduction Algorithm allowing Column Generation”, *SIAM Journal of Optimization*, n.2, pp.7-20, 1992.
- [50] YUDIN, D., NEMIROVSKY, A. “Informational Complexity and Efficient Methods for Solving Convex Extremal Problems”, *Ekonomika i matem. metody*, v.12, n.2, pp.357-369, 1976.