

ALGORITMOS DE PONTOS INTERIORES PARA PROGRAMAÇÃO LINEAR

COMBINANDO FASE 1 E FASE 2

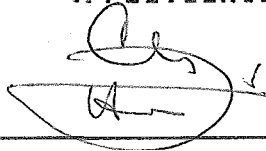
Marco Antonio Figueiredo Menezes

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Aprovada por:



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



Prof. Carlos Humes Júnior, Ph.D.



Prof. Paulo Roberto Oliveira, D.Ing.

Rio de Janeiro, RJ - Brasil

Junho/1991

Menezes, Marco Antonio Figueiredo

Algoritmos de Pontos Interiores para Programação
Linear Combinando Fase 1 e Fase 2
7, 126p. 29,7 cm COPPE/UFRJ, M.Sc.,
Programa de Engenharia de Sistemas e Computação
1991

Tese - Universidade Federal do Rio de
Janeiro, COPPE.

1. Programação Linear
2. Métodos de Pontos Interiores
3. Complexidade

I. COPPE/UFRJ. II. Título (série).

III

Para Guiomar†, Manoel†,
Sélio, Sidney, Cláudia e
Polange.

† em memória

AGRADECIMENTOS

Como não tenho a pretensão de escrever um livro nesta seção exclusiva aos agradecimentos, tentarei ser sintético.

Agradeço a todo o povo latino-americano. Em particular, ao meu orientador Clóvis Caesar Gonzaga; aos dois membros da Banca, Paulo Roberto Oliveira e Carlos Humes Júnior; e às digitadoras Denise Mina Firmiano e Marilza José Murta, cujos esforços estão explicitados nesta Tese.

Também, agradeço à existência da COPPE/UFRJ, IMPA e IPRJ pelas suas contribuições ao desenvolvimento da ciência no Brasil. Em particular, ao nosso grupo de Pontos Interiores.

Finalmente, peço desculpas ao meu orientador Clóvis Caesar Gonzaga para quem não tenho palavras para expressar meus agradecimentos por tudo que vivemos juntos e o quanto aprendi. Confesso que às vezes penso na palavra amiga, noutras, na palavra pai....

Marco Antonio Figueiredo Menezes

Resumo da tese apresentada à COPPE/UFRJ, como parte dos requisitos necessários para a obtenção do grau de mestre em ciências (M.Sc.).

ALGORITMOS DE PONTOS INTERIORES PARA PROGRAMAÇÃO LINEAR

COMBINANDO FASE 1 E FASE 2

Marco Antonio Figueiredo Menezes

Junho, 1991.

Orientador: Clóvis Caesar Gonzaga

Programa : Engenharia de Sistemas e Computação

Neste trabalho construímos dois algoritmos primais de pontos interiores para programação linear combinando fase 1 (viabilidade) e fase 2 (otimalidade): um algoritmo afim e outro projetivo. Isto evita o uso de uma técnica "big M", que introduz dificuldades numéricas para encontrar um ponto viável inicial para métodos de pontos interiores. Ambos os procedimentos para determinar direções de busca e a busca unidirecional, são baseados sobre uma estratégia *min-max*. Possibilitamos passos longos na busca unidirecional, o que difere dos métodos existentes na literatura. Finalmente, mostramos uma complexidade polinomial de $O(nL)$ iterações para ambos algoritmos.

Abstract of thesis presented to COPPE/UFRJ, as partial fulfillment of the requirements for the degree of master of science (M.Sc.).

INTERIOR POINT METHODS FOR LINEAR PROGRAMMING

COMBINING PHASE 1 AND PHASE 2

Marco Antonio Figueiredo Menezes

June, 1991

Thesis Supervisor: Clóvis Caesar Gonzaga

Department = Systems Engineering and Computer Sciences

In this work we propose two primal interior point algorithms for linear programming combining phase 1 (feasibility) and phase 2 (optimality): an affine and a projective algorithm. This avoids the usage of a "big M" technique, which introduces numerical difficulties in finding an initial feasible point for interior point methods. Both the procedure for determining search directions and the line searches are based on a *min-max* strategy. We allow large steps in the line searches, which differs from the methods existing in the literature. Finally we show a polynomial complexity of $O(nL)$ iterations for both algorithms.

INDICE

	Página
CAPÍTULO 1 - INTRODUÇÃO.....	1
CAPÍTULO 2 - ALGORITMOS DE PONTOS INTERIORES EM PROGRAMAÇÃO LINEAR: UMA INTRODUÇÃO	5
2.1 - Notação e Ferramentas	5
2.2 - Funções potenciais e mudança de escala	11
2.3 - O algoritmo de Karmarkar	18
CAPÍTULO 3 - OS PROBLEMAS	25
3.1 - Formulações	25
3.2 - Os algoritmos	27
3.3 - Alguns resultados	
CAPÍTULO 4 - LIMITANTES INFERIORES	45
CAPÍTULO 5 - DIREÇÕES	51
5.1 - Uma abordagem afim	51
5.2 - Uma abordagem projetiva	--
5.3 - Equivalência sob certa condição..	98

	Página
CAPÍTULO 6 - BUSCA UNIDIRECIONAL	105
6.1 - Uma abordagem afim	106
6.2 - Uma abordagem projetiva	113
CAPÍTULO 7 - CONVERGÊNCIA	119
REFERÊNCIAS BIBLIOGRÁFICAS.....	124

CAPITULO 1

INTRODUÇÃO

Estudamos a resolução, através de algoritmos de pontos interiores, de um problema de programação linear primal

$$\begin{aligned} & \text{minimizar } c'x \\ & \text{sujeito a: } Ax = b \\ & \quad \quad \quad x \geq 0, \end{aligned}$$

onde $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, A é uma matriz $m \times n$ de rank máximo $\min\{m, n\}$.

Algoritmos de pontos interiores para programação linear partem normalmente de um ponto viável inicial, que deve ser fornecido por algum procedimento de inicialização. O procedimento mais usual parte de um ponto $x^0 \geq 0$ não viável, introduz uma nova variável γ e modifica a restrição para

$$Ax + \gamma(b - Ax^0) = b.$$

O ponto

$$(x^0, 1) \in \mathbb{R}^{n+1}$$

é viável para este problema transformado e deve-se reduzir γ a zero para obter um ponto viável ao problema original.

Pode-se reduzir γ associando a esta variável um custo muito alto M (método do "big M "), o que força γ a diminuir rapidamente com qualquer algoritmo de minimização. Este método é sujeito a problemas numéricos, a obtenção de M deselegante, além de gastar algumas iterações.

Em nosso estudo, procuramos reduzir simultaneamente a variável γ e o critério original $c'x$, como em um problema de critérios múltiplos, sem associar custo à variável γ . Este tipo de algoritmo é conhecido como de "fase 1 - fase 2", e já foi desenvolvido por Anstreicher [2, 3] e Todd [12, 14] - veja também Todd e Wang [13] -, para algoritmos primais do tipo projetivo e afim. De Ghellinck e Vial [5] também propuseram um algoritmo primal do tipo projetivo que não requer viabilidade.

Inspiramo-nos fortemente nos artigos de Todd [12, 14], Anstreicher [3] e Gonzaga [7, 8]. Neste sentido, desenvolvemos dois algoritmos primais de pontos interiores para programação linear combinando fase 1 e fase 2: um afim e outro projetivo. Nosso procedimento para atualizar limitantes inferiores é uma generalização de Gonzaga [9], o qual é equivalente ao desenvolvido em Todd [14]. Mostramos uma equivalência, sob certa condição,

entre as direções de Todd [12, 14] e as nossas. Tanto os limitantes inferiores quanto as direções são obtidas com complexidade $O(n)$, onde n é o número de variáveis. Nosso procedimento de busca unidirecional fornece um tamanho do passo maior do que o de Todd [12, 14]. Permitimos, através desta busca, passos longos nos algoritmos o que não ocorre necessariamente em Todd [12, 14]. Ainda, provamos a convergência de ambos os algoritmos, cuja complexidade é da ordem de nL iterações, onde n é o número de variáveis e L é o "tamanho dos dados de entrada" do problema original. Naturalmente, os dados de entrada são inteiros.

Procuramos organizar este trabalho de maneira construtiva. Desta forma, o capítulo 2 trata-se de uma introdução aos algoritmos de pontos interiores em programação linear. O Capítulo 3 reformula o problema original, fornece os algoritmos e alguns resultados necessários aos conteúdos seguintes. O procedimento de atualização de limitantes inferiores é apresentado no capítulo 4. A essência deste trabalho é tratada em detalhes no capítulo 5, onde desenvolvemos uma estratégia para encontrar direções. No capítulo 6 fornecemos o procedimento de busca unidirecional. Finalmente, enunciamos e provamos os resultados sobre convergência no capítulo 7.

Meus agradecimentos para o meu orientador Clovis Caesar Gonzaga por ter me possibilitado acompanhar dinamicamente todos os enunciados e demonstrações do

Capítulo 4 em diante, cujos resultados são todos de sua autoria. Deve-se ressaltar a maioria dos resultados dos Capítulos 2 e 3, inclusive. Estes resultados deram consistência para a Tese como um todo. Não resta dúvida que a originalidade deste trabalho se deve ao meu orientador.

CAPITULO 2

ALGORITMOS DE PONTOS INTERIORES EM PROGRAMAÇÃO LINEAR:

UMA INTRODUÇÃO

Consideremos um problema de programação linear na forma padrão

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

onde $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $0 < m < n$ e $A \in \mathbb{R}^{m \times n}$ é uma matriz com m linhas (ou colunas) linearmente independentes.

Trataremos conceitos importantes ao melhor entendimento dos capítulos seguintes - veja Gonzaga [6].

2.1 - Notação e ferramentas

Um quadrado claro \square , denotará o final de definições, exemplos, procedimentos, algoritmos e fatos sem demonstração. Um quadrado escuro, \blacksquare , denotará o final de um enunciado com demonstração. A transposta de uma matriz A será denotada por A' , sua inversa por A^{-1} , a matriz identidade por I , e o produto escalar entre dois vetores u , $v \in \mathbb{R}^n$ por $u'v$. Sub-índices serão usados normalmente para

denotar as componentes x_1, x_2, \dots , de um vetor x . Super-índices denotam vetores em \mathbb{R}^n . A letra e denotará o vetor de uns de dimensão apropriada $e = [1 \dots 1]'$. Estas notações seguirão todo o restante deste trabalho.

A um vetor $x \in \mathbb{R}^n$ associa-se a norma euclidiana

$$\|x\| = \sqrt{(x'x)},$$

e a norma 1 (da soma)

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

Estabelecemos a seguir um fato acerca de desigualdades associadas às normas, o qual usaremos à vontade em todo o trabalho.

Fato 2.1: Considere $x, y \in \mathbb{R}^n$. Então:

(i) (Desigualdade de Cauchy-Schwartz): $|x'y| \leq \|x\| \|y\|$;

(ii) (Desigualdade Triangular): $\|x + y\| \leq \|x\| + \|y\|$; e

(iii) $\|x\| \leq \|x\|_1$. \square

Uma transformação linear definida em \mathbb{R}^n com valores em \mathbb{R}^m é representada por uma matriz $A \in \mathbb{R}^{m \times n}$. Com este fato bem entendido, pode-se confundir a notação e falar de uma transformação $A \in \mathbb{R}^{m \times n}$.

Rank: o *rank* de uma transformação linear A , $r(A)$, é igual ao número de linhas (colunas) linearmente independentes. A é uma matriz de *rank* máximo se $r(A) = \min(m, n)$. \square

Note que em (P), $r(A) = m$. Ou seja, supomos que se a matriz A tem m linhas linearmente dependentes, então as redundâncias serão eliminadas de modo a reduzir o valor de m a $r(A)$. Isto é permitido quando uma matriz está sendo usada para representar um sistema de equações lineares $Ax = b$, como será o nosso caso.

A uma transformação $A \in \mathbb{R}^{m \times n}$ associam-se dois espaços importantes: o espaço imagem de A ,

$$I(A) = \left\{ y \in \mathbb{R}^m \mid y = Ax, x \in \mathbb{R}^n \right\},$$

e o espaço nulo de A ,

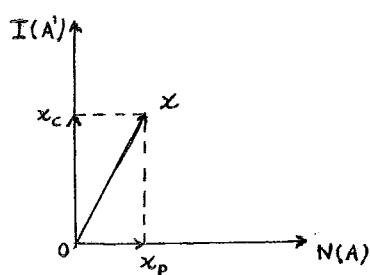
$$N(A) = \left\{ x \in \mathbb{R}^n \mid Ax = \theta \right\}.$$

Existe uma relação geométrica entre $N(A)$ e o espaço imagem da transposta de A , $I(A')$: esses dois espaços são subespaços ortogonais de \mathbb{R}^n e geram o espaço todo.

Deste fato conclui-se que \mathbb{R}^n é a soma direta das subespaços $N(A)$ e $I(A')$. Ou seja, dado qualquer vetor $x \in \mathbb{R}^n$, ele pode ser expresso como

$$x = x_p + x_c,$$

onde $x_p \in N(A)$ e $x_c \in I(A')$. O vetor x_p é a projeção de x sobre $N(A)$ e x_c é o complemento ortogonal de x em relação ao $N(A)$, isto é, $x_c \perp N(A)$. Veja figura abaixo.



- Figura 2.1 -

O próximo fato estabelece o cálculo da projeção sobre $N(A)$, onde P_A é a matriz de projeção sobre $N(A)$. Ainda, denotamos em todo o trabalho o subscrito "p" para designar os vetores projetados no $N(A)$ tal que $Ax = b$.

Fato 2.2: Seja $A \in \mathbb{R}^{m \times n}$ uma matriz de rank m , $0 < m < n$ e $x \in \mathbb{R}^n$ um vetor arbitrário. Então:

(i) A matriz AA' é não singular; e

$$(ii) \quad x_c = A'(AA')^{-1}Ax \text{ e } x_p = P_A x,$$

$$\text{onde } P_A = I - A'(AA')^{-1}A. \quad \square$$

O cálculo da projeção inclui a resolução de um sistema linear, o que é um procedimento trabalhoso aos algoritmos de pontos interiores em programação linear, em geral.

As seguintes definições devem ser entendidas para o desenvolvimento de algoritmos de pontos interiores.

Definição 2.3: Considere um problema (P) como definimos anteriormente. Então:

(i) O conjunto viável de (P) é definido por

$$S = \{x \in \mathbb{R}^n \mid Ax = b, \quad x \geq 0\};$$

(ii) $\hat{x} \in \mathbb{R}^n$ é uma solução ótima para (P) se, e só se, $\hat{x} \in S$ e $c'\hat{x} \geq c'x$ para todo $x \in S$;

(iii) Um ponto $x \in \mathbb{R}^n$ é ponto interior de (P) se, e só se, $x \in S$ e $x > 0$;

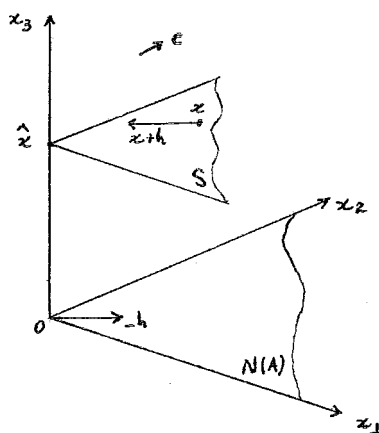
(iv) Direção é um vetor; e

(v) Um vetor $h \in \mathbb{R}^n$ é uma direção viável a partir de $x \in S$ se, e só se, existe $\bar{\lambda} > 0$ tal que para qualquer $\lambda \in [0, \bar{\lambda}]$, $x + \lambda h \in S$. \square

Segue-se um fato importante acerca de direções.

Fato 2.4: Se $x \in S$ é um ponto interior, onde S é o conjunto viável, então $h \in \mathbb{R}^n$ é uma direção viável a partir de x se, e só se, $h \in N(A)$. \square

Uma representação geométrica em \mathbb{R}^3 da Definição 2.3 e do Fato 2.4 está na figura abaixo; para $\lambda = 1$.



- Figura 2.2 -

Como sempre trabalharemos com pontos interiores, as direções viáveis serão normalmente confundidas com os vetores do espaço nulo da matriz A .

Na próxima seção definiremos funções potenciais, mudança de escala, e estabelecemos uma relação entre elas. Além disso, introduzimos mais notações e ferramentas de acordo com nossas necessidades nos capítulos seguintes,

2.2 - Funções potenciais e mudança de escala

Usaremos a seguinte notação ao primeiro ortante:

$$\mathbb{R}_+^n = \left\{ x \in \mathbb{R}^n \mid x \geq 0 \right\},$$

e

$$\mathbb{R}_{++}^n = \left\{ x \in \mathbb{R}^n \mid x > 0 \right\}.$$

\mathbb{R}_+^n é o primeiro ortante, enquanto \mathbb{R}_{++}^n refere-se aos vetores estritamente positivos em \mathbb{R}_+^n . Dado um vetor $x \in \mathbb{R}^n$, a letra maiúscula correspondente será utilizada para indicar a matriz diagonal

$$X = \text{diag}(x_1, \dots, x_n).$$

Derivada direcional: Considere uma função $f: T \rightarrow \mathbb{R}$, onde $T \subset \mathbb{R}^n$ é um conjunto aberto. A derivada de $f(\cdot)$ em $x \in T$ na direção $h \in \mathbb{R}^n$ é

$$f'(x, h) = \lim_{\lambda \rightarrow 0} \frac{f(x + \lambda h) - f(x)}{\lambda}$$

se o limite existir. Além disso, se $f(\cdot)$ é diferenciável em $x \in T$, então

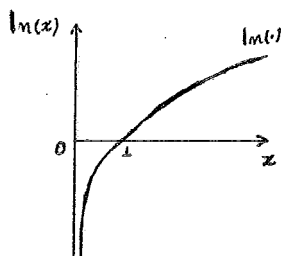
$$f'(x, h) = \nabla f(x)' h. \quad \square$$

Alguns exemplos de funções diferenciáveis são dados a seguir.

Exemplo 2.5: (i) (Função logaritmo): A função logaritmo, $\ln(\cdot)$, é definida por

$$\ln(\cdot) : \mathbb{R}_{++} \rightarrow \mathbb{R}.$$

Seu gráfico está representada na figura abaixo.



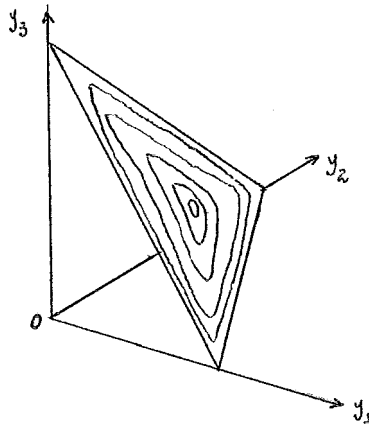
- Figura 23 -

Note que quando x tende a zero, $\ln x$ tende a $-\infty$. Além disso, a função logaritmo é crescente: e

(ii) (Função barreira): A função barreira, $p(\cdot)$ é definida por

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

Seu gráfico está representado abaixo pelas suas curvas de nível restritas a um poliedro em \mathbb{R}^3 .



- Figura 2.4 -

Note que quando alguma componente y_i , $i=1, \dots, n$, do vetor $y \in \mathbb{R}_{++}^n$ tende a zero, a função barreira tende a $+\infty$. \square

Trabalhando com problemas de programação linear, Karmarkar E113 definiu uma "função potencial" que toma o logaritmo da custo acrescido da função barreira. Desta forma, é necessário garantir que o custo seja sempre positivo. Faremos então uma "hipótese forte" provisória para esta seção, que pode ser removida com a utilização de limitantes inferiores.

Hipótese forte: O problema de programação linear (P) tem solução ótima \hat{x} e seu custo é nulo, isto é, $c' \hat{x} = 0$. \square

Com esta "hipótese forte" definimos:

Definição 2.6: A função potencial é definida por

$$f(x) = q \ln c'x - \sum_{i=1}^n \ln x_i,$$

onde $q \geq n$. \square

A função potencial de Karmarkar é definida para $q = n$. Gonzaga [7] definiu uma função potencial para $q \geq n + \sqrt{n}$ com $q = O(n)$.

É padrão em algoritmos de pontos interiores em programação linear trabalhar sobre um problema escalado, obtido através de uma mudança de escala.

Definição 2.7 : Considere o problema (P) como definimos anteriormente, um ponto viável $x^0 \in S$ tal que $x^0 > 0$, e a matriz diagonal $X_0 = \text{diag}(x_1^0, \dots, x_n^0)$. Definimos uma mudança de escala como uma transformação afim

$$x \longmapsto \bar{x} = X_0^{-1} x. \quad \square$$

Observações:

(i) O problema (P) torna-se um problema escalado (feita uma mudança de escala)

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

onde $\bar{A} = AX_0$ e $\bar{c} = X_0c$; e

(ii) Para um dado vetor $x^0 \in \mathbb{R}_{++}^n$ viável em (P), pode-se definir um problema (\bar{P}) em que o ponto e é viável e corresponde a x^0 . \square

O problema dual: O dual de (P) é dado por

$$\begin{aligned}
 & \text{maximizar } b'y \\
 & \text{sujeito a: } A'y + s = c \\
 & s \geq 0. \quad \square
 \end{aligned}$$

Condições de otimalidade para o problema de programação linear: As condições de otimalidade em programação linear são dadas pelas condições de Karush-Kuhn-Tucker, a saber: $x^* \in \mathbb{R}_+^n$ é uma solução ótima para o problema (P) se, e só se, existem multiplicadores $y_i \in \mathbb{R}$, $i = 1, \dots, m$ e $s_i \in \mathbb{R}$; $i = 1, \dots, n$ tais que

$$c - \sum_{i=1}^m y_i A_i' - \sum_{i=1}^n s_i I_i = 0, \quad (2.1)$$

$$s_i \geq 0, \quad i = 1, \dots, n, \quad (2.2)$$

$$\sum_{i=1}^n s_i x_i = 0, \quad (2.3)$$

x viável. \square

As condições (2.1) e (2.2) indicam que $y \in \mathbb{R}^m$ e $s \in \mathbb{R}^n$ são soluções dual viáveis. O *gap* de dualidade é caracterizado pela condição (2.3) conforme fato a seguir. Note que a folga complementar (2.3) é nula em qualquer solução ótima em programação linear.

Fato 2.8: Dado qualquer par (x, s) onde $x \in S = \{x \in \mathbb{R}_+^n \mid Ax = b\}$ e (y, s) é viável para o dual do problema (P) para algum $y \in \mathbb{R}^m$, então $x's = c'x - b'y$. \square

Funções potenciais

Uma propriedade chave dessas funções potenciais é sua invariância sob escalas positivas das variáveis. Ou seja, considere $x^0 > 0$ e $q \geq n$. Considere uma mudança de escala

$$\bar{x} = X_0^{-1} x.$$

Então o valor da função potencial

$$x > 0 \mapsto f(x) = q \ln c'x + p(x)$$

difere de

$$\bar{x} > 0 \mapsto \bar{f}(\bar{x}) = q \ln \bar{c}'\bar{x} + p(\bar{x})$$

por uma constante dependendo somente de x^0 . Vejamos isto:

Considere um ponto $x > 0$. Então

$$\begin{aligned} f(x) &= f(X_0 \bar{x}) = q \ln (c'X_0 \bar{x}) - \sum_{i=1}^n \ln x_i^0 \bar{x}_i = \\ &= q \ln (\bar{c}'\bar{x}) - \sum_{i=1}^n \ln \bar{x}_i - \sum_{i=1}^n \ln x_i^0 = \\ &= \bar{f}(\bar{x}) - \sum_{i=1}^n \ln x_i^0. \end{aligned}$$

Então, para reduzir $f(\cdot)$ pode-se reduzir $\bar{f}(\cdot)$ e voltar para a escala original, uma vez que $f(\cdot)$ e $\bar{f}(\cdot)$ diferem por uma constante, apenas.

Na próxima seção desenvolvemos uma variante do algoritmo de Karmarkar - veja Gonzaga [8].

2.3 - O algoritmo de Karmarkar

O problema será enunciado em um formato especial

$$\begin{aligned}
 (K): \text{ minimizar } c'x \\
 \text{sujeito a: } Ax = 0 \\
 a'x = 1 \\
 x \geq 0,
 \end{aligned}$$

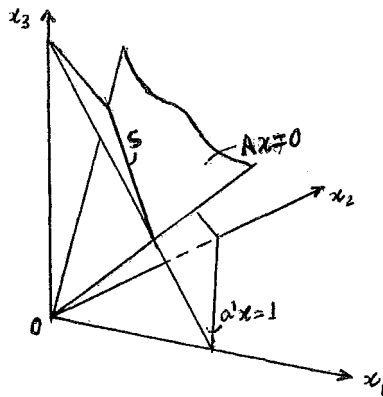
onde $c, a \in \mathbb{R}^n$, $a \geq 0$, $A \in \mathbb{R}^{m \times n}$ é uma matriz de rank máximo, $0 < m < n$.

O formato é tão geral como o padrão. Isto é, para colocar um problema no formato padrão no formato (K), basta introduzir uma variável o com a restrição $o = 1$ e escrever $Ax - bo = 0$. Naturalmente, o vetor $a \in \mathbb{R}^n$ será então $[0 \dots 0 1]'$. Note que o vetor c muda de dimensão também.

Hipótese para o algoritmo de Karmarkar: Fazemos as hipóteses comuns e mais uma "hipótese forte":

- (i) O conjunto viável é limitado;
- (ii) Seja $S = \{x \in \mathbb{R}_+^n \mid Ax = 0, a'x = 1\}$ o conjunto viável. Um ponto inicial viável $x^0 > 0$, $x^0 \in S$ é dado; e
- (iii) (Hipótese forte): O custo de qualquer solução ótima \hat{x} é $c' \hat{x} = 0$. \square

O conjunto viável S pode ser visualizado como a interseção de um subespaço, $Ax = 0$, e um hiperplano, $a'x = 1$, no primeiro ortante, $x \geq 0$. Veja a figura abaixo para $n = 3$. Denominamos o conjunto $\{x \in \mathbb{R}_+^n \mid Ax = b\}$ de cone.



- Figura.2.5 -

Karmarkar [11] definiu uma função potencial ($q = n$) e reduziu a problema de programação linear a encontrar um valor muito negativo para essa função. Formalmente, o problema a ser resolvido é:

"Encontrar $x \in S$ tal que $f(x) = q \ln c'x + p(x) < -3nL$ ".

Aqui n é o número de variáveis e L é o "tamanho das dadas de entrada" do problema. Uma solução deste problema terá um custo muito pequeno da ordem de 2^{-L} . É aqui onde se

inicia um processo denominado "algoritmo de purificação", o que não trataremos aqui (veja Gonzaga [6], Cap.6, Seção 6.3).

Homogeneidade da função potencial de Karmarkar: A função potencial de Karmarkar é homogênea de grau zero. Isto é, para qualquer escalar $\alpha > 0$ e qualquer vetor $x > 0$,

$$f(\alpha x) = f(x) = n \ln c'x - \sum_{i=1}^n \ln x_i.$$

Este fato pode ser verificado da seguinte maneira: Considere um ponto $x > 0$. Usando propriedade do logaritmo,

$$f(x) = \ln \frac{(c'x)^n}{\prod_{i=1}^n x_i}.$$

Por outro lado,

$$f(\alpha x) = \ln \frac{\alpha^n c'x}{\alpha^n \prod_{i=1}^n x_i} = f(x).$$

Projeção cônica

A homogeneidade da função potencial de Karmarkar permite uma conclusão muito interessante:

"A restrição $a'x = 1$ pode ser sumariamente ignorada".

Vejamos isto: pela figura 2.5 note que o cone é um conjunto de raias partindo da origem e passando por S . A cada ponto de S está associada uma única raia. E, a cada ponto $x \neq 0$ no cone, corresponde um único ponto viável: o ponto em que a raia passando por x intercepta a restrição $a'x = 1$. Este ponto é chamado de projeção cônica, $K(x)$, de x sobre S , e sua determinação é imediata

$$K(x) = \frac{x}{a'x},$$

pois

$$a'K(x) = \frac{a'x}{a'x} = 1.$$

Portanto, como a função potencial de Karmarkar é constante em uma raia qualquer, segue-se que trabalhar no cone ou em S são equivalentes a menos de uma projeção cônica.

Finalmente, algoritmos de pontos interiores evoluem por uma seqüência de iterações, cada uma contendo a escolha de uma direção viável e uma busca unidirecional ao longo dessa direção, de modo a obter um novo ponto com melhor valor para algum critério. Naturalmente, os pontos dessa seqüência são todos interiores no conjunto viável.

Assim necessita-se de um método para a obtenção de um ponto ótimo. Em programação linear, existe um vértice que é solução ótima - veja Humes e Humes [10]. A obtenção de uma solução ótima exata se dá através do "algoritmo de purificação", o qual procura um vértice ótimo (veja Gonzaga [6]).

A seguir vamos escrever o algoritmo de Karmarkar. Denotaremos "argmin" ao resultado da minimização. O subscrito "q" denotará vetores projetados no espaço nulo da matriz A tal que $Ax = 0$. Estas notações seguirão todo o restante deste trabalho. Além disso, definimos $\bar{f}(\cdot)$ como a função potencial de Karmarkar após a mudança de escala.

Algoritmo 2.9: Karmarkar

Dados: $x^0 > 0$, $x^0 \in S$ e $L > 0$.

$k := 0$.

Repita:

Mudança de escala: $X_k := \text{diag}(x_1^k, \dots, x_n^k)$.

Calcule $A_k := AX_k$, $c^k := X_k c$.

Projeção: $P_{A_k} := I - A_k'(A_k A_k')^{-1} A_k$.

Calcule $c_q := P_{A_k} c^k$.

Direção: $h := -P_{A_k} \nabla \bar{f}(e)$.

Busca unidirecional:

$$\bar{\lambda} := \operatorname{argmin} \{ \bar{f}(e + \lambda h) \mid \lambda \geq 0, e + \lambda h \succ 0 \}.$$

Passo: $\bar{y} := e + \lambda h$.

Novo ponto: $x^{k+1} := K(X_k \bar{y})$.

$k := k + 1$.

Até que $f(x^k) < -3nL$. \square

Observações:

(i) A projeção do vetor e no $N(A)$ é ele mesmo, isto é, $e_q = e$.

(ii) A projeção cônica está embutida em "Novo ponto". Observe que pode-se introduzir a projeção cônica após "Repita - Até que". Neste caso, teríamos um ponto viável somente na última iteração. \square

Agora retiramos a "hipótese forte". Neste sentido, podemos definir a função potencial como

$$x \in \mathbb{R}_{++}^n \mapsto f(x, z) = q \ln(c'x - z) - \sum_{i=1}^n \ln x_i,$$

onde $q \geq n$ e z é um limitante inferior obtido por algum procedimento (No nosso caso veja Capítulo 4.).

Com sua estratégia de resolução (algoritmos de pontos interiores), Karmarkar enriqueceu ainda mais a

Programação Matemática como uma disciplina que se auto-interliga de uma maneira muito forte. Por exemplo, a função potencial $f(\cdot)$ é uma função não linear.

No próximo capítulo daremos ênfase à construção do nosso problema que é equivalente ao problema original.

CAPITULO 3

OS PROBLEMAS

Consideremos o problema de programação linear na forma padrão

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

onde $\vec{A} \in \mathbb{R}^{m \times n}$ é uma matriz de *rank* máximo, $0 < m < n$.

Faremos reformulações necessárias em (P) para duas abordagens distintas, as quais serão utilizadas durante todo o trabalho. Além disso, fornecemos os algoritmos e demonstramos alguns resultados acerca de determinados conteúdos adiante.

3.1 - Formulações

Nesta seção faremos duas formulações: uma afim e outra projetiva.

Formulação afim

Algoritmos de pontos interiores para programação linear partem normalmente de um ponto viável interior

inicial, *que* deve ser fornecido por algum procedimento de inicialização (fase 1).

O procedimento mais usual parte de um ponto

$$e = [1 \dots 1]' > 0,$$

não viável (se o for é desnecessária a fase 1).

Introduz-se uma nova variável τ , e modifica-se a restrição

$$\bar{A}x = b \text{ para}$$

$$\bar{A}x + (b - \bar{A}e)\tau = b,$$

onde $[x' \tau]' = [e' 1]$ satisfaz a equação acima.

Desta forma, podemos escrever um problema de programação linear fase 1,

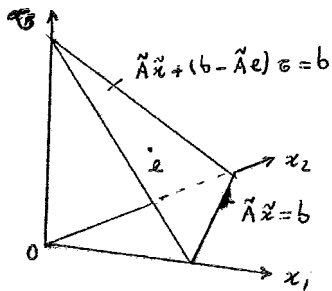
minimizar τ

$$\text{sujeito a: } \bar{A}x + (b - \bar{A}e)\tau = b,$$

$$\tau \geq 0$$

$$x \geq 0.$$

O ponto $e \in \mathbb{R}^{n+1}$ é viável para este problema transformado, e deve-se reduzir τ a zero para obter um ponto viável ao problema original (P). Isto pode ser visualizado na figura abaixo, onde $n = 2$.



- Figura 3.1 -

Uma alternativa comum para reduzir τ é o método "big M", (Veja Adler, Resende, Veiga e Karmarkar [1], onde se usa uma "heurística" sobre o "big M"). Este método consiste em reduzir τ associando a esta variável um custo muito alto, M , a saber:

$$\begin{aligned} &\text{minimizar} && \bar{c}'\bar{x} + M\tau \\ &\text{sujeito a:} && \bar{A}\bar{x} + (b - \bar{A}e)\tau = b \\ &&& \tau \geq 0 \\ &&& \bar{x} \geq 0. \end{aligned}$$

Assim força-se a diminuir rapidamente com qualquer algoritmo de minimização.

Este método é sujeito a problemas numéricos: por exemplo, soma a multiplicação de números muito pequenos (queremos reduzir a função objetivo), com um número muito grande, M ; a obtenção do número M é deselegante, "chutando-se" um valor; além de gastar algumas iterações.

Em nosso estudo procuramos reduzir simultaneamente a variável τ e o critério original $\bar{c}'x$, como em um problema de critérios múltiplos, sem associar custo à variável τ .

Neste sentido, definindo

$$c' = [c' \ 0], \quad A = [\bar{A} \ b - \bar{A}e], \quad \xi' = [0' \ 1] \text{ e } x' = (x' \ \tau),$$

obtemos um problema equivalente a (P),

$$\begin{aligned} \text{(PA):} \quad & \text{minimizar } c'x \\ & \text{sujeito a: } Ax = b \\ & \xi'x = 0 \\ & x \geq 0. \end{aligned}$$

Podemos considerar o seguinte conjunto de trabalho:

$$(SA)' = \{x \in \mathbb{R}_+^n \mid Ax = b\}.$$

O conjunto

$$(SA) = \{x \in \mathbb{R}_+^n \mid Ax = b, \xi'x = 0\},$$

é o conjunto viável que queremos atingir. Note que $(SA) \subset (SA)'$. Além disso, note que $\xi'x \geq 0$ por definição.

Pelo fato de trabalharmos em $(SA)'$, dizemos que o problema (PA) permite desenvolver algoritmos afins. Isto é, permite algoritmos que geram pontos interiores no conjunto $(SA)'$.

Vejamos a geometria de (PA) através de um exemplo.

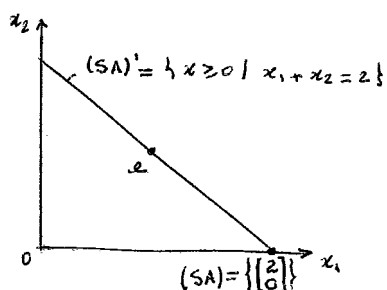
Exemplo 3.1: Considere o problema de programação linear na forma padrão,

$$\begin{aligned} (P) = \text{minimizar} & \quad -x_1 \\ \text{sujeito a} & \quad x_1 = 2 \\ & \quad x_1 \geq 0. \end{aligned}$$

Então o conjunto viável é igual a (2) e a solução ótima igual ao ponto $x_1 = 2$. Seu valor ótimo -2. A partir do problema (P), obtemos um problema equivalente

$$\begin{aligned} (PA) = \text{minimizar} & \quad -x_1 \\ \text{sujeito a} & \quad x_1 + x_2 = 2 \\ & \quad x_2 = 0 \\ & \quad x_1, x_2 \geq 0. \end{aligned}$$

A figura a seguir representa o conjunto viável transformado (SA) e o conjunto transformado $(SA)'$.



- Figura 3.2 -

Note que queremos atingir (SA) através de $(SA)'$. \square

Hipóteses: Fazemos as hipóteses comuns sobre o problema (PA) :

(i) O conjunto viável (SA) é limitado;

(ii) $0 \notin (SA)$; e

(iii) Um limitante inferior inicial $z \leq \hat{z}$ é conhecida, \hat{z} denota o valor ótimo de (PA) . \circ

A hipótese (i) é fundamental nas demonstrações, enquanto (ii) evita a situação trivial quando $0 \in (SA)$ (0 é a solução ótima ou o problema (PA) é ilimitado), e (iii) pode ser descartada devido o próprio procedimento de atualização que utilizaremos (capítulo 4), isto é, podemos utilizá-lo a partir do primeiro ponto dado $x^0 := e$.

Funções Potenciais:

Definimos duas funções potenciais as quais serão tratadas com mais detalhes no Capítulo 5,

$$f_1(x) = q \ln \xi'x - \sum_{i=1}^n \ln x_i,$$

e

$$f_2(x, z) = \begin{cases} q \ln(c'x - z) - \sum_{i=1}^n \ln x_i, & \text{se } c'x - z > 0, \\ -\infty, & \text{se } c'x - z \leq 0, \end{cases}$$

onde $q \geq n + \sqrt{n}$ e z é um limitante inferior dado. Note pelo Exemplo 3.1 que podemos ter $c'x > z$ ou $c'x \leq z$.

Os gradientes no ponto e das funções potenciais logarítmicas $f_1(\cdot)$ e $f_2(\cdot, \cdot)$ são dados, respectivamente por

$$g_1 = \frac{q\xi}{\xi'e} - e,$$

e

$$g_2 = \frac{qc}{c'e - z} - e.$$

Definimos dois vetores importantes, os quais são as projeções sobre $N(A)$, dos gradientes g_1 e g_2 no ponto e , dados respectivamente por

$$d_1 = \frac{q\xi_P}{\xi'e} - e_{P'}$$

e

$$d_2 = \frac{qc_P}{c'e-z} - e_{P'}. \quad \square$$

Formato Projetivo

A partir de (\bar{P}) podemos obter um problema equivalente no "formato especial" conforme seção 2.3 do capítulo anterior quando estudamos algoritmo de Karmarkar. De modo análogo ao raciocínio anterior acerca da formulação afim, num contexto de inicialização, podemos obter um problema equivalente a (\bar{P}) , definindo

$$c' = [c' \ 00], \quad A = [A \ -b \ b \ -Ae], \quad w' = [0' \ 10],$$

$$\xi' = [0' \ 01] \text{ e } x' = [x' \ 0 \ \tau],$$

a saber:

$$\begin{aligned}
 \text{(PP)}: & \text{ minimizar } c'x \\
 & \text{sujeito a: } Ax = 0 \\
 & \quad \quad \quad w'x = 1 \\
 & \quad \quad \quad \xi'x = 0 \\
 & \quad \quad \quad x \geq 0.
 \end{aligned}$$

Podemos considerar o seguinte conjunto de trabalhos:

$$(SP)' = \{x \in \mathbb{R}_+^n \mid Ax = 0, w'x = 1\}.$$

O conjunto

$$(SP) = \{x \in \mathbb{R}_+^n \mid Ax = 0, w'x = 1, \xi'x = 0\},$$

é o conjunto viável que queremos atingir. Note que $(SP) \subset (SP)'$. Além disso note que $\xi'x \geq 0$ por definição.

Pelo fato de trabalharmos em $(SP)'$, dizemos que o problema (PP) permite desenvolver algoritmos projetivos. Isto é, permite algoritmos que geram pontos interiores no cone gerado por $(SP)'$.

Vejamos a geometria de (PP) através de um exemplo.

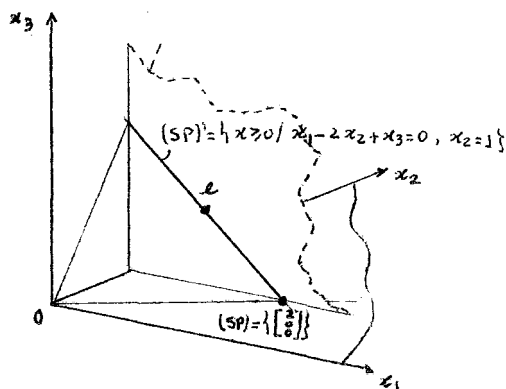
Exemplo 3.2 : Considere o problema de programação linear na forma padrão

$$\begin{aligned}
 \text{(PA):} & \text{ minimizar } -x_1 \\
 & \text{sujeito a: } x_1 = 2 \\
 & x_1 \geq 0.
 \end{aligned}$$

Então, o conjunto viável é igual a $\{2\}$ e a solução ótima igual ao ponto $x_1 = 2$. Seu valor ótimo -2 . A partir do problema (P), obtemos um problema equivalente

$$\begin{aligned}
 \text{(PA):} & \text{ minimizar } -x_1 \\
 & \text{sujeito a: } x_1 - 2x_2 + x_3 = 0 \\
 & x_2 = 1 \\
 & x_3 = 0 \\
 & x_1, x_2, x_3 \geq 0.
 \end{aligned}$$

A figura abaixo representa o conjunto viável transformado (SP) e o conjunto transformado (SP)'.



- Figura 3.3 -

Note que queremos atingir (SP) através de (SP)'. □

Hipóteses: Pela equivalência entre os problemas (P), (PA) e (PP), as hipóteses aqui serão as mesmas que definimos anteriormente. □

Funções Potenciais

Definimos duas funções potenciais, as quais serão tratadas com mais detalhes no capítulo 5:

$$F_1(x) = n \ln \xi'x - \sum_{i=1}^n \ln x_i,$$

e

$$F_2(x, z) = \begin{cases} n \ln(c - zw)'x - \sum_{i=1}^n \ln x_i, & \text{se } c'x - z > 0, \\ -\infty, & \text{se } c'x - z \leq 0, \end{cases}$$

onde z é um limitante inferior dado. Note pelo Exemplo 3.2 que podemos ter $c'x > z$ ou $c'x \leq z$.

Os gradientes no ponto e das funções potenciais de Kamarkar $F_1(\cdot)$ e $F_2(\cdot, \cdot)$ são dados, respectivamente por

$$G_1 = \frac{n\xi}{\xi'e} - e,$$

e

$$G_2 = \frac{n(c - zw)}{c'e - z} - e.$$

Definimos dois vetores importantes, os quais são as projeções sobre $N(A)$, dos gradientes G_1 e G_2 no ponto e , dados respectivamente por

$$D_1 = \frac{n\xi_q}{\xi'e} - e,$$

e

$$D_2 = \frac{n(c_q - zw_q)}{c'e - z} - e.$$

Aqui, definimos para todo o restante deste trabalho, na formulação afim, para $x \in (SA)'$ e z um limitante inferior para (PA),

$$f(x, z) = \max \{f_1(x), f_2(x, z)\}.$$

E, na formulação projetiva, para $x \in (SP)'$ e z um limitante inferior para (PP),

$$F(x, z) = \max \{F_1(x), F_2(x, z)\}.$$

3.2 - Os Algoritmos

Uma opção relevante na tentativa de desenvolver um trabalho de maneira construtiva é adiantar onde se pretende chegar, quando possível.

Dito isto, vamos construir dois algoritmos de pontos interiores para programação linear combinando fase 1 (viabilidade) e fase 2 (otimalidade): um afim e o outro projetivo. Os algoritmos reduzem $f(x,z)$ e $F(x,z)$, respectivamente, a $-O(nL)$ (n é o número de variáveis do problema original e L é o tamanho dos dados de entrada deste problema), e com isto obtêm um ponto x e um valor z tais que $\xi'x \leq 2^{-L}$ e $c'x - z \leq 2^{-L}$.

O modelo dos dois algoritmos de pontos interiores aqui apresentados consta de um "algoritmo mestre" que chama um "algoritmo interno". O algoritmo mestre consta de uma mudança de escala, do cálculo da matriz de projeção após a mudança de escala e da chamada do algoritmo interno que fornece um novo ponto. Termina por reverter a mudança de escala, determinando um ponto x^{k+1} . O algoritmo interno atualiza limitantes inferiores, calcula uma direção e executa uma busca unidirecional.

Notação: Consideremos k , a k -ésima iteração do algoritmo. Denotaremos em todo este trabalho, z^k um limitante inferior

tal que $z^k \leq \hat{z}$, onde \hat{z} é o valor ótimo do problema (PA) (ou (PP)) e P_{A_k} é a matriz de projeção sobre o $N(A_k)$. \square

Vejamos os algoritmos, cujos algoritmos internos serão tratados detalhadamente nos próximos capítulos. É importante salientar ao leitor neste momento, que os dois algoritmos aqui apresentados têm o objetivo de esclarecer onde pretendemos chegar. Se além da idéia, o leitor entender os algoritmos em seus mínimos detalhes, esperamos que o restante deste trabalho não seja uma leitura desagradável.

No algoritmo a seguir consideremos d_1 e d_2 os vetores definidos no início deste capítulo na parte referente a funções potenciais; no formato afim.

Algoritmo 3.3: Afim combinando fases 1 e 2.

Dados: $x^0 := e$, $z^0 \leq \hat{z}$ e uma precisão $\epsilon > 0$.

$k := 0$.

Repita:

Mudança de escala: $X_k := \text{diag}(x_1^k, \dots, x_n^k)$;

Calcule: $c^k := X_k c$, $A_k := AX_k$ e $\xi^k := X_k \xi$.

Projeção: $P_{A_k} := I - A_k'(A_k A_k')^{-1} A_k$;

Calcule: $c_p^k := P_{A_k} c^k$, $\xi_p^k := P_{A_k} \xi^k$ e $e_p^k := P_{A_k} e$.

Algoritmo interno:

Atualizar limitantes inferiores $z^{k+1} \geq z^k$, e fazer

$$\Delta_2^k = (c^k)' e - z^{k+1}.$$

Calcular uma direção de busca h^k com as características:

Se $\Delta_2^k > 0$, então

$$h^k = \operatorname{argmin} \{ \max \{ d_i'(h) \} \},$$

$$\|h\| \leq 1 \quad i \in \{1, 2\}$$

Senão

$$h^k = \operatorname{argmin} \{ d_i'(h) \mid c_p'(h) \leq -\Delta_2^k, i \in \{1, 2\} \},$$

$$\|h\| = 1$$

Busca unidirecional: Resolver aproximadamente

$$\alpha^k = \operatorname{argmin}_{\substack{e + \alpha h^k > 0 \\ \alpha \geq 0}} \{ \max \{ f_1(e + \alpha h^k) - \gamma, f_2(e + \alpha h^k, z^{k+1}) \} \},$$

onde

$$\gamma = \begin{cases} f_1(e) - f_2(e, z^{k+1}), & \text{se } \Delta_2^k > 0, \\ 0, & \text{se } \Delta_2^k \leq 0. \end{cases}$$

$$\text{Passo: } \bar{x} := e + \alpha^k h^k.$$

$$\text{Novo ponto: } x^{k+1} := X_k \bar{x}.$$

$$k := k + 1.$$

Até que $\xi'x^k \leq e$ e $c'x^k - z^{k+1} \leq e$. \square

Algoritmo 3.4: Projetivo combinado fases 1 e 2: Basta trocar $d_i(\cdot, \cdot)$, $i \in \{1, 2\}$, por $D_i(\cdot, \cdot)$, $i \in \{1, 2\}$ no Algoritmo 3.3, onde D_1 e D_2 estão definidos no início deste capítulo na parte referente a funções pontenciais; no formato projetivo. Além disso, deve-se fazer uma projeção cônica em "Novo ponto" da seguinte maneira:

$$x^{k+1} := K(X_k \bar{x}) \quad \square$$

Nosso objetivo daqui por diante é determinar precisamente os algoritmos internos (afim e projetivo), além de provar a convergência do Algoritmo 3.3 e do Algoritmo 3.4.

Antes, porém, vejamos alguns resultados na próxima seção.

3.3 - Alguns resultados

Vamos enunciar e demonstrar alguns resultados não tão evidentes nas demonstrações dos próximos capítulos.

Lema 3.5: Considere um problema de minimização de uma função linear em uma bola,

$$\begin{aligned} &\text{minimizar} && g'h \\ &\text{sujeito a:} && Ah = 0 \\ &&& \|h\| = 1. \end{aligned}$$

Então, o problema tem uma única solução, dada por

$$h = \frac{-g_p}{\|g_p\|},$$

onde $g_p = P_A g$.

Demonstração: Pode-se escrever

$$g = g_p + g_c,$$

onde $g_p \in N(A)$ e $g_c \perp N(A)$. Para qualquer direção viável h , $h \in N(A)$. Então,

$$g_c' h = 0.$$

Segue que

$$g' h = g_p' h.$$

O problema é portanto equivalente a

$$\begin{aligned} & \text{minimizar } g_p' h \\ & \text{sujeito a: } \|h\| = 1, \end{aligned}$$

com solução

$$h = \frac{-g_p}{\|g_p\|}. \quad \blacksquare$$

Lema 3.6: Considere $\alpha \in (0, 1)$. Então,

(i) Se $x > -1$, então $\ln(1+x) \leq x$;

(ii) Se $|x| \leq \alpha$, então $\ln(1+x) \geq x - \frac{x^2}{2(1-\alpha)}$.

Demonstração: Demonstraremos (i) e (ii) juntos. A primeira desigualdade é uma consequência da concavidade do logaritmo. Para obter a segunda desigualdade, desenvolva em série de Taylor

$$\begin{aligned} \ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots = \\ &= x - \frac{x^2}{2} \left(1 - \frac{2x}{3} + \frac{2x^2}{4} - \dots \right) \geq \\ &\geq x - \frac{x^2}{2} \left(1 + \frac{2|x|}{3} + \frac{2|x|^2}{4} + \dots \right) \geq \\ &\geq x - \frac{x^2}{2} \left(1 + |x| + |x|^2 + \dots \right) = \\ &= x - \frac{x^2}{2} \left(\frac{1}{1-|x|} \right), \end{aligned}$$

esta última expressão obtida da série geométrica. Mas, do fato de $|x| \leq \alpha$, segue-se que

$$\ln(1+x) \geq x - \frac{x^2}{2(1-\alpha)},$$

finalizando a demonstração. ■

Lema 3.7: Considere uma função $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Se $g: \mathbb{R} \rightarrow \mathbb{R}$ é uma função monotônica não-decrescente, então o problema

$$(1): \text{ minimizar } f(x) \\ \text{sujeito a: } x \in X$$

é equivalente ao problema

$$(2): \text{ minimizar } g(f(x)) \\ \text{sujeito a: } x \in X.$$

Demonstração: $\hat{x} \in X$ é solução ótima para o problema (1) se, e só se, $f(x) \geq f(\hat{x})$ para todo $x \in X$, ou equivalentemente, se e só se $g(f(x)) \geq g(f(\hat{x}))$ para todo $x \in X$, pela monotonicidade de $g(\cdot)$. ■

Lema 3.8: Considere $x, e \in \mathbb{R}_{++}^n$, onde $e = [1 \dots 1]'$. Então,

$$\sum_{i=1}^n \ln x_i \leq n \ln \frac{e'x}{n}.$$

Demonstração: Sabemos que a média geométrica de um conjunto de números positivos é menor ou igual a média aritmética desses números, isto é,

$$\left(\prod_{i=1}^n x_i\right)^{1/n} \leq \frac{\sum_{i=1}^n x_i}{n}.$$

Como a função logaritmo é crescente, podemos aplicá-la em ambos os lados da fórmula acima, mantendo a desigualdade. Em seguida, aplicando a propriedade de logaritmo da potência no lado esquerdo obtemos,

$$\frac{1}{n} \ln \prod_{i=1}^n x_i \leq \ln \left[\frac{\sum_{i=1}^n x_i}{n} \right]. \quad (3.1)$$

Por hipótese $x \in \mathbb{R}_{++}^n$ e $e = [1 \dots 1]'$. Logo,

$$e'x = \sum_{i=1}^n x_i. \quad (3.2)$$

Aplicando a propriedade de logaritmo do produto no lado esquerdo de (3.1) e usando (3.2) segue-se que

$$\frac{1}{n} \sum_{i=1}^n \ln x_i \leq \ln \frac{e'x}{n}.$$

Finaliza-se a demonstração multiplicando por n ambos os lados desta última desigualdade. ■

CAPÍTULO 4

LIMITANTES INFERIORES

Neste capítulo desenvolvemos um procedimento para atualizar limitantes inferiores em cada iteração. Esse procedimento é uma generalização de Gonzaga [9] e equivalente ao procedimento da Fraley, conforme Todd [14]. Assumimos feita uma mudança de escala para uma dada iteração.

Consideremos o problema de programação linear escalado reformulado

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

onde $A \in \mathbb{R}^{m \times n}$ tem rank máximo, $0 < m < n$.

Considere o conjunto viável

$$\text{(SA)} = \left\{ x \in \mathbb{R}_+^n \mid Ax = b, \xi'x = 0 \right\},$$

e o conjunto de trabalho

$$(SA)' = \{x \in \mathbb{R}_+^n \mid Ax = b\}.$$

Assumimos que na iteração corrente $e \in (SA)'$, e que um limitante inferior corrente é z (possivelmente $-\infty$).

Seja

$$\Delta_1 = \xi' e$$

a inviabilidade corrente, assumida estritamente positiva.

Iniciamos com um lema de dualidade (veja Gonzaga [9]). Aqui o subscrito "p" indica projeção sobre $N(A)$.

Lema 4.1: $s \in \mathbb{R}^n$ é uma folga dual viável para o problema (PA) se, e somente se,

$$(i) \quad s \geq 0 \text{ e } s_p = c_p - \lambda \xi_p,$$

ou equivalentemente,

$$(ii) \quad s \geq 0 \text{ e } s = c_p - \lambda \xi_p - \gamma \text{ com } \gamma \perp N(A).$$

Demonstração: A equivalência entre (i) e (ii) é imediata da definição de projeção, isto é, de (ii)

$$s_p = (c_p - \lambda \xi_p - \gamma)_p = c_p - \lambda \xi_p \text{ com } \gamma \perp N(A).$$

Agora, consideremos o dual de (PA),

(DA): maximizar $b'y$

$$\text{sujeito a: } A'y + \lambda\xi + s = c$$

$$s \geq 0.$$

s é uma folga dual viável se, e só se, $s \geq 0$ e $A'y = c - \lambda\xi - s$ para algum $y \in \mathbb{R}^m$ ($\lambda \in \mathbb{R}$). Esta última igualdade é equivalente a $c - \lambda\xi - s \perp N(A)$, uma vez que $c - \lambda\xi - s \in I(A')$. Mas isto é equivalente a $(c - \lambda\xi - s)_p = 0$, estabelecendo (i) e completando a demonstração. ■

Deste lema, vemos que para obtermos uma folga dual viável devemos tentar encontrar um vetor $\gamma \perp N(A)$ tal que $c_p - \lambda\xi_p - \gamma \geq 0$. A melhor escolha é certamente um γ onde uma folga dual ótima é obtida. Mas obtê-la é tão difícil quanto resolver o problema.

Para γ devemos ter componentes "muito negativas". $\gamma = -e$ deveria ser bom, mas isto em geral não é ortogonal ao $N(A)$. Vetores com o formato $\gamma = \mu(e - e_p)$, onde $\mu \in \mathbb{R}$ e $(e - e_p) \perp N(A)$ fornece bons resultados. Com isto em mente, estamos prontos para enunciarmos o procedimento para gerar um limitante inferior.

Consideremos uma folga dual viável

$$s = c_p - \lambda \xi_p - \mu(e - e_p), \quad (4.1)$$

onde $\lambda, \mu \in \mathbb{R}$, tal que $s \geq 0$. Ainda, consideremos um problema dual (DA). Seja um *gap*, denotado Δ , associado ao problema (PA), dado por

$$\Delta = c'e - b'y. \quad (4.2)$$

Temos da definição da matriz A (pela reformulação), que $b = Ae$. Além disso, pelo dual (DA), $-e'A'y = -e'(c - \lambda\xi - s)$. Substituindo em (4.2) a primeira igualdade e, posteriormente, a segunda, obtemos

$$\Delta = \lambda\xi'e + s'e. \quad (4.3)$$

Substituindo (4.1) em (4.3), e, ajeitando,

$$\Delta = c_p'e + \lambda(\xi - \xi_p)'e + \mu(e_p'e - n).$$

Como queremos o menor *gap* possível, segue-se um problema de minimização bi-dimensional para programação linear, semelhante ao desenvolvido em Todd [14],

$$\begin{aligned} \text{(LI)} = \quad & \text{minimizar } \lambda(\xi - \xi_p)'e + \mu(e_p'e - n) + c_p'e \\ & \text{sujeito a: } c_p - \lambda\xi_p - \mu(e - e_p) \geq 0. \end{aligned}$$

Então,

$$z = c'e - \bar{\Delta} \quad (4.4)$$

é o limitante inferior, onde $\bar{\Delta}$ é o valor ótimo de (LI), se existir. Por outro lado, se (LI) for inviável, permitimos $\Delta = +\infty$. Se (LI) tem solução ilimitada, permitimos $\Delta = -\infty$.

Analisando as soluções de (LI) e relacionado-as com (4.4) temos:

- (i) Se (LI) tem uma solução ótima finita, então z assume um valor finito. Logo, existe uma solução viável para o dual (DA);
- (ii) Se (LI) tem solução ilimitada, isto é, $\Delta = -\infty$, então $z = +\infty$. Logo, o dual (DA) tem solução ilimitada. Pela dualidade, (PA) é inviável, logo, (P). Aqui temos um critério de parada; e
- (iii) Se (LI) é inviável, isto é, $\Delta = +\infty$, então $z = -\infty$. Logo, não podemos dizer coisa alguma acerca do dual (DA). Aqui, devemos nos contentar com o limitante inferior corrente.

Portanto, concluímos o seguinte procedimento referente aos limitantes inferiores, considerando \bar{z} um limitante inferior corrente:

Procedimento 4.2: Limitante inferior.

Dado: \bar{z} .

Calcule (LI).

Teste:

Se (LI) é ilimitado, então:

"parada do algoritmo"

Senão

$$z := \max \{ \bar{z}, c' e - \Delta \}. \quad \square$$

O próximo capítulo é o fundamento deste trabalho como um todo.

CAPITULO 5

DIREÇÕES

Neste capítulo vamos desenvolver uma estratégia (*min-max*) para encontrar direções em algoritmos de pontos interiores para programação linear combinando fase 1 (viabilidade) e fase 2 (otimalidade). Esta estratégia é desenvolvida num conjunto afim (estratégias diferentes estão em Anstreicher [3] e Todd [14], em um espaço projetivo, (veja Anstreicher [3] e Todd [12] - veja também Todd e Wang [13]). Nossos resultados garantem um comportamento polinomial aos algoritmos que estamos construindo. Além disso, demonstramos equivalência entre as direções de Todd e as nossas para certa condição ("gap positivo"). em ambas as abordagens.

Importante: Nos Capítulos 5 e 6 assumimos feita uma mudança de escala, e que temos um limitante inferior z para uma dada iteração, obtido pelo Procedimento 4.2. \square

5.1 - Uma abordagem afim

Consideremos o problema de programação linear escalado reformulado

$$\begin{aligned}
 \text{(PA):} & \text{ minimizar } c'x \\
 & \text{sujeito a: } Ax = b \\
 & \quad \quad \quad \xi'x = 0 \\
 & \quad \quad \quad x \geq 0.
 \end{aligned}$$

onde $A \in \mathbb{R}^{m \times n}$ é uma matriz de *rank* máximo, $0 < m < n$. Além disso, consideremos o conjunto viável

$$(SA) = \left\{ x \in \mathbb{R}_+^n / Ax = b, \zeta'x = \theta \right\},$$

e o conjunto de trabalho

$$(SA)' = \left\{ x \in \mathbb{R}_+^n / Ax = b \right\}$$

Assumimos que na iteração corrente partimos do ponto $e \in (SA)'$.

Notação: Denotamos $\Delta_1 = \zeta'e$ e $\Delta_2 = c'e - z$ para um limitante inferior z , "p" relativo a vetores projetados no espaço nulo da matriz A_k , e "k" será omitido por se tratar de uma análise sobre uma iteração arbitrária. Em todo o trabalho usaremos os subscritos "1" com referência à fase 1 e "2" à fase 2. \square

O restante desta seção assumirá

$$\Delta_1 = \zeta'e > \theta.$$

Pelo exemplo 3.1, observe que é possível termos $c'e > z$, assim como $c'e \leq z$. Afinal, não estamos viáveis ($\Delta_1 > \theta$). Desta forma,

$$\Delta_2 = c'x - z \in \mathbb{R},$$

para um limitante inferior z .

Aqui a idéia é definir uma direção de busca no $N(A)$ que nos aproxime simultaneamente viabilidade e otimalidade.

Definimos duas funções potenciais logarítmicas variantes à de Karmarkar no Capítulo 3, a saber:

$$f_1(x) = q \ln c'x - \sum_{i=1}^n \ln x_i,$$

e

$$f_2(x, z) = \begin{cases} q \ln(c'x - z) - \sum_{i=1}^n \ln x_i, & \text{se } \Delta_2 > 0, \\ -\infty, & \text{se } \Delta_2 \leq 0, \end{cases}$$

onde $x \in (SA)'$, z é um limitante inferior e $q \geq n + \sqrt{n}$.

Nosso candidato para aproximar viabilidade está associado à função potencial $f_1(\cdot)$, enquanto para otimalidade à função $f_2(\cdot, \cdot)$.

Os gradientes no ponto e das funções potenciais $f_1(\cdot)$ e $f_2(\cdot, \cdot)$ são dados respectivamente por

$$g_1 = \frac{q \xi}{\Delta_1} - e,$$

e

$$g_2 = \frac{q c}{\Delta_2} - e.$$

As projeções sobre $N(A)$ dos gradientes g_1 e g_2 no ponto e definem dois vetores dados por

$$d_1 = \frac{q \xi_P}{\Delta_1} - e_P,$$

e

$$d_2 = \frac{q c_P}{\Delta_2} - e_P.$$

Segue-se um lema auxiliar para a demonstração dos Teoremas 5.5 e 5.11 adiante.

Lema 5.1: Considere o problema (PA) e o conjunto $(SA)'$. Tome $x = e \in (SA)'$ e $\Delta_1 = \xi' e > 0$. Então, $\|d_1\| \geq 1$.

Demonstração: Considere \hat{x} uma solução viável para (PA) e

$$h = \frac{\hat{x} - e}{\|\hat{x} - e\|}. \quad \text{Pela definição de } g_1,$$

$$g_1'h = \left(\frac{q \xi}{\Delta_1} - e \right)' \frac{\hat{x} - e}{\|\hat{x} - e\|} =$$

$$= \frac{1}{\|\hat{x} - e\|} \left[\frac{q \xi \hat{x}}{\Delta_1} - \frac{q \Delta_1}{\Delta_1} - e' \hat{x} + e'e \right]$$

Desenvolvendo esta igualdade, e usando o fato de \hat{x} ser uma solução viável para (PA), segue-se que

$$g_1'h = \frac{1}{\|\hat{x} - e\|} \left[-q - e' \hat{x} + \|e\|^2 \right].$$

Mas, $\hat{x} \geq 0$ implica $e' \hat{x} = \|\hat{x}\|_1 \geq \|\hat{x}\|$. Ainda, $q \geq n + \sqrt{n}$ e $\|e\|^2 = n$. Logo,

$$g_1'h \leq \frac{1}{\|\hat{x} - e\|} \left[-n - \sqrt{n} - \|\hat{x}\| + n \right] =$$

$$= \frac{-1}{\|\hat{x} - e\|} \left[\|\hat{x}\| + \sqrt{n} \right].$$

Note que $\|\hat{x} - e\| \leq \|\hat{x}\| + \|e\| = \|\hat{x}\| + \sqrt{n}$. Então,

$$g_1'h \leq -1.$$

Pelo fato de que $h \in N(A)$, tem-se $g_1'h = d_1'h$ e portanto

$$-d_1'h \geq 1.$$

Observe que $-d_1'h \leq \|d_1\| \|h\| = \|d_1\|$.

Finalmente,

$$\|d_1\| \geq 1. \quad \blacksquare$$

Separamos nossa estratégia para encontrar direções em dois casos, segundo $\Delta_2 > \emptyset$ ou $\Delta_2 \leq \emptyset$.

1.º Caso: Considere $\Delta_2 > \emptyset$, onde $\Delta_2 = c'e - z$ piava um limitante inferior z .

Queremos desenvolver uma estratégia para obtermos uma direção "boa". O que significa isto? Significa definir uma direção ao longo da qual as funções potenciais $f_1(\cdot)$ e $f_2(\cdot, \cdot)$ "caem bastante". Mostraremos que as derivadas direcionais de $f_1(\cdot)$ e $f_2(\cdot, \cdot)$ ao longo da direção obtida, são menores ou iguais a menos um.

Direção

Existe uma estratégia para encontrar direções com um suporte geométrico e algébrico simples e resultados eficientes.

Esta é uma direção obtida por um problema de *min-max*, a saber:

$$(EDA) = \begin{array}{ll} \text{minimizar} & \max_{i \in \{1,2\}} d_i' h. \\ \text{||h||} \leq 1 & \end{array}$$

Se r é o valor da solução de (EDA), e se h é a solução, então $d_1' h \leq r$ e $d_2' h \leq r$. Assim, r é o menor valor possível para o qual estas duas desigualdades são satisfeitas.

A seguir construiremos um procedimento sobre o problema (EDA): Considere β o valor ótimo de (EDA) e \hat{h} uma solução de (EDA). Obviamente $\beta \leq 0$ (veja Lema 5.1). Se $\beta < 0$, então $\|\hat{h}\| = 1$ e a derivada direcional de ambas as funções $f_1(\cdot)$ e $f_2(\cdot, \cdot)$ é negativa ao longo da solução \hat{h} , isto é, $g_1' \hat{h} \leq \beta$ e $g_2' \hat{h} \leq \beta$. Se $\beta = 0$, então não existe direção de decréscimo para ambas as funções, e pode haver uma indeterminação na escolha de \hat{h} . Para evitar essa indeterminação, o procedimento de busca de direção será:

Procedimento 5.2: Sobre (EDA).

Resolva (EDA).

Se o valor da solução for zero, faça $\hat{h} = 0$. □

No próximo lema mostraremos uma equivalência entre o Procedimento 5.2 para encontrar direções *min-max* e o seguinte procedimento (comum em otimização convexa):

Procedimento 5.3: Sobre (EDA)'.

Obtenha a solução ótima \hat{d} para o problema

(EDA)': minimizar $\|d\|$

sujeito a: $d = td_1 + (1 - t)d_2$

$t \in [0, 1]$.

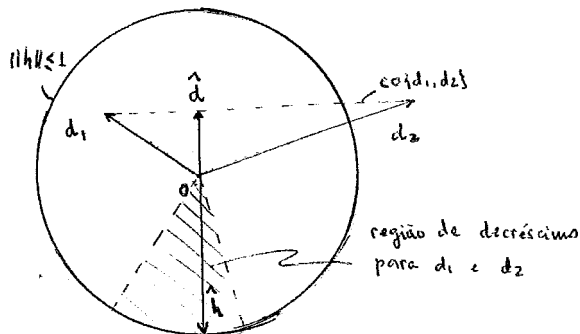
Faça $\hat{h} = -\frac{\hat{d}}{\|\hat{d}\|}$, se $\hat{d} \neq 0$.

Caso contrário, tome $\hat{h} = 0$. \square

O problema de encontrar direções em (EDA)' significa encontrar a menor distância da origem à envoltória convexa entre os vetores d_1 e d_2 (denotada por $\{d_1, d_2\}$), cujo resultado é um vetor \hat{d} . Daí, tome

$\hat{h} = \frac{-\hat{d}}{\|\hat{d}\|}$ se $\hat{d} \neq 0$. Veja figura abaixo. Note que as

direções obtidas estão no $N(A)$.



- Figura 5.1 -

Lema 5.4 - Os Procedimentos 5.2 e 5.3 são equivalentes e o valor da solução de (EDA) é o negativo do valor da solução de (EDA)'.

Demonstração: Considere o problema (EDA) e seja β o valor de sua solução ótima. Para h fixo, o problema

$$\begin{aligned} & \text{maximizar} && d'h \\ & \text{sujeito a} && d \in \text{co}\{d_1, d_2\}, \end{aligned}$$

é um problema de programação linear, e portanto tem uma solução ótima em um vértice. Como os vértices de $\text{co}\{d_1, d_2\}$ são d_1 e d_2 , os problemas (EDA) e

$$\begin{aligned} & \text{minimizar} && \max_{d \in \text{co}\{d_1, d_2\}} d'h \\ & \|h\| \leq 1 \end{aligned}$$

são equivalentes. Agora, como $d'h$ é uma função linear nas variáveis h e d , sobre conjuntos convexos compactos não vazios, $\|h\| \leq 1$ e $\text{co}\{d_1, d_2\}$, podemos inverter a ordem dos operadores *min-max* (veja Bazaraa e Shetty [4]; exercício 6.30) e escrever a equivalência

$$\begin{aligned} & \text{maximizar} && \min_{\|h\| \leq 1} d'h. \\ & d \in \text{co}\{d_1, d_2\} \end{aligned} \tag{5.1}$$

O problema interno em (5.1)

$$\begin{aligned} & \text{Minimizar } d'h \\ & \text{sujeito a: } \|h\| \leq 1 \end{aligned}$$

tem a seguinte solução:

(i) Se $d = 0$, então h é arbitrário e pode-se tomar $h = 0$.

(ii) Se $d \neq 0$, então $h = -\frac{d}{\|d\|}$; conforme Lema 3.5.

De (i), $d'h = 0$ e, (ii), $d'h = -\|d\| < 0$. Podemos portanto examinar dois casos:

(iii) Se $0 \in \text{co}\{d_1, d_2\}$, então $\hat{d} = 0$ é solução de (5.1)

(iv) Caso contrário, \hat{d} é solução de

$$\begin{aligned} & \text{maximizar } -\|d\| \\ & \text{sujeito a: } d \in \text{co}\{d_1, d_2\}. \end{aligned}$$

O valor da solução ótima é $-\|\hat{d}\| = \beta$. Finalmente, este problema é equivalente a

$$\begin{aligned} & \text{minimizar } \|d\| \\ & \text{sujeito a: } d \in \text{co}\{d_1, d_2\}. \end{aligned}$$

Isto conclui a equivalência com (EDA)', finalizando a demonstração. ■

O teorema a seguir relaciona nosso procedimento de atualização de limitantes inferiores do capítulo anterior com nossa direção obtida pelo Procedimento 5.3, com $\Delta_2 > 0$.

Teorema 5.5: Considere o problema (PA) e o conjunto (SA)'. Tome $x = e \in (SA)'$, $\Delta_1 > 0$ e $\Delta_2 = c'e - z > 0$ para um limitante inferior z obtido pelo procedimento 4.2. Então, $\|t d_1 + (1 - t) d_2\| \geq 1$ para todo $t \in [0, 1]$.

Demonstração: Faça

$$d = t d_1 + (1 - t) d_2, \text{ para } t \in [0, 1].$$

Se $t = 1$, então $d = d_1$ e, pelo Lema 5.1, $\|d_1\| \geq 1$. Seja, então, $t \in [0, 1)$. Suponha por absurdo que $\|d\| < 1$. Logo, $e + d > 0$ já que $-1 < d_j < a$, $j = 1, \dots, n$, onde d_j são as componentes do vetor d . Vamos construir uma folga dual viável no formato (4.1) usado pela atualização de limitantes inferiores do capítulo anterior: desenvolvendo a expressão para d ,

$$d = t \left[\frac{q \xi_p}{\Delta_1} - e_p \right] + (1 - t) \left[\frac{q c_p}{\Delta_2} - e_p \right] =$$

$$= \frac{(1-t)q}{\Delta_2} c_p + \frac{tq}{\Delta_1} \xi_p - e_p.$$

Segue-se que

$$e + d = \frac{(1-t)q}{\Delta_2} c_p + \frac{tq}{\Delta_1} \xi_p + e - e_p > 0.$$

Para obter uma folga dual viável no formato (4.1), façamos

$$\begin{aligned} s &= \frac{\Delta_2}{(1-t)q} (e + d) = \\ &= c_p + \frac{t \Delta_2}{(1-t) \Delta_1} \xi_p + \frac{\Delta_2}{(1-t)q} (e - e_p) > 0. \end{aligned} \tag{5.2}$$

A expressão acima está no formato (4.1) usado pelo Procedimento 4.2 do capítulo anterior e deve-se ter por definição de Δ_2 ,

$$\Delta(s) > \Delta_2,$$

onde $\Delta(s)$ é o *gap* associado ao s de (5.2). Finalmente, vamos calcular $\Delta(s)$ e estabelecer uma contradição: temos

$$s = \frac{\Delta_2}{(1-t)q} (e + d).$$

O *gap* correspondente é dado por (4.3), onde $\lambda = -\tau$, isto é,

$$\Delta(s) = e's - \tau \Delta_1,$$

onde $\tau = \frac{t\Delta_2}{(1-t)\Delta_1}$ é extraído de (5.2). Então,

$$\begin{aligned} \Delta(s) &= e's - \tau \Delta_1 = \\ &= \frac{\Delta_2}{(1-t)q} (e + d)' e - \frac{t\Delta_2}{(1-t)\Delta_1} \Delta_1 = \\ &= \frac{\Delta_2}{(1-t)q} (n + d'e) - \frac{t\Delta_2}{(1-t)}. \end{aligned}$$

Com a hipótese de que $\|d\| < 1$, tem-se $d'e \leq \|d\| \|e\| < \sqrt{n}$.

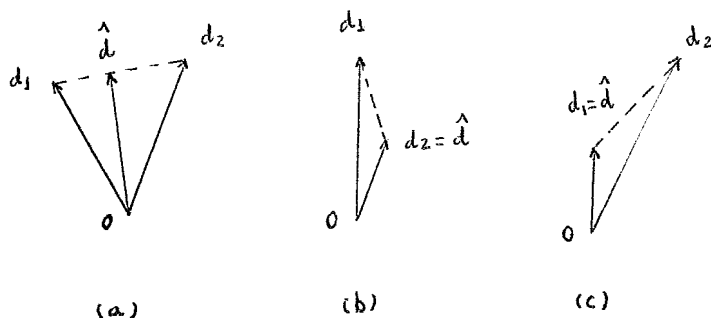
Portanto, $n + d'e < q$, já que $q \geq n + \sqrt{n}$.

Segue-se que

$$\Delta(s) < \frac{\Delta_2}{(1-t)} - \frac{t\Delta_2}{(1-t)} = \Delta_2,$$

completando a demonstração por absurdo. ■

Este teorema afirma que não obtemos uma solução $h = \emptyset$. Logo, podemos apresentar a geometria de (EDA) ou (EDA)' caracterizando três situações, dadas nas figuras a seguir.



- Figuras 5.2

Note que as direções de busca nas Figuras 5.2 são

$$h = \frac{-\hat{d}}{\|\hat{d}\|}.$$

Ainda pelo Teorema 5.5, podemos enunciar (EDA) na seguinte forma:

$$\begin{aligned} (\overline{\text{EDA}}) = & \text{Minimizar} \quad \max_{i \in \{1,2\}} d_i' h, \\ & \|h\| = 1 \end{aligned}$$

As soluções ótimas \hat{h} de $(\overline{\text{EDA}})$ e \hat{d} de $(\text{EDA})'$ estão relacionadas por $\hat{h} = \frac{-\hat{d}}{\|\hat{d}\|}$.

Devemos mostrar que a direção obtida ao resolvermos $(\overline{\text{EDA}})$ (ou $(\text{EDA})'$ que é mais conveniente computacionalmente) reduz as funções potenciais $f_1(\cdot)$ e $f_2(\cdot, \cdot)$ bastante. Trataremos disto na Perna a seguir.

Lema 5.6: Considere o problema (PA) e o conjunto (SA)'. Tome $x = e \in (SA)'$, $\Delta_1 = \xi'e > 0$, $\Delta_2 = c'e - z > 0$, para um limitante inferior $z \leq \hat{z}$, onde \hat{z} é o valor ótimo de (PA) e z é obtido pelo Procedimento 4.2, e tome \hat{h} a direção obtida resolvendo (\overline{EDA}) . Então,

$$(i) \quad g_1' \hat{h} \leq -1; \text{ e}$$

$$(ii) \quad g_2' \hat{h} \leq -1.$$

Demonstração: Demonstraremos (i) e (ii) juntos. Pela otimalidade de \hat{h} em (\overline{EDA}) , e pelo Lema 5.4,

$$g_i' \hat{h} \leq \min_{\|h\|=1} \max_{i \in \{1,2\}} d_i' h = - \min_{d \in \text{co}\{d_1, d_2\}} \|d\|.$$

Pelo Teorema 5.5, $\|d\| \geq 1$. Então,

$$g_i' \hat{h} \leq -1, \quad i \in \{1,2\},$$

finalizando a demonstração. ■

Agora estamos prontos para demonstrar um teorema fundamental à análise de convergência do algoritmo afim.

Teorema 5.7: Considere as hipóteses do Lema 5.6. Então, existe um α , $\alpha \in (0, 1)$ tal que

$$(i) f_1(e + \hat{\alpha}h) - f_1(e) \leq -0,25; \text{ e}$$

$$(ii) f_2(e + \hat{\alpha}h, z) - f_2(e, z) \leq -0,25.$$

Demonstração: (i) $f_1(e + \hat{\alpha}h) - f_1(e) =$

$$= q \ln \xi'(e + \hat{\alpha}h) - \sum_{i=1}^n \ln (e + \hat{\alpha}h)_i - \left[q \ln \Delta_1 - \sum_{i=1}^n \ln e_i \right].$$

Desenvolvendo esta igualdade, usando o fato de que

$$\sum_{i=1}^n \ln e_i = 0, \text{ e pelas propriedades do logaritmo,}$$

$$\begin{aligned} f_1(e + \hat{\alpha}h) - f_1(e) &= \\ &= q \ln \left(1 + \frac{\alpha \xi' h}{\Delta_1} \right) - \sum_{i=1}^n \ln (1 + \alpha \hat{h}_i). \end{aligned}$$

Pelo fato de $\|\hat{h}\| = 1$ implicar que $|\alpha \hat{h}_i| \leq \alpha < 1$, $i = 1, \dots, n$, e como

$$\left| \frac{\alpha \xi' h}{\Delta_1} \right| \leq \alpha,$$

uma vez que o valor absoluto da razão acima é o resultado de operações com a última componente com suas respectivas grandezas no problema escalado, podemos usar o Lema 3.6 obtendo,

$$\begin{aligned} f_1(e + \alpha \hat{h}) - f_1(e) &\leq \\ &\leq q \frac{\alpha \xi' \hat{h}}{\Delta_1} - \sum_{i=1}^n \alpha \hat{h}_i + \sum_{i=1}^n \frac{(\alpha \hat{h}_i)^2}{2(1-\alpha)}. \end{aligned}$$

Resolvendo os somatórios, e ajeitando,

$$f_1(e + \alpha \hat{h}) - f_1(e) \leq \alpha \left[\frac{q\xi}{\Delta_1} - e \right]' \hat{h} + \frac{\alpha^2 \|\hat{h}\|^2}{2(1-\alpha)}.$$

Pela definição de g_1 , e como $\|\hat{h}\| = 1$,

$$f_1(e + \alpha \hat{h}) - f_1(e) \leq \alpha g_1' \hat{h} + \frac{\alpha^2}{2(1-\alpha)}.$$

Pelo Lema 5.6 e tomando $\alpha = 0,5$,

$$f_1(e + \alpha \hat{h}) - f_1(e) \leq -0,25, \text{ c.q.d..}$$

$$(ii) f_2(e + \alpha \hat{h}, z) - f_2(e, z) =$$

$$= q \ln [c'(e + \alpha \hat{h}) - z] - \sum_{i=1}^n \ln(e + \alpha \hat{h})_i - \left[q \ln \Delta_2 - \sum_{i=1}^n \ln e_i \right].$$

Desenvolvendo esta igualdade, usando o fato de que

$$\sum_{i=1}^n \ln e_i = \emptyset, \text{ e pelas propriedades do logaritmo,}$$

$$\begin{aligned} f_2(e + \alpha \hat{h}, z) - f_2(e, z) &= \\ &= q \ln \left(1 + \frac{\alpha c' \hat{h}}{\Delta_2} \right) - \sum_{i=1}^n \ln (1 + \alpha \hat{h}_i). \end{aligned}$$

Neste ponto, consideremos $\frac{\alpha c' \hat{h}}{\Delta_2} > -1$, porque caso contrário, $\frac{\alpha c' \hat{h}}{\Delta_2} \leq -1$ implica em $c'(e + \alpha \hat{h}) - z \leq \emptyset$, que pela definição de $f_2(\cdot)$ fornece um resultado desejado.

Assim, como a razão $\frac{\alpha c' \hat{h}}{\Delta_2} > -1$, e pelo fato de $\|\hat{h}\| = 1$ implicar que $|\alpha \hat{h}_i| \leq \alpha < 1$, $i = 1, \dots, n$, podemos usar o Lema 3.6, obtendo

$$\begin{aligned} f_2(e + \alpha \hat{h}, z) - f_2(e, z) &\leq \\ &\leq \frac{q \alpha c' \hat{h}}{\Delta_2} - \sum_{i=1}^n \alpha \hat{h}_i + \sum_{i=1}^n \frac{(\alpha \hat{h}_i)^2}{2(1-\alpha)}. \end{aligned}$$

Resolvendo os somatórios, e ajeitando,

$$f_2(e + \alpha \hat{h}, z) - f_2(e, z) \leq$$

$$\leq \alpha \left(\frac{qc}{\Delta_2} - e \right) \|\hat{h}\| + \frac{\alpha^2 \|\hat{h}\|^2}{2(1-\alpha)}.$$

Pela definição de g_2 , e como $\|\hat{h}\| = 1$,

$$f_2(e + \alpha\hat{h}, z) - f_2(e, z) \leq \alpha g_2 \|\hat{h}\| + \frac{\alpha^2}{2(1-\alpha)}.$$

Pelo Lema 5.6 e tomando $\alpha = 0,5$, finalizamos a demonstração. ■

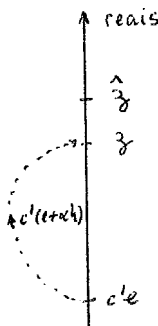
Agora imaginemos uma situação onde estamos no ponto e , no problema escalado de (PA), tal que $c'e - z \leq 0$, para um certo limitante inferior disponível. Neste caso, $f_2(\dots)$ não faz sentido, isto é, $f_2(e, z) = -\infty$. Uma pergunta natural é: para onde ir, a partir do ponto e , quando $\Delta_2 = c'e - z \leq 0$?

2.º Caso: Considere $\Delta_2 \leq 0$, onde $\Delta_2 = c'e - z$ para um limitante inferior z .

Aqui uma direção razoável será aquela que mantém a função potencial $f_2(\dots)$ em $-\infty$ ($\Delta_2 \leq 0$) e reduz $f_1(\dots)$. Isto pode incorrer em passos curtos. No capítulo 6 melhoraremos este comportamento.

Como manter $\Delta_2 \leq 0$?

Considere um passo a partir do ponto e , no problema escalado de (PA), dado por $e + \alpha h$, onde α o tamanho do passo numa direção h dada, para um limitante inferior z disponível. Consideremos $c'e - z \leq 0$ e z o valor ótimo do problema (PA). Então queremos atingir um custo $c'(e + \alpha h)$ tal que $c'(e + \alpha h) \leq z$. Veja figura a seguir.



- Figura 5.3 -

Isto pode ser obtido pela seguinte restrição

$$c'h \leq -\Delta_2, \quad (5.3)$$

onde $c \in \mathbb{R}^n$ é o vetor custo do problema (PA) e $\Delta_2 = c'e - z \leq 0$ para um dado limitante inferior z .

Como reduzir $f_1(\cdot)$?

Tomando um problema (cujo significado bem sabemos; conforme Lema 3.5),

$$\begin{aligned} &\text{minimizar} && g_1' h \\ &\text{sujeito a:} && Ah = 0 \\ &&& \|h\| = 1. \end{aligned}$$

Aqui g_1 é o gradiente de $f_1(\cdot)$ no ponto e , $A \in \mathbb{R}^{m \times n}$ é uma matriz no problema (PA) escalado e $\|h\| = 1$ porque $g_1 \neq 0$ pelo Lema 5.1.

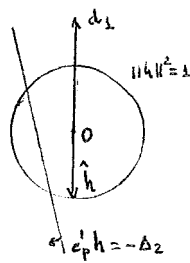
Direção

Unindo a restrição (5.3) no problema acima, tomando $\|h\|^2 = 1$ e lembrando que $h \in N(A)$, obtemos o problema

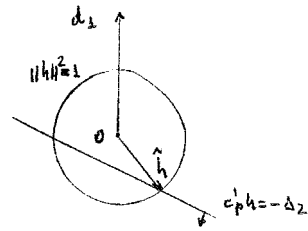
$$\begin{aligned} \text{(EDA)}_{\Delta_2} &= \text{minimizar } d_1' h \\ &\text{sujeito a: } \|h\|^2 = 1 \\ & c_p' h \leq -\Delta_2. \end{aligned}$$

Considere \hat{h} uma solução de (EDA)_{Δ_2} . A geometria do problema (EDA)_{Δ_2} está representada nas figuras abaixo. Observe que se $c_p' \left(\frac{-d_1}{\|d_1\|} \right) \leq -\Delta_2$, então $\hat{h} = \frac{-d_1}{\|d_1\|}$ porque

\hat{h} é justamente a direção de máximo declive. Caso contrário, também podemos obter uma expressão para \hat{h} , onde $c_p' \hat{h} = -\Delta_2$.



(a)



(b)

- Figuras 5.4 -

O próximo lema fornece a solução do problema $(EDA)_{\Delta_2}$ explicitamente.

Lema 5.8: Considere h uma solução do problema $(EDA)_{\Delta_2}$.

Então:

$$(i) \quad h = \frac{-d_1}{\|d_1\|}, \text{ ou}$$

$$(ii) \quad h = - \frac{\sqrt{1 - \left(\frac{\Delta_2}{\|c_p\|} \right)^2}}{\|d_r\|} d_r - \frac{\Delta_2}{\|c_p\|^2} c_p,$$

onde

$$d_r = d_1 - \frac{c_p' d_1}{\|c_p\|^2} c_p.$$

Demonstração: Demonstraremos (i) e (ii) juntos. Pelas condições de Karush-Kuhn-Tucker, h é solução ótima de $(EDA)_{\Delta_2}$ se, e só se, existem $\alpha, \beta \in \mathbb{R}$ tais que

$$d_1 + 2\alpha h + \beta c_p = 0 \quad (5.4)$$

$$\beta \geq 0 \quad (5.5)$$

$$\|h\|^2 = 1 \quad (5.6)$$

$$c_p' h \leq -\Delta_2 \quad (5.7)$$

$$(c_p' h + \Delta_2)\beta = 0. \quad (5.8)$$

1.º caso: $\beta = 0$. Então, de (5.4) segue-se que existe um α , $\alpha = \frac{\|d_1\|}{2}$, tal que

$$h = \frac{-d_1}{\|d_1\|},$$

concluindo (i).

2.º caso: $\beta > 0$. Então, pela condição (5.8),

$$c_p' h = -\Delta_2. \quad (5.9)$$

Desta forma podemos escrever d_1 como o vetor soma da projeção de d_1 sobre e_p mais um vetor, d_r , no nulo de c_p , (isto é, $d_r \perp c_p$), dado por

$$d_1 = d_r + \frac{c_p' d_1}{\|c_p\|^2} c_p, \quad (5.10)$$

e podemos substituir (5.10) na condição (5.4) tal que

$$2\alpha h + \beta c_p = -d_r - \frac{c_p' d_1}{\|c_p\|^2} c_p.$$

Desenvolvendo esta igualdade obtemos,

$$h = -\frac{1}{2\alpha} d_r - \left(\frac{\beta \|c_p\|^2 + c_p' d_1}{2\alpha \|c_p\|^2} \right) c_p.$$

$$\text{Tomando } \mu = \frac{1}{2\alpha} \text{ e } \nu = \frac{\beta \|c_p\|^2 + c_p' d_1}{2\alpha \|c_p\|^2},$$

$$h = -\mu d_r - \nu c_p. \quad (5.11)$$

Pela condição (5.6), usando (5.11), e pela ortogonalidade de d_r e c_p ,

$$1 = \mu^2 \|d_r\|^2 + \nu^2 \|c_p\|^2. \quad (5.12)$$

Pelas restrições (5.9) e (5.11),

$$c_p' (-\mu d_r - \nu c_p) = -\mu c_p' d_r - \nu \|c_p\|^2 = -\Delta_2.$$

Como $d_r \perp c_p$, segue-se que

$$\nu = \frac{\Delta_2}{\|c_p\|^2}. \quad (5.13)$$

Também, de (5.12) e (5.13),

$$\mu = \frac{\sqrt{1 - \frac{\Delta_2^2}{\|c_p\|^2}}}{\|d_r\|}. \quad (5.14)$$

Finalmente, de (5.11), (5.13) e (5.14),

$$h = - \frac{\sqrt{1 - \left[\frac{\Delta_2}{\|c_p\|} \right]^2}}{\|d_r\|} d_r - \frac{\Delta_2}{\|c_p\|^2} c_p,$$

onde

$$d_r = d_1 - \frac{c_p' d_1}{\|c_p\|^2} c_p,$$

por (5.10), finalizando a demonstração. ■

Deste lema podemos enunciar o seguinte procedimento:

Procedimento 5.9: Solução para $(EDA)_{\Delta_2}$.

Dados: c_p, d_r e $\Delta_2 \leq 0$.

Teste:

Se $c_p' \left(\frac{-d_1}{\|d_1\|} \right) \leq -\Delta_2$, então tome

$$h = - \frac{d_1}{\|d_1\|}$$

Senão, tome

$$h = - \sqrt{1 - \left(\frac{\Delta_2}{\|c_p\|} \right)^2} \frac{d_r}{\|d_r\|} - \frac{\Delta_2}{\|c_p\|^2} c_p. \quad \square$$

Segue-se um lema auxiliar para a demonstração do próximo teorema.

Lema 5.10: Considere \hat{h} a solução do problema $(EDA)_{\Delta_2}$.

Então,

$$d_1' \hat{h} \leq -\|d_r\|,$$

onde

$$d_r = d_1 - \frac{c_p' d_1}{\|c_p\|^2} c_p.$$

Demonstração: Podemos tomar $h = -\frac{d_r}{\|d_r\|}$ uma solução viável, pois $\|h\|^2 = 1$ e $c_p' h = 0 \leq -\Delta_2$, pela ortogonalidade entre d_r e c_p . Então,

$$d_1' \hat{h} \leq d_1' h = -\|d_r\|,$$

já que $d_1 = d_r + d_c$, onde $d_c \perp d_r$. ■

Teorema 5.11: Considerere \hat{h} a solução do problema $(EDA)_{\Delta_2}$.

Então:

(i) Se $q \geq n + \sqrt{n}$, então $d_1' \hat{h} \leq -0,5$; e

(ii) $c'(\hat{e} + \alpha \hat{h}) \leq z$, para qualquer $\alpha \in (0,1)$, onde z é o limitante inferior fornecido pelo Procedimento 4.2.

Demonstração: (i) Se $\hat{h} = \frac{-d_1}{\|d_1\|}$, então a função objetivo de $(EDA)_{\Delta_2}$, obtém

$$d_1' \hat{h} \leq -1 \leq -0,5,$$

pelo Lema 5.1. Consideremos, então,

$$\hat{h} = -\sqrt{1 - \left(\frac{\Delta_2}{\|c_p\|}\right)^2} \frac{d_r}{\|d_r\|} - \frac{\Delta_2}{\|c_p\|^2} c_p. \quad (5.15)$$

onde

$$d_r = d_1 - \frac{c_p' d_1}{\|c_p\|^2} c_p \quad (5.16)$$

Por absurdo, vamos supor que $d_1' h \hat{=} -0,5$.

Inicialmente, vamos estabelecer dois resultados úteis:

(i) Multiplicando escalarmente (5.15) por d_1 ,

$$d_1' h \hat{=} - \sqrt{1 - \left[\frac{\Delta_2}{\|c_p\|} \right]^2} \|d_r\| - \frac{\Delta_2}{\|c_p\|^2} c_p' d_1,$$

pois $d_1' d_r = \|d_r\|^2$. Conseqüentemente,

$$-0,5 < d_1' h \hat{=} \leq \frac{-\Delta_2}{\|c_p\|^2} c_p' d_1 \quad (5.17)$$

(ii) Pelo Lema 5.10 e pela hipótese de absurdo,

$$-0,5 < d_1' h \hat{=} \leq -\|d_r\|. \quad (5.18)$$

Agora vamos construir uma variável dual viável a partir da hipótese de absurdo. De (5.18), obtemos $\|d_r\| < 0,5$ e portanto $e + d_r \hat{=} > 0$. Mas por (5.16),

$$d_r = \frac{-c'_p d_1}{\|c_p\|^2} c_p + \frac{q}{\Delta_1} \zeta_p - e_p.$$

Segue-se que

$$\begin{aligned} s &= \frac{\|c_p\|^2}{-c'_p d_1} (e + d_r) \\ &= c_p + \frac{\|c_p\|^2 q}{-c'_p d_1 \Delta_1} \zeta_p + \frac{\|c_p\|^2}{-c'_p d_1} (e - e_p) \end{aligned}$$

é uma folga dual viável, no formato (4.1), com

$$\tau = \frac{\|c_p\|^2 q}{-c'_p d_1 \Delta_1}.$$

Note que $c'_p \left(\frac{-d_1}{\|d_1\|} \right) > -\Delta_2$. Logo, $-c'_p d_1 > 0$.

O gap associado a s é dado por

$$\Delta(s) = e's - \tau \Delta_1.$$

Voltando à definição de s ,

$$s = \frac{\|c_p\|^2}{-c'_p d_1} (e + d_r)$$

e portanto,

$$\Delta(s) = \frac{\|c\|_p^2}{-c'_p d_1} (n + e'_r d_r) - \frac{\|c\|_p^2 q}{-c'_p d_1}.$$

Multiplicando ambos os lados da igualdade acima por menos um,

$$\begin{aligned} -\Delta(s) &= \frac{\|c\|_p^2}{-c'_p d_1} (q - n - e'_r d_r) \geq \\ &\geq \frac{\|c\|_p^2}{-c'_p d_1} (\sqrt{n} - \sqrt{n} \|d_r\|), \end{aligned}$$

pois $q - n \geq \sqrt{n}$ e $-e'_r d_r \geq -\sqrt{n} \|d_r\|$. Usando (5.18), $-0,5 < -\|d_r\|$, e portanto

$$-\Delta(s) > 0,5 \sqrt{n} \frac{\|c\|_p^2}{-c'_p d_1}.$$

Finalmente, usando (5.17), obtemos

$$0,5 \frac{\|c\|_p^2}{-c'_p d_1} > -\Delta_2.$$

Unindo as duas últimas expressões,

$$-\Delta(s) > -\sqrt{n} \Delta_2,$$

e como $\sqrt{n} \geq 1$, segue-se que

$$\Delta(s) < \Delta_2,$$

contradizendo a definição de Δ_2 . Isto conclui a demonstração do item (i).

(ii) Considere o problema $(EDA)_{\Delta_2}$. Seja $\Delta_2 = c'e - z \leq 0$, onde z é o limitante inferior obtido pelo Procedimento 4.2. A restrição de desigualdade de $(EDA)_{\Delta_2}$ fornece para $\hat{h} \in N(A)$,

$$\alpha c' \hat{h} \leq -\Delta_2,$$

para qualquer $\alpha \in (0,1)$, já que $\Delta_2 = c'e - z \leq 0$. Segue-se que

$$c'(e + \alpha \hat{h}) \leq z,$$

para qualquer $\alpha \in (0,1)$, finalizando a demonstração. ■

Observe neste teorema que se $q \geq n + \sqrt{n} + 1$, então obtemos $d_1' \hat{h} \leq -1$.

Novamente estamos prontos para enunciar um teorema fundamental à análise de convergência do algoritmo afim.

Teorema 5.12: Considere \hat{h} uma solução do problema $(EDA)_{\Delta_2}$ e $q \geq n + \sqrt{n}$. Então,

$$(i) \quad f_1(e + 0,2 \hat{h}) - f_1(e) \leq -0,075; e$$

(ii) $c'(e + 0,2 \hat{h}) \leq z$, onde z é o limitante inferior fornecido pelo Procedimento 4.2.

Demonstração: (i) Análoga ao Teorema 5.7 (i), já que $f_1(\cdot)$ independe do sinal de Δ_2 , e podemos usar o Teorema 5.11(i).

(ii) Imediata. ■

Neste momento cabe um procedimento para encontrar direções para o algoritmo afim.

Procedimento 5.13: Direções

Dado: $\Delta_2 = c'e - z$, onde z é o limitante inferior fornecido pelo Procedimento 4.2.

Teste:

Se $\Delta_2 > 0$, então calcule $(EDA)'$.

Senão, calcule $(EDA)_{\Delta_2}$ pelo Procedimento 5.9. □

Note que neste procedimento estamos num problema escalado de (PA).

Na próxima seção usaremos a mesma estratégia para encontrar direções numa abordagem não afim.

5.2 - Uma abordagem projetiva

Consideremos o problema de programação linear escalado reformulado

$$\begin{aligned}
 \text{(PP)} = & \text{ minimizar } c'x \\
 & \text{sujeito a: } Ax = 0 \\
 & w'x = 1 \\
 & \xi'x = 0 \\
 & x \geq 0.
 \end{aligned}$$

onde $A \in \mathbb{R}^{m \times n}$ é uma matriz de rank máximo, $0 < m < n$. Além disso, consideremos o conjunto viável

$$\text{(SP)} = \left\{ x \in \mathbb{R}_+^n / Ax = 0, w'x = 1, \xi'x = 0 \right\},$$

e o conjunto de trabalho

$$\text{(SP)'} = \left\{ x \in \mathbb{R}_+^n / Ax = 0, w'x = 1 \right\}.$$

Assumimos que na iteração corrente partimos do ponto $e \in \text{(SP)'}$.

Notação: Denotamos $\delta_1 = \xi'e$, $\delta_2 = c'e - z$ para um limitante inferior z , "q" relativo a vetores projetados no espaço

nulo da matriz A_k , e "k" será omitido por se tratar de uma análise sobre uma iteração arbitrária. \square

O restante desta seção assumirá

$$\delta_1 = \xi'e > 0.$$

Pelo Exemplo 3.2, observe que é possível termos $c'e > z$, assim como $c'e \leq z$. Afinal, não estamos viáveis ($\delta_1 > 0$)

Desta forma,

$$\delta_2 = c'e - z \in \mathbb{R},$$

para um limitante inferior z .

A idéia é a mesma da seção anterior. Isto definir uma direção de busca no $N(A)$ que nos aproxime simultaneamente viabilidade e otimalidade.

Definimos duas funções potenciais de Karmarkar no Capítulo 3, a saber:

$$F_1(x) = n \ln \xi'e - \sum_{i=1}^n \ln x_i,$$

e

$$F_2(x, z) = \begin{cases} n \ln (c - zw)'x - \sum_{i=1}^n \ln x_i, & \text{se } \delta_2 > 0, \\ -\infty, & \text{se } \delta_2 \leq 0, \end{cases}$$

onde $x \in (SP)'$ e z é um limitante inferior.

Nosso candidato para aproximar viabilidade está associado à função potencial $F_1(\cdot)$, enquanto para otimalidade à função $F_2(\cdot, \cdot)$.

Os gradientes no ponto e das funções potenciais $F_1(\cdot)$ e $F_2(\cdot, \cdot)$ são dados respectivamente por

$$G_1 = \frac{n \xi}{\delta_1} - e$$

e

$$G_2 = \frac{n (c - zw)}{\delta_2} - e.$$

As projeções sobre $N(A)$ dos gradientes G_1 e G_2 no ponto e definem dois vetores dados por

$$D_1 = \frac{n \xi_q}{\delta_1} - e,$$

$$D_2 = \frac{n (c_q - zw_q)}{\delta_2} - e,$$

onde $e = e_q$.

Segue-se um lema auxiliar para a demonstração do Teorema 5.17 adiante.

Lema 5.14: Considere o problema (PP) e o conjunto (SP)'. Tome $x = e \in (SP)'$, $\delta_1 = \zeta'e > 0$. Então $\|D_1\| \geq 1$.

Demonstração: Aqui vamos explorar a homogeneidade da função potencial de Karmarkar. Seja \hat{x} uma solução viável do problema (PP). Então podemos considerar um ponto

$$\tilde{x} \in \{x \in \mathbb{R}_+^n \mid Ax = 0, e'x = n, \zeta'x = 0\},$$

tal que $\tilde{x} = \lambda \hat{x}$, $\lambda > 0$. Note que $\lambda = \frac{n}{e'\hat{x}} > 0$, pois $e'\tilde{x} = n$ e $e'\hat{x} \neq 0$. Desta forma, uma projeção cônica

$$K(x) = \frac{\tilde{x}}{w'x} = \frac{\lambda \hat{x}}{\lambda w'\hat{x}} = \frac{\hat{x}}{w'\hat{x}} = \hat{x}.$$

Vamos escolher a direção $\bar{H} = \frac{\tilde{x} - e}{\|x - e\|} \in N(A)$. Temos, então,

$$G_1'\bar{H} = G_1' \frac{\tilde{x} - e}{\|x - e\|} = \left(\frac{n \zeta}{\delta_1} - e \right)' \frac{\tilde{x} - e}{\|x - e\|}.$$

Desenvolvendo esta igualdade, usando o fato de $\zeta'\tilde{x} = 0$ e $\|e\|^2 = n$,

$$G_1'\bar{H} = - \frac{e'\tilde{x}}{\|x - e\|}.$$

Mas $x \neq 0$ e, então $e'x = \|x\|_1 \geq \|x\|$. Daí,

$$G_1' H \leq \frac{\|x\|}{\|x - e\|}.$$

Como $e'x = n$ segue-se que $e \perp (x - e)$. Então, obtemos por Pitágoras

$$\|e\|^2 + \|x - e\|^2 = \|x\|^2.$$

Dividindo ambos os lados desta igualdade por $\|x - e\|^2 \neq 0$,

$$\frac{\|x\|^2}{\|x - e\|^2} = \frac{\|e\|^2}{\|x - e\|^2} + 1 \geq 1.$$

Então

$$\frac{\|x\|}{\|x - e\|} \geq 1.$$

Logo, existe uma direção $\vec{H} = \frac{x - e}{\|x - e\|}$ que depende de x tal

que

$$G_1' H \leq -1.$$

Multiplicando esta desigualdade por menos um, e pelo fato de que $\vec{H} \in N(A)$,

$$-D_1' \bar{H} \geq 1.$$

Observe que

$$-D_1' \bar{H} \leq \|D_1\| \|\bar{H}\| = \|D_1\|.$$

Finalmente,

$$\|D_1\| \geq 1. \quad \blacksquare$$

Novamente separaremos nossa estratégia para encontrar direções em dois casos, segundo $\delta_2 > 0$ ou $\delta_2 \leq 0$.

1.º Caso: Considere $\delta_2 > 0$, onde $\delta_2 = c'e - z$ para um limitante inferior z .

Queremos desenvolver uma estratégia para encontrar direções semelhante à da seção anterior. Isto é, mostraremos que as derivadas direcionais de $F_1(\cdot)$ e $F_2(\dots)$ ao longo da direção obtida são menores ou iguais a menos um.

Direção

Como a estratégia para encontrar direções é a mesma da seção anterior, podemos enunciar os seguintes procedimentos:

Procedimento 5.15: Sobre (EDP).

Resolva

$$(EDP) = \underset{\|x\| \leq 1}{\text{minimizar}} \quad \max_{i \in \{1, 2\}} D_i' H$$

Se o valor da solução for zero, faça $\hat{H} = 0$. \square

Procedimento 5.16: Sobre (EDP)'.

Obtenha a solução ótima \hat{D} para o problema

$$(EDP)' = \underset{\substack{D = tD_1 + (1-t)D_2 \\ t \in [0,1]}}{\text{minimizar}} \|D\|$$

$$\text{Faça } \hat{H} = - \frac{\hat{D}}{\|\hat{D}\|} \text{ se } \hat{D} \neq 0.$$

Caso contrário, tome $\hat{H} = 0$. \square

A equivalência entre os Procedimentos 5.15 e 5.16, e a relação do valor da solução de (EDP) como sendo o negativo do valor da solução de (EDP)' é garantido pelo Lema 5.4.

O teorema a seguir relaciona nosso procedimento de atualização de limitantes inferiores do capítulo anterior com nossa direção obtida pelo Procedimento 5.16, com $\delta_2 > 0$.

Teorema 5.17: Considere o problema (PP) e o conjunto (SP)'. Tome $x = e \in (SP)'$, $\delta_1 > 0$ e $\delta_2 = c'e - z > 0$ para um limitante inferior z obtido pelo Procedimento 4.2. Então, $\|tD_1 + (1-t)D_2\| \geq 1$ para todo $t \in [0,1]$.

Demonstração: Faça

$$D = tD_1 + (1-t)D_2, \text{ para } t \in [0,1].$$

Se $t=1$, então $D = D_1$ e, pelo Lema 5.14, $\|D_1\| \geq 1$. Seja então $t \in [0,1)$. Suponha por absurdo que $\|D\| < 1$. Logo, $e + D > 0$, já que $-2 < D_j < 1$, $j = 1, \dots, n$, onde D_j as componentes do vetor D . Vamos construir uma folga dual viável no formato (4.1) usado pela atualização de limitantes inferiores do capítulo anterior: desenvolvendo a expressão para D ,

$$\begin{aligned} D &= t \left[\frac{n \xi_q}{\delta_1} - e \right] + (1-t) \left[\frac{n(c_q - zw_q)}{\delta_2} - e \right] = \\ &= \frac{tn\xi_q}{\delta_1} + \frac{(1-t)n(c_q - zw_q)}{\delta_2} - e. \end{aligned}$$

Segue-se que

$$e + D = \frac{tn \xi_q}{\delta_1} + \frac{(1-t)n(c_q - zw_q)}{\delta_2} > 0.$$

Mas,

$$e + D = \frac{tn \xi_p}{\delta_1} + \frac{(1-t) nc_p}{\delta_2} + \left(\frac{tn}{\delta_1} \mu_1 + \frac{(1-t)n}{\delta_2} \mu_2 - \frac{(1-t)nz}{\delta_2} \right) w_q > 0,$$

pois $\xi_q = \xi_p + \mu_1 w_q$, $E_q = c_p + \mu_2 w_q$, com $\mu_1, \mu_2 \in \mathbb{R}$.

Tomando,

$$\frac{tn \mu_1}{\delta_1} + \frac{(1-t)n}{\delta_2} \mu_2 - \frac{(1-t)nz}{\delta_2} = v,$$

então,

$$e + D = \frac{tn \xi_p}{\delta_1} + \frac{(1-t) nc_p}{\delta_2} + (e - e_p) > 0,$$

já que $e_q = e = e_p + v w_q$. Para obter folga dual viável no formato (4.1), fazemos

$$s = (e + D) \frac{\delta_2}{(1-t)n} = c_p + \frac{t\delta_2}{(1-t)\delta_1} \xi_p + \frac{\delta_2}{(1-t)n} (e - e_p) > 0. \quad (5.19)$$

A expressão acima está no formato (4.1) usado pelo Procedimento 4.2 do capítulo anterior e deve-se ter por definição de δ_2 ,

$$\delta(s) > \delta_2,$$

onde $\delta(s)$ é o *gap* associado ao s de (5.19). Finalmente, vamos calcular $\delta(s)$ e estabelecer uma contradição: temos

$$s = (e + D) \frac{\delta_2}{(1-t)n}$$

O *gap* correspondente é dado por (4.3), onde $\lambda = -\tau$, isto é,

$$\delta(s) = e's - \tau\delta_1,$$

onde $\tau = \frac{t\delta_2}{(1-t)\delta_1}$ é extraído de (5.19). Então,

$$\begin{aligned} \delta(s) &= e's - \tau\delta_1 = e'(e+D) \frac{\delta_2}{(1-t)n} - \frac{t\delta_2}{(1-t)\delta_1} \delta_1 = \\ &= \frac{n\delta_2}{(1-t)n} - \frac{t\delta_2}{(1-t)}, \end{aligned} \quad (5.20)$$

pois,

$$\begin{aligned} e'D &= e' \left\{ t \left[\frac{n\xi_q}{\delta_1} - e \right] + (1-t) \left[\frac{n(c_q - zw_q)}{\delta_2} - e \right] \right\} = \\ &= t \left[\frac{n}{\delta_1} \xi_q' e - n \right] + (1-t) \left[\frac{n(c_q - zw_q)' e}{\delta_2} - n \right] = \\ &= t \left[\frac{n\delta_1}{\delta_1} - n \right] + (1-t) \left[\frac{n\delta_2}{\delta_2} - n \right] = 0. \end{aligned}$$

Portanto, em (5.20),

$$\delta(s) = \delta_2,$$

o que é um absurdo, porque o Procedimento 4.2 resolve um problema de programação linear, cuja solução está na fronteira. Observe que $s > 0$ e $\delta(s)$ não é o menor valor. Isto completa a demonstração por absurdo. ■

Este teorema afirma que não obtemos uma solução $\hat{H} = 0$. Logo, podemos representar a geometria de (EDP) ou (EDP)' caracterizando três situações, conforme Figuras 5.2, apresentada na seção anterior.

Ainda pelo Teorema 5.17, podemos enunciar (EDP) na seguinte forma:

$$\begin{aligned} \overline{\text{(EDP)}} = & \text{minimizar} & \max & D_i' H. \\ & \|H\| = 1 & i \in \{1, 2\} \end{aligned}$$

As soluções ótimas \hat{H} de $\overline{\text{(EDP)}}$ e \hat{D} de (EDP)' estão relacionadas por $\hat{H} = -\frac{\hat{D}}{\|\hat{D}\|}$.

Devemos mostrar que a direção obtida ao resolvermos $\overline{\text{(EDP)}}$ (ou (EDP)' que é mais conveniente) reduz as funções potenciais $F_1(\cdot)$ e $F_2(\cdot, \cdot)$ bastante. Trataremos disto no lema a seguir.

Lema 5.18: Considere o problema (PP) e o conjunto $(SP)'$. Tome $x = e \in (SP)'$, $\delta_1 = \xi'e > 0$, $\delta_2 = c'e - z > 0$ para um limitante inferior $z \leq \hat{z}$ onde \hat{z} é o valor ótimo de (PP) e \hat{z} é obtido pelo Procedimento 4.2, e tome \hat{H} a direção obtida resolvendo (\overline{EDP}) . Então,

$$(i) \quad G_1^{\hat{H}} \leq -1; \quad e$$

$$(ii) \quad G_2^{\hat{H}} \leq -1.$$

Demonstração: (i) e (ii): Análogas às do Lema 5.6. ■

Agora estamos prontos para demonstrar um teorema fundamental à análise de convergência do algoritmo projetivo.

Teorema 5.19: Considere as hipóteses do Lema 5.18. Então existe um α , $\alpha \in (0,1)$ tal que

$$(i) \quad F_1(e + \alpha \hat{H}) - F_1(e) \leq -0,25; \quad e$$

$$(ii) \quad F_2(e + \alpha \hat{H}, z) - F_2(e, z) \leq -0,25.$$

Demonstração: (i) e (ii): Análogas às do Teorema 5.7. ■

Agora analisemos o 2.^o caso.

2.º Caso: Considere $\delta_2 \leq \theta$, onde $\delta_2 = c'_q e - z$ para um limitante inferior z .

Direção

Aqui desenvolvemos a mesma estratégia da seção anterior, cujo comportamento pode ser melhorado através da busca unidirecional no próximo capítulo, já que reduzir $F_1(\cdot)$ e manter $F_2(\dots)$ em $-\infty$ ($\delta_2 \leq \theta$) pode incorrer em passos curtos. Então podemos definir diretamente o problema para encontrar direções

$$\begin{aligned} \text{(EDP)}_{\delta_2} &= \text{minimizar } D_1' H \\ &\text{sujeito a: } \|H\|^2 = 1 \\ & \quad c'_q H \leq -\delta_2. \end{aligned}$$

Sua solução foi calculada no Lema 5.8 da seção anterior. Então podemos enunciar um procedimento semelhante.

Procedimento 5.20: Solução para (EDP)_{δ_2} .

Dados: c'_q , D_1 e $\delta_2 \leq \theta$.

Teste:

$$\text{Se } c'_q \left(\frac{-D_1}{\|D_1\|} \right) \leq -\delta_2, \text{ então tome } H = \frac{-D_1}{\|D_1\|}.$$

$$\text{Senão, tome } H = -\sqrt{1 - \left(\frac{\delta_2}{\|c'_q\|} \right)^2} \frac{D_r}{\|D_r\|} - \frac{\delta_2}{\|c'_q\|^2} c'_q,$$

onde

$$D_r = D_1 - \frac{c'_q D_1}{\|c'_q\|^2} c'_q. \quad \square$$

Uma vez que necessitamos garantir que a nossa direção é razoável formalmente, segue-se um teorema, o qual não desmonstraremos aqui. Deve ser possível demonstrá-lo pelo mesmo caminho do Teorema 5.11 quando do caso afim. Além disso, é possível deduzir a validade deste resultado a partir da análise feita por Todd [12].

Teorema 5.21: Considere \hat{H} a solução do problema (EDP) $_{\delta_2}$.

Então:

(i) $D'_1 \hat{H} \leq -0,5$; e

(ii) $c'(\epsilon + \alpha \hat{H}) \leq z$, para qualquer $\alpha \in (0,1)$, onde z é o limitante inferior fornecido pelo Procedimento 4.2. \square

Novamente estamos prontos para enunciar um teorema fundamental à análise de convergência do algoritmo projetivo.

Teorema 5.22: Considere \hat{H} uma solução do problema (EDP) $_{\delta_2}$.

Então:

$$(i) F_1(e + 0,2H) - F_1(e) \leq -0,075; e$$

$$(ii) c'(e + 0,2H) \leq z, \text{ onde } z \text{ é o limitante inferior fornecido pelo Procedimento 4.2.}$$

Demonstração: (i) Análoga ao Teorema 5.19 (i), já que $F_1(\cdot)$ independe do sinal de δ_2 , e podemos usar o Teorema 5.21 (i).

(ii) Imediata. ■

Neste momento cabe um procedimento para encontrar direções para o algoritmo projetivo.

Procedimento 5.23: Direções.

Dados: $\delta_2 = c'e - z$, onde z é o limitante inferior fornecido pelo Procedimento 4.2.

Teste:

Se $\delta_2 > 0$, então calcule $(EDP)'$.

Senão calcule $(EDP)_{\delta_2}$ pelo Procedimento 5.20. □

Note que neste procedimento estamos num problema escalado de (PP).

Na próxima seção demonstraremos uma equivalência existente entre as direções de Todd [14, 12] - veja também

Todd e Wang [13] - e as nossas, quando $\Delta_2 \succ \emptyset$ na formulação afim, e quando $\delta_2 \succ \emptyset$ na formulação projetiva.

5.3 - Equivalência sob certa condição

Aqui vamos mostrar uma equivalência entre as direções de Todd e as nossas, tanto para algoritmos afins quanto para projetivos, quando $\Delta_2 \succ \emptyset$ no primeiro caso, e $\delta_2 \succ \emptyset$ no segundo. O símbolo " $\hat{\cdot}$ " designará uma solução ótima do problema envolvido (se existir).

Abordagem afim

Inicialmente consideremos uma abordagem afim, em que assumimos feita uma mudança de escala, $\Delta_1 = \xi' e \succ \emptyset$ e $\Delta_2 = c'e - z \succ \emptyset$ para um limitante inferior z , fornecido pelo Procedimento 4.2. Sejam d_1 e d_2 como definimos anteriormente.

A direção de Todd [14], denotamos \hat{h}_T , é uma direção obtida resolvendo o seguinte problema:

$$\begin{aligned} \text{(SP):} \quad & \text{minimizar } \|\hat{h}_T\| \\ & \text{sujeito a: } \hat{h}_1' \hat{h}_T \leq -\Delta_1 \\ & \hat{h}_2' \hat{h}_T \leq -\Delta_2, \end{aligned}$$

onde $\hat{h}_1 = \Delta_1 d_1$ e $\hat{h}_2 = \Delta_2 d_2$ ($\Delta_2 \succ \emptyset$). Omitimos em (SP) uma restrição de monotonicidade,

$$c_p' h_T \leq 0 \quad (\text{se } \Delta_2 > 0),$$

apesar de podermos introduzi-la no nosso problema a seguir.

Consideremos nossa direção, denotada \hat{h} , para $\Delta_2 > 0$, dada pela resolução do problema

$$\begin{aligned} \text{(EDA)'} &= \text{minimizar } \|d\| \\ &\text{sujeito a: } d = td_1 + (1-t)d_2 \\ &\quad t \in [0,1]. \end{aligned}$$

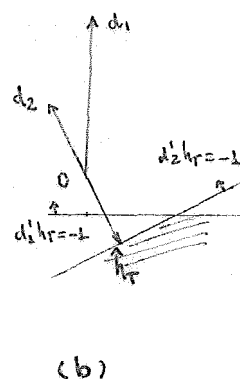
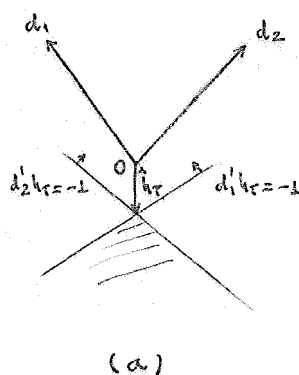
$$\text{Tome } \hat{h} = \frac{-d}{\|d\|}.$$

Como $\Delta_1, \Delta_2 > 0$, então dividindo ambos os lados das duas primeiras restrições em (SP) por Δ_1 e Δ_2 , respectivamente, obtemos

$$\begin{aligned} \overline{\text{(SP)}} &= \text{minimizar } \|h_T\| \\ &\text{sujeito a: } d_1' h_T \leq -1 \\ &\quad d_2' h_T \leq -1. \end{aligned}$$

Observe que $\hat{h}, \hat{h}_T \in N(A)$ já que d_1 e d_2 os pertence (em $\overline{\text{(SP)}}$ e (EDA)').

A geometria do problema $\overline{\text{(SP)}}$ é dada nas figuras a seguir.



- Figuras 5.5 -

Ambas as Figuras 5.5(a) - (b), refletem a menor distância da origem ao conjunto formado pela interseção dos semi-espacos $d_i/h_T \leq -1$, $i \in \{1,2\}$.

Segue-se um teorema que associa as direções h_T e h pela equivalência dos problemas (\overline{SP}) e (\overline{EDA}) (o problema (\overline{EDA}) surge como consequência do Teorema 5.5.).

Teorema 5.24: Considere $\Delta_2 > 0$. Então (\overline{EDA}) é equivalente a (\overline{SP}) .

Demonstração: Partindo de (\overline{EDA}) , e tomando $\sigma = \max_{i \in \{1,2\}} d_i/h_T$

obtemos a equivalência

(T) = minimizar σ

sujeito a: $d_i' h \leq \sigma, \quad i = 1, 2$

$$\|h\| = 1.$$

Agora, partindo de (\overline{SP}) , e tomando

$$h_T = \gamma h \text{ com } \|h\| = 1 \text{ e } \gamma > 0,$$

isto é $\gamma = \|h_T\|$, segue-se a equivalência

minimizar γ

sujeito a: $d_1' h \leq -\frac{1}{\gamma}$

$$d_2' h \leq -\frac{1}{\gamma}$$

$$\|h\| = 1.$$

Como a função $\gamma \in \mathbb{R}_{++} \rightarrow -\frac{1}{\gamma}$ é monotônica não decrescente, pelo Lema 3.7, este último problema é equivalente a

minimizar $-\frac{1}{\gamma}$

sujeito a: $d_1' h \leq -\frac{1}{\gamma}$

$$d_2' h \leq -\frac{1}{\gamma}$$

$$\|h\| = 1.$$

Este último problema é idêntico ao problema (T), com $\sigma = -\frac{1}{\gamma}$.

Isto finaliza a demonstração. ■

Abordagem projetiva

Finalmente, consideremos uma abordagem projetiva, assumindo feita uma mudança de escala, $\delta_1 = \xi' e > 0$ e $\delta_2 = (c'e - z) > 0$, para um limitante inferior z obtido pelo Procedimento 4.2. Sejam D_1 e D_2 como definimos anteriormente.

A direção de Todd [12] - veja também Todd e Wang [13] -, denotamos \hat{H}_T , é uma direção obtida resolvendo o seguinte problema:

$$\begin{aligned}
 \text{(DFSP):} \quad & \text{minimizar } \|H_T\| \\
 & \text{sujeito a: } AH_T = 0 \\
 & e'H_T = 0 \\
 & \xi'H_T \leq -\delta_1 \\
 & (c-zw)'H_T \leq -\delta_2.
 \end{aligned}$$

Omitimos em (DFSP) uma restrição de monotonicidade,

$$(c - (c'e)w)'H_T \leq 0 \quad (\text{se } \delta_2 > 0),$$

apesar de podermos introduzi-la no nosso problema a seguir.

Consideremos nossa direção, denotada \hat{H} , para $\delta_2 > 0$, dada pela resolução do problema

(EDP)' : minimizar $\|D\|$

$$\text{sujeito a: } D = tD_1 + (1-t)D_2$$

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

$$\text{Tome } \hat{H} = - \frac{\hat{D}}{\|\hat{D}\|}.$$

Partindo de $\xi'H_T \leq -\delta_1$ e $(c-zw)'H_T \leq -\delta_2$, onde $\delta_1, \delta_2 > 0$, e pelo fato de $AH_T = 0$ e $e'H_T = 0$, obtemos por construção, $D_1'H_T \leq -n$ e $D_2'H_T \leq -n$. Escalando estas desigualdades por $n > 0$, segue-se um problema

(DFSP) : minimizar $\|H_T\|$

$$\text{sujeito a: } D_1'H_T \leq -1$$

$$D_2'H_T \leq -1.$$

Agora, observando que o problema (DFSP) fornece uma solução $H_T \in N(A)$, e escalando suas restrições por $n > 0$, segue-se as desigualdades

$$\left[\frac{n\xi}{\delta_1} - e \right]'H_T \leq -n \quad \text{e} \quad \left[\frac{n(c-zw)}{\delta_2} - e \right]'H_T \leq -n.$$

Ora, $\xi'H_T \leq -\delta_1$ e $(c-zw)'H_T \leq -\delta_2$ se, e somente se, $e'H_T = 0$. Finalmente podemos concluir que os problemas (DFSP) e (DFSP) são equivalentes.

Observe que $\hat{H}, \hat{H}_T \in N(A)$ já que D_1 e D_2 os pertence (em (EDP)' e (DFSP)).

A geometria do problema (\overline{DFSP}) é semelhante às das Figuras 5.5, dada anteriormente.

Segue-se um teorema que associa as direções H_T e H pela equivalência dos problemas (\overline{DFSP}) e (\overline{EDP}) (o problema (\overline{EDP}) surge como consequência do Teorema 5.17).

Teorema 5.25: Considere $\delta_2 > 0$. Então (\overline{EDP}) é equivalente a (\overline{DFSP}) .

Demonstração: Análoga ao Teorema 5.24.

CAPÍTULO 6

BUSCA UNIDIRECIONAL

No Capítulo 5 desenvolvemos um procedimento para encontrar direções. Sua obtenção distingue-se pelo sinal do parâmetro $\Delta_2 = c'e - z$ (ou δ_2), onde z é um limitante inferior fornecido pelo Procedimento 4.2. Isto é, quando $\Delta_2 > 0$ ($\delta_2 > 0$) trabalhamos com as funções potenciais $f_1(\cdot)$ e $f_2(\cdot, \dots)$ ($F_1(\cdot)$ e $F_2(\cdot, \dots)$), e adotamos uma estratégia que procura reduzir simultaneamente ambas as funções. Quando $\Delta_2 \leq 0$ ($\delta_2 \leq 0$), $f_2(\cdot, \dots) = -\infty$ ($F_2(\cdot, \dots) = -\infty$), adotamos uma estratégia que se preocupa por um lado em reduzir $f_1(\cdot)$ ($F_1(\cdot)$) e, por outro, em manter $\Delta_2 \leq 0$ ($\delta_2 \leq 0$), para passos de comprimento menor ou igual a um.

Uma busca unidirecional determina o tamanho do passo que daremos a partir de um ponto (neste caso o vetor e), numa direção h (ou H) sem violar viabilidade (em (SAD)' - em (SP)').

Nossa estratégia aqui será essencialmente diferente de Anstreicher [2,3] e Todd [12, 13, 14], os quais condenam $\Delta_2 \leq 0$ ($\delta_2 \leq 0$) para sempre. Aumentamos o tamanho do passo em relação ao deles sem comprometer a convergência dos algoritmos que estamos construindo parte por parte neste trabalho. Além disso, condenar $\Delta_2 \leq 0$ ($\delta_2 \leq 0$) durante todo o restante do algoritmo, pode resultar em passos curtos.

6.1 - Uma Abordagem Afim

Consideremos as funções potenciais definida anteriormente,

$$f_1(x) = q \ln c'x - \sum_{i=1}^n \ln x_i,$$

$$f_2(x, z) = \begin{cases} q \ln(c'x - z) - \sum_{i=1}^n \ln x_i, & \text{se } c'x - z > 0, \\ -\infty, & \text{se } c'x - z \leq 0, \end{cases}$$

onde $q \geq n + \sqrt{n}$ e z é um limitante inferior dado.

Faremos nossa análise da busca considerando o sinal do parâmetro Δ_2 .

1.º caso: Considere $\Delta_2 > 0$, onde $\Delta_2 = c'e - z$ para um limitante inferior z .

Neste caso queremos reduzir as funções potenciais $f_1(\cdot)$ e $f_2(\cdot, \cdot)$ simultaneamente, ao longo de uma direção h , a partir do ponto e .

O estudo de convergência é feito para passo de 0,2. Desta forma a busca unidirecional deve ser

um procedimento com $O(i)$ iterações, cuja finalidade é obter um decréscimo de

$$\max\left\{f_1(e + \alpha h) - f_1(e), f_2(e + \alpha h, z) - f_2(e, z)\right\} \quad (6.1)$$

superior ao associado a $\alpha = 0,2$. Note que (6.1) é equivalente a (veja Algoritmo 3.3)

$$\max\left\{f_1(e + \alpha h) - f_1(e) + f_2(e, z), f_2(e + \alpha h, z)\right\}.$$

Para $\alpha = 0$ em (6.1), as diferenças associadas tanto a $f_1(\cdot)$ quanto a $f_2(\cdot, \cdot)$ são nulas. Pelo Teorema 5.7, para $\alpha = 0,5$,

$$f_1(e + \alpha h) - f_1(e) \leq -0,25,$$

$$f_2(e + \alpha h, z) - f_2(e, z) \leq -0,25.$$

Isto é, por iteração $f_1(\cdot)$ e $f_2(\cdot, \cdot)$, ao longo da direção h , decrescem pelo menos 0,25. Portanto, a nossa escolha acerca da busca é dada pela resolução de um problema *min-max*,

$$(BA) = \text{minimizar } \max \left\{ f_1(e+\alpha h) - f_1(e), f_2(e+\alpha h, z) - f_2(e, z) \right\}$$

sujeito a: $e + \alpha h > 0$
 $\alpha \geq 0.$

O problema acima está bem definido. Vejamos por quê: Se $h > 0$, então a busca é feita na semi-reta $e + \alpha h$, $\alpha \geq 0$. Anstreicher prova (ver demonstração do Teorema 7.2 adiante) que a região em que $f_1(x) \leq f_1(e)$ e $f_2(x, z) \leq f_2(e, z)$, se $x \in (SA)'$, é limitada e portanto o problema (BA) está bem definido. Finalmente, se h tem alguma componente negativa, então seja $\bar{\alpha} = \max \{ \alpha \mid e + \alpha h \geq 0 \text{ e } \alpha \geq 0 \}$. Para $\alpha \in [0, \bar{\alpha})$ considere

$$f_1(e + \alpha h) = q \ln \xi'(e + \alpha h) + p(e + \alpha h),$$

$$f_2(e + \alpha h, z) = q \ln [c'(e + \alpha h) - z] + p(e + \alpha h).$$

Se $\xi'(e + \bar{\alpha}h) > 0$ então $\lim_{\alpha \rightarrow \bar{\alpha}} f_1(e + \alpha h) = +\infty$. Por outro

lado, se $c'(e + \bar{\alpha}h) - z > 0$ então $\lim_{\alpha \rightarrow \bar{\alpha}} f_2(e + \alpha h, z) = +\infty$.

Em qualquer destes casos,

$$\lim_{\alpha \rightarrow \bar{\alpha}} \max \{ f_1(e + \alpha h) - f_1(e), f_2(e + \alpha h, z) - f_2(e, z) \} = +\infty.$$

(6.2)

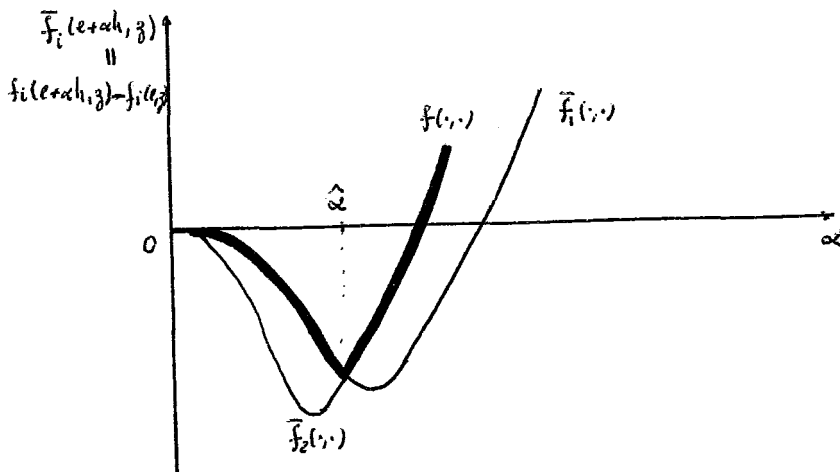
Então, a função $\max(\cdot)$ tem mínimo em $[0, \bar{\alpha})$. Para que (6.2) não seja verdade, é necessário que

$$\xi'(e + \bar{\alpha}h) = 0,$$

$$c'(e + \bar{\alpha}h) - z = 0,$$

que é justamente quando $e + \bar{\alpha}h$ é a solução ótima do problema (PA). Neste caso, o algoritmo pode parar com a solução ótima.

A Figura 6.1, mostra a geometria do problema unidimensional (BA), onde $\hat{\alpha}$ é a solução de (BA).



- Figura 6.1 -

Mostraremos adiante como se pode usar o método de Seção Áurea para uma busca aproximada sobre (6.1).

2.º caso: Considere $\Delta_2 \leq 0$, onde $\Delta_2 = c'e - z$ para um limitante inferior z .

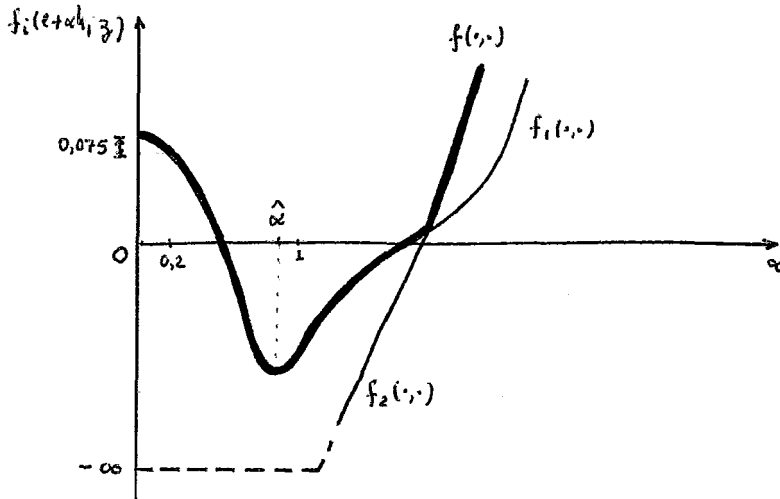
Neste caso queremos reduzir a função potencial $f_1(\cdot)$ e possibilitar mudança de sinal no parâmetro Δ_2 . Desta forma podemos continuar dando passos longos.

Uma vez que o Teorema 5.12 garante um comportamento razoável para $\alpha = 0,2$, uma estratégia interessante para a busca é a seguinte:

$$\begin{aligned} (BA)_{\Delta_2} = \text{minimizar} \quad & \max \left\{ f_1(e + \alpha h), f_2(e + \alpha h, z) \right\} \\ \text{sujeito a:} \quad & e + \alpha h > 0 \\ & \alpha \geq 0. \end{aligned}$$

O problema $(BA)_{\Delta_2}$ está bem definido pela mesma análise feita para o problema (BA) .

Na Figura abaixo sejam $i \in \{1, 2\}$ e $\hat{\alpha}$ uma solução ótima do problema $(BA)_{\Delta_2}$. Observe que a função potencial $f_1(\cdot)$ decresce ao longo da direção h pelo menos $0,075$ para $\alpha = 0,2$. E, a função $f_2(e + \alpha h, z) = -\infty$, para $\alpha \in (0, 1)$ (Teorema 5.12).



- Figura 6.2 -

O algoritmo de busca unidirecional definitivo ($\Delta_2 > 0$ ou $\Delta_2 \leq 0$), com a regra de parada é dado a seguir. A idéia é resolver aproximadamente os problemas (BAD) e $(\text{BAD})_{\Delta_2}$, com $O(1)$ contagens. Utilizamos o método de Seção Áurea, simplesmente. Ou seja, poderíamos usar qualquer método de busca a valores finitos com $O(1)$ contagens.

Antes, porém, definimos o seguinte teste e um critério de parada do algoritmo, a saber:

Teste do razão 6.1: Considere uma direção $h \in \mathbb{R}^n$, a partir do ponto e . Se h tem alguma componente negativa, então o teste de razão

$$\bar{\lambda} = \min \left\{ \frac{-1}{h_i} \mid h_i < 0, \quad i = 1, \dots, n \right\} \quad (6.3)$$

fornece o maior passo possível ao longo de h , a partir do ponto e . \square

Critério de Parada: Se o mínimo de (6.3) é atingido somente em $i = n$, e se a função potencial $f_2(\dots)$ para a fase 2 (exclui-se a n -ésima coordenada $i + \bar{\lambda} h_n = 0$ de novo ponto obtido pelo algoritmo fase 1-fase2) não é maior do que $f_2(e, z)$ da iteração anterior, isto é,

$$q \ln c'(e + \bar{\lambda} h) - \sum_{i=1}^{n-1} \ln (e + \bar{\lambda} h)_i \leq f_2(e, z),$$

então deve-se parar o algoritmo fase1-fase2 concluindo a fase 1 (Note que este novo ponto para fase 2 não estará "próximo" da fronteira em (SA), a menos do ponto ótimo.). \square

Algoritmo 6.2: Busca unidirecional para (BA).

Dados: um intervalo $[u, v]$ e uma precisão $\epsilon > 0$.

Repita:

$$\alpha_1 := u + 0,38 (v - u).$$

$$\alpha_2 := u + 0,62 (v - u).$$

Se $\bar{f}(\alpha_1) < \bar{f}(\alpha_2)$ então $v := \alpha_2$.

Senão $u := \alpha_1$.

Até que $v - u \leq \epsilon$.

$\alpha := u$. \square

Observações 6.3:

(i) Tome $u = 0,2$ inicialmente.

(ii) A função $\bar{f}(\cdot) : \mathbb{R}_+ \rightarrow \mathbb{R}$ é definida por

$$\alpha \geq 0 \mapsto \bar{f}(\alpha) = \max \left\{ f_1(e + \alpha h) - \gamma, f_2(e + \alpha h, z) \right\},$$

onde

$$\gamma = \begin{cases} f_1(e) - f_2(e, z), & \text{se } \Delta_2 > 0, \\ 0, & \text{se } \Delta_2 \leq 0. \end{cases}$$

(iii) Tome v inicial da seguinte maneira:

Se $h > 0$, então

$$u := 0,2; \quad \lambda := 0,4; \quad v := 0,8.$$

Enquanto $\bar{f}(v) < \bar{f}(\lambda)$ faça

$$u := \lambda; \quad \lambda := v \quad v := 2v.$$

Senão use o Teste de Razão 6.1.

(iv) Por (i), e pelo fato de $\alpha = u$ na final do Algoritmo 6.2, garantimos então $\bar{f}(\alpha) \leq \bar{f}(0,2)$. \square

Segue-se um procedimento acerca da busca unidirecional para uma abordagem afim.

Procedimento 6.4: Busca unidirecional.

Dados: $x := e$, h , z e Δ_2

Use o Algoritmo 6.2. \square

6.2 - Uma Abordagem Projetiva

Consideremos as funções potenciais de Karmarkar

$$F_1(x) = q \ln \xi'x - \sum_{i=1}^n \ln x_i,$$

e

$$F_2(x, z) = \begin{cases} q \ln(c - zw)'x - \sum_{i=1}^n \ln x_i, & \text{se } c'x - z > 0 \\ -\infty, & \text{se } c'x - z \leq 0, \end{cases}$$

onde $q = n$ e z é um limitante inferior.

Aqui a estratégia é a mesma da seção anterior. Logo, ao considerarmos os sinais do parâmetro $\delta_2 = c'e - z$, incorremos no mesmo problema e consequentes algoritmo e procedimento como se seguem. Naturalmente, o critério de parada desenvolvido para a abordagem afim é análogo ao caso projetivo.

Problema 6.5:

$$\text{minimizar } \max \left\{ F_1(e + \alpha H) - \gamma, F_2(e + \alpha H, z) \right\}$$

$$\text{sujeito a: } e + \alpha H > 0$$

$$\alpha \geq 0,$$

onde

$$\gamma = \begin{cases} F_1(e) - F_2(e, z), & \text{se } \delta_2 > 0, \\ 0, & \text{se } \delta_2 \leq 0. \end{cases} \quad \square$$

Algoritmo 6.6: Busca unidirecional: Faça

$$\bar{f}(\alpha) = \max \left\{ F_1(e + \alpha H) - \gamma, F_2(e + \alpha H, z) \right\}$$

no Algoritmo 6.2. H é uma direção obtida resolvendo o problema (EDP)' se $\delta_2 > 0$, ou o problema (EDP) $_{\delta_2}$ se $\delta_2 \leq 0$.

Naturalmente este algoritmo resolve aproximadamente o Problema 6.5. \square

Segue-se um procedimento acerca da busca unidirecional para uma abordagem projetiva.

Procedimento 6.7: Busca unidirecional.

Dados: $x := e, H, z$ e δ_2

Use o Algoritmo 6.6. \square

O teorema a seguir nos diz que o Algoritmo 3.3 (ou Algoritmo 3.4) através da nossa busca unidirecional fornece um decréscimo da função máximo $f(.,.)$ (ou $F(.,.)$) de pelo menos 0,075 por iteração. Mostraremos para o caso afim. O caso projetivo é análogo.

Teorema 6.8: Considere o Algoritmo 3.3 aplicado sobre o problema (PA). Seja

$$f(e, z) = \max \{f_1(e), f_2(e, z)\},$$

para $e \in (5A)'$ e z obtido pelo Procedimento 4.2. Então, o resultado da busca unidirecional, α , através do Procedimento 6.4, fornece

$$f(e + \alpha h, z) \leq f(e, z) - 0,075,$$

onde h é obtido pelo Procedimento 5.13.

Demonstração: Considere o passo α obtido pelo Procedimento 6.4, a função máxima $\bar{f}(\cdot)$ definida conforme Observações 6.3(ii), e a função

$$f(e + \alpha h, z) = \max \{f_1(e + \alpha h), f_2(e + \alpha h, z)\},$$

definida para todo α , tal que $e + \alpha h > 0$ e $\alpha \geq 0$. Pela Observações 6.3(iv), obtemos

$$\bar{f}(\alpha) = \max\{f_1(e + \alpha h) - \gamma, f_2(e + \alpha h, z)\} \leq \bar{f}(0, 2). \quad (6.4)$$

Se $\Delta_2 \leq 0$, então $\gamma = 0$ e

$$f(e + \alpha h, z) = \bar{f}(\alpha) \leq \bar{f}(0, 2) = f(e + 0, 2h, z).$$

Pelo Teorema 5.12

$$f_1(e + 0, 2h) \leq f_1(e) - 0,075$$

$$f_2(e + 0, 2h, z) = -\infty.$$

Logo,

$$f(e + ah, z) \leq f(e + 0,2h, z) \leq f(e, z) - 0,075.$$

Se $\Delta_2 > 0$, então $\gamma = f_1(e) - f_2(e, z)$. Por (6.4) e pelo Teorema 5.7 segue-se que

$$\bar{f}(\alpha) \leq \bar{f}(0, z) \leq \bar{f}(0) - 0,075.$$

Por construção,

$$\bar{f}(0) = f_1(e) - \gamma = f_2(e, z).$$

Desenvolvendo $\bar{f}(\alpha)$,

$$f_1(e + ah) - \gamma \leq \bar{f}(0) - 0,075 = f_1(e) - \gamma - 0,075.$$

Isto é equivalente a

$$f_1(e + ah) \leq f_1(e) - 0,075. \quad (6.5)$$

Ainda,

$$f_2(e + ah, z) \leq \bar{f}(0) - 0,075 = f_2(e, z) - 0,075. \quad (6.6)$$

Portanto, tomando o máximo em (6.5) e (6.6),

$$f(e + \alpha h, z) \leq f(e, z) - 0,075,$$

finalizando a demonstração. ■

Agora estamos prontos para encerrar nosso trabalho no próximo capítulo, onde basta-nos demonstrar a convergência dos nossos dois algoritmos.

CAPITULO 7

CONVERGÊNCIA

Neste capítulo mostramos que os algoritmos de pontos interiores afim e projetivo para programação linear combinando fases 1 e 2 desenvolvidos nos capítulos anteriores, convergem com complexidade de $O(nL)$ iterações. Aqui, n é o número de variáveis do problema original, enquanto L é o tamanho desse problema, avaliado em número de bits relativo aos dados de entrada.

Assuma que z^k , $k = 0, 1, \dots$, é o limitante inferior obtido pelo Procedimento 4.2 e \hat{z} é o valor ótimo do problema original.

Vamos considerar o problema original resolvido quando se encontrar um ponto x^k tal que $\zeta'x^k \leq 2^{-L}$ e $c'x^k - z^k \leq 2^{-L}$. A partir deste ponto uma "purificação" detecta uma solução ótima.

Segue-se um lema fundamental para a demonstração de convergência de ambos os algoritmos (afim e projetivo).

Lema 7.1: Considere as seqüências x^k , z^k , $k = 0, 1, \dots$, geradas pelo Algoritmo 3.3 (ou Algoritmo 3.4) aplicado sobre o problema (PA) (ou (PP)). Seja

$$f(x^k, z^k) = \max \left\{ f_1(x^k), f_2(x^k, z^k) \right\},$$

para $x^k \in (SA)'$ (ou $(SP)'$) e z^k obtido pelo Procedimento 4.2. Então,

$$f(x^{k+1}, z^{k+1}) \leq f(x^k, z^k) - 0,075,$$

para todo $k \geq 0$.

Demonstração: Considere uma iteração $k \geq 0$. Por definição,

$$f(x^{k+1}, z^{k+1}) = \max \left\{ f_1(x^{k+1}), f_2(x^{k+1}, z^{k+1}) \right\},$$

para $x^{k+1} \in (SA)'$ (ou $(SP)'$) e z^{k+1} obtido pelo Procedimento 4.2. Pelo Teorema 6.8 segue-se que

$$f(x^{k+1}, z^{k+1}) \leq f(x^k, z^{k+1}) - 0,075.$$

Como $c'x^k - z^{k+1} \leq c'x^k - z^k$, uma vez que o Procedimento 4.2 gera um limitante inferior $z^{k+1} \geq z^k$, tem-se

$$f(x^{k+1}, z^{k+1}) \leq f(x^k, z^k) - 0,075,$$

completando a demonstração. ■

Teorema 7.2: Suponha que $x^0 := e$, $z^0 \leq \hat{z}$ e o Algoritmo 3.3 (ou Algoritmo 3.4) é aplicado para (PA) (ou (PP)). Então, o problema é resolvido em não mais que $O(qL)$ iterações.

Demonstração: Considere as seqüências x^k , z^k , $k = 0, 1, \dots$, geradas pelo Algoritmo 3.3 (ou 3.4) aplicado sobre o problema (PA) (ou (PP)). Seja

$$f(x^k, z^k) = \max \left\{ f_1(x^k), f_2(x^k, z^k) \right\},$$

para $x^k \in (SA)'$ (ou $(SP)'$) e z^k obtido pelo Procedimento 4.2. Então pelo Lema 7.1,

$$f(x^{k+1}, z^{k+1}) \leq f(x^k, z^k) - 0,075, \quad \forall k \geq 0.$$

Segue-se que

$$f(x^k, z^k) \leq f(x^0, z^0) - 0,075 k.$$

Mas, por definição de máximo e pela desigualdade acima

$$f_1(x^k) \leq f(x^k, z^k) \leq f(x^0, z^0) - 0,075 k,$$

$$f_2(x^k, z^k) \leq f(x^k, z^k) \leq f(x^0, z^0) - 0,075 k.$$

Temos que

$$f(x^0, z^0) = \max \left\{ f_1(x^0), f_2(x^0, z^0) \right\} = O(qL),$$

pois $f_1(x^0) = 0$ e $f_2(x^0, z^0) = q \ln(c'e - z^0) = O(qL)$

(observe que podemos ter $f_2(x^0, z^0) = -\infty$ ($f_1(x^0) = 0$)).

Assim, em particular,

$$f_1(x^k) = q \ln \xi'x^k + p(x^k) \leq O(qL) - 0,075k,$$

e portanto

$$q \ln \xi'x^k \leq O(qL) - 0,075k - p(x^k).$$

Pelo Lema 3.8,

$$-p(x^k) = \sum_{i=1}^n \ln x_i^k \leq n \ln \frac{e'x^k}{n}.$$

Kurt Anstreicher prova (comunicação particular - veja Anstreicher [2,3] e Todd [12, 14]), que

$$\frac{e'x^k}{n} = 2^{O(L)}.$$

Logo,

$$-p(x^k) \leq n \ln 2^{O(L)} = O(nL).$$

Então,

$$q \ln \xi'x^k \leq q O(L) - 0,075k + n O(L).$$

Para obter $\|x^k\| \geq 2^{-L}$ (uma vez que esta desigualdade vale em qualquer iteração do algoritmo) deve-se satisfazer

$$0,075 k \leq q_0(L) + n_0(L) - q \ln(2^{-L}),$$

o que se pode fazer com $k = O(nL)$, uma vez que $q = O(n)$. De modo análogo podemos desenvolver $f_2(x^k, z^k)$ tal que

$$\|c'x^k - z^k\| \leq 2^{-L}$$

finalizando a demonstração. ■

Neste teorema demonstramos convergência dos algoritmos afim e projetivo para programação linear combinando fases 1 e 2, cuja complexidade é de $O(nL)$ iterações.

As operações aritméticas dos dois algoritmos incluem uma inversão de matrizes no procedimento do cálculo da matriz de projeção. Isto significa que sua complexidade é de $O(n^3)$ operações aritméticas, totalizando $O(n^4L)$.

Atualmente, em algoritmos primais de pontos interiores para programação linear combinando fase 1 e fase 2 é interessante pesquisar algoritmos $O(\sqrt{n} L)$ iterações (para algoritmo primal-dual já foi feito por Ye, cuja referência não dispomos). Computacionalmente, foram desenvolvidos por enquanto apenas testes um tanto limitados sobre algoritmos projetivos (veja Todd e Wang [13]).

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] I. Adler, M.G.C. Resende, G.Veiga and N. Karmarkar, "An Implementation of Karmarkar Algorithm for Linear Programming", *Mathematical programming* 44 (1989) 297-335.
- [2] K.M. Anstreicher, "A Combined Phase I - Phase II Projective Algorithm for Linear Programming", *Mathematical Programming*, 43 (1989) 209-223.
- [3] K.M. Anstreicher, "A Combined Phase I - Phase II Scaled Potential Algorithm for Linear Programming", CORE Discussion Paper 8939, Catholic University of Louvian, Belgium, 1990.
- [4] M.S. Bazaraa and C.M.Shetty, "Nonlinear Programming - Theory and Algorithms", by John Wiley & Sons, 1979.
- [5] G.de Ghellinck and J.P.Vial, "A Polynomial Newton Method for Linear Programming", *Algorithmica* 1 (1986) 425-453.
- [6] C.C. Gonzaga. "Algoritmos de Pontos Interiores para Programação Linear", 17.^o Colóquio Brasileiro de Matemática, IMPA, 1989.
- [7] C.C. Gonzaga, "Polynomial Affine Algorithms for Linear Programming", *Mathematical Programming* 49 (1990) 7-21.

- [83] C.C. Gonzaga, "A Simple Presentation of Karmarkar's Algorithm", COPPE, Universidade Federal do Rio de Janeiro, Brasil, 1988.
- [91] C.C. Gonzaga, "On Lower Bound Updates in Primal Potential Reduction Methods for Linear Programming", Relatório Técnico, COPPE, Universidade Federal do Rio de Janeiro, maio de 1990. A ser publicado em mathematical Programming.
- [10] Humes, C. Jr. e A.F.P.C. Humes: Programação Linear - Um primeiro curso, Livro, Sociedade Brasileira de Matemática Aplicada e Computacional, 1986.
- [11] N. Karmarkar, "A New Polynomial Time Algorithm for Linear Programming", Combinatória 4 (1984) 373-395.
- [12] M.J. Todd, "On Anstreicher's Combined Phase I - Phase II Projective Algorithm for Linear Programming", to Appear in Mathematical Programming.
- [13] M.J. Todd and Y. Wang, "A Combined Phase I - Phase II Projective Methods for Linear Programming", Technical Report n.º 877 Scholl of OR & IE, Cornell University, Ithaca, NY, 1989.

[14] M.J. Todd, "A Combined Phase E - Phase II in a Potential Reduction Algorithm for Linear Programming", Technical Report n.º 907, School of OR & IE, Cornell University, Ithaca, NY, 1990.