

UM ALGORITMO DE PONTO-INTERIOR-INVIÁVEL COM
COMPLEXIDADE $O(\sqrt{n}L)$ ITERAÇÕES PARA PROGRAMAÇÃO LINEAR

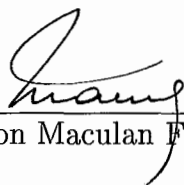
Marco Antonio Figueiredo Menezes

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 DOUTOR EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Aprovada por:



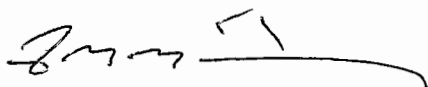
Prof. Clóvis Caesar Gonzaga, D. Sc.



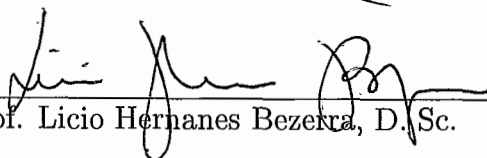
Prof. Nelson Maculan Filho, D. Sc.



Prof. Adilson Elias Xavier, D. Sc.



Prof. José Mario Martínez, D. Sc.



Prof. Licio Hernanes Bezerra, D. Sc.

RIO DE JANEIRO, RJ – BRASIL

AGOSTO DE 1998

MENEZES, MARCO ANTONIO FIGUEIREDO

Um Algoritmo de Ponto-Interior-Inviável
com Complexidade $O(\sqrt{n}L)$ Iterações para
Programação Linear [Rio de Janeiro] 1998

VIII, 88p. 29.7 cm (COPPE/UFRJ, D. Sc.,
Engenharia de Sistemas e Computação, 1998)

Tese - Universidade Federal do Rio de Ja-
neiro, COPPE

1. Programação Linear
2. Pontos Interiores
3. Ponto-Interior-Inviável
4. Complexidade

I. COPPE/UFRJ II. Título (Série).

*Para Hélia, Cláudia e Elis Regina.
Em memória, para Manoel e Guiomar
sobre uma música....*

Agradecimentos

Para todos aqueles que disseram algo ou não souberam o que dizer.

Para a CAPES pelo apoio financeiro, para o Departamento de Matemática e Computação Científica da Universidade Federal de Santa Catarina pelos três anos e pouco de acolhimento e para o Programa de Engenharia de Sistemas e Computação da COPPE/UFRJ pela sua existência.

Para o meu orientador e professor Clóvis Caesar Gonzaga, pela burocracia inicial para o meu ingresso neste Programa quando do mestrado, pela sua paciência para comigo, pelo aprendizado e por termos juntos, separados, “estrepados”, conseguido o resultado que jamais abandonamos. E, a quem lamento profundamente o que talvez seja irreparável pelo resto da minha vida: minha ignorância e minha incompetência.

Para o professor Nelson Maculan Filho, cuja palavra que me vem na cabeça é mãe. Também, desde a burocracia inicial para o meu ingresso neste Programa quando do mestrado até a burocracia final para a conclusão deste trabalho, passando pelas notas de aulas, as aulas propriamente ditas, um dos meus exames de qualificação. E, principalmente, confesso, minha admiração profissional, porque é um exemplo gratificante saber que é possível uma pessoa conduzir bem ensino, pesquisa e extensão.

Finalmente: Obrigado Hélia! Obrigado Adilson!

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

Um Algoritmo de Ponto-Interior-Inviável com Complexidade $O(\sqrt{n}L)$ Iterações
para Programação Linear

Marco Antonio Figueiredo Menezes

Agosto de 1998

Orientador: Clóvis Caesar Gonzaga

Programa: Engenharia de Sistemas e Computação

Neste trabalho descrevemos a teoria de ponto-interior-inviável existente na literatura e construímos um algoritmo de ponto-interior-inviável para programação linear. Iniciamos o estudo deste algoritmo com um algoritmo conceitual primal que usa um oráculo e nos fornece uma estratégia para manipularmos os parâmetros associados à viabilidade e otimalidade. Este algoritmo indica um caminho para o estudo da complexidade em iterações. Então, especificamos o oráculo e desenvolvemos um algoritmo completo que atinge a complexidade $O(\sqrt{n}L)$ iterações para passos curtos, com uma abordagem primal-dual tal que permitimos inviabilidade primal, mas mantemos viabilidade dual.

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

An $O(\sqrt{nL})$ -Iteration Infeasible-Interior-Point Linear Programming Algorithm

Marco Antonio Figueiredo Menezes

August 1998

Advisor: Clóvis Caesar Gonzaga

Department: Systems Engineering and Computer Science

In this work we describe the theory of infeasible-interior-point existing in the literature and we propose an infeasible-interior-point algorithm for linear programming. We start the study of this algorithm with a primal conceptual algorithm that uses an oracle, and provides a strategy to manipulate the parameters associated with feasibility and optimality. This indicates a way toward the study of complexity in iterations. Then, we specify the oracle and develop a complete algorithm which achieves the complexity $O(\sqrt{nL})$ -iteration for short steps with a primal-dual approach such that we permit primal infeasibility, but we maintain dual feasibility.

Índice

1	Introdução	1
2	Métodos de pontos interiores	6
2.1	O problema primal-dual	6
2.2	A trajetória central	12
2.3	A medida de proximidade e o passo de Newton	15
2.4	Complexidade	19
2.5	Alguns algoritmos	20
2.5.1	Passos curtos	21
2.5.2	Redução potencial	22
3	Métodos de ponto-interior-inviável	25
3.1	O problema primal-dual e a superfície de centros	26
3.2	A medida de proximidade e o passo de Newton	29
3.3	Alguns algoritmos	31
3.3.1	Preditor-corretor	32
3.3.2	Homogêneo e auto-dual	38
4	Um algoritmo conceitual primal	43
4.1	O problema	43
4.2	O algoritmo conceitual	52
5	Um algoritmo de ponto-interior-inviável	58
5.1	Sobre a notação	59

5.2	Os problemas primais-duais	62
5.3	As medidas de proximidade	65
5.4	O passo de Newton	70
5.5	O algoritmo de passos curtos	74
Referências Bibliográficas		81

Capítulo 1

Introdução

Existe um problema em aberto em programação linear (PL) que consiste na obtenção de um algoritmo de ponto-interior-inviável com complexidade polinomial em $O(\sqrt{n}L)$ iterações para a formulação não artificial do problema de PL, onde n é o número de variáveis do problema e L é o tamanho do problema para dados inteiros. Com a exceção do método homogêneo e auto-dual para PL, Ye, Todd e Mizuno [71] (e para problemas de complementaridade linear com a hipótese de monotonicidade, PCL, Ye [68]) que possui $O(\sqrt{n}L)$ iterações para uma formulação artificial do problema de PL, a menor complexidade em iterações para a formulação não artificial é $O(nL)$. Neste trabalho estudamos e resolvemos este problema permitindo inviabilidade primal, mas mantendo viabilidade dual. Na inicialização do nosso algoritmo é necessário um procedimento de centralização, que é feito em geral nos algoritmos de pontos interiores (viáveis). Este procedimento de centralização tem complexidade $O(\sqrt{n}L)$ iterações (veja Gonzaga [17]).

Estudamos a resolução de um problema de PL

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

onde $c \in R^n$, $b \in R^m$ e A é uma matriz de *rank* máximo, $0 < m \leq n$.

Algoritmos de pontos interiores para PL partem normalmente de um ponto

viável inicial que deve ser fornecido por algum procedimento de inicialização. Um procedimento bem conhecido parte de um ponto inicial $x^0 > 0$ não viável, introduz uma nova variável γ e modifica a restrição para

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

O ponto

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

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

Pode-se reduzir γ associando a esta variável um custo muito alto M (método *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 a mais. Todavia, veja um trabalho relevante devido a Kojima, Mizuno e Yoshise [24].

Reduzir simultaneamente a variável γ e o critério original $c^T x$, como em um problema de critérios múltiplos, sem associar custo à variável γ , é conhecido como algoritmos combinando fases 1 e 2 e já foi desenvolvido por Anstreicher [2] e [3], Todd [52] e [53], Todd e Wang [54], Freund [10] e [11] e Menezes [32]. De Ghellinck e Vial [13] também propuseram um algoritmo primal do tipo projetivo que não requer viabilidade.

Algoritmo de ponto-interior-inviável é um algoritmo de pontos interiores (pontos no primeiro octante com coordenadas estritamente positivas) tal que o ponto inicial não é restrito a ser um ponto viável. Recentemente, os algoritmos de ponto-interior-inviável primais-duais para PL e PCL são conhecidos como os algoritmos de pontos interiores mais eficientes na prática (veja Lustig, Marsten e Shanno [26] e, veja também, Marsten, Subramanian, Saltzman, Lustig e Shanno [28]).

Historicamente, porém não cronologicamente, Kojima, Megiddo e Mizuno [25] provaram a convergência de um algoritmo primal-dual de ponto-interior-inviável

para PL. Zhang [72] demonstrou convergência para PCL horizontal com uma certa condição sobre o ponto inicial, a qual foi retirada mais tarde por Billups e Ferris [6]. O resultado principal de Zhang foi a convergência polinomial em $O(n^2L)$ iterações detectando inviabilidade para uma certa região. Zhang e Zhang [73] estudaram o algoritmo de Zhang para PL, enquanto Wright [62] para PCL e, ambos, demonstraram convergência superlinear. Wright [63] construiu uma variante para o algoritmo de Zhang com convergência superlinear. Para esta variante Wright [64] abandonou a função de mérito usada por Zhang, que é a soma do *gap* de complementaridade e a norma do resíduo de inviabilidade, permitindo maior flexibilidade na escolha de parâmetros e do ponto inicial, mantendo convergência superlinear. Wright e Zhang [65] mostraram convergência superquadrática. Monteiro e Wright [41] desenvolveram um algoritmo com convergência superlinear onde as direções são afim-escala em $O(n^2L^2)$ iterações. Mizuno [35] demonstrou $O(n^2L)$ para o algoritmo de Kojima, Megiddo e Mizuno e desenvolveu uma variante deste obtendo $O(nL)$ iterações. Ambos algoritmos detectam inviabilidade para uma certa região. A variante de Mizuno foi um algoritmo preditor-corretor. Quase simultaneamente, Potra [43] obteve a mesma complexidade em iterações. Ainda, Potra [44] demonstrou convergência quadrática e $O(\sqrt{n}L)$ iterações sob certas condições para o ponto inicial. Além disso, estendeu e melhorou seus resultados para PCL, Potra [45] e Sheng e Potra [47].

Vários outros trabalhos foram desenvolvidos. Mizuno, Todd e Ye [37] definiram uma superfície de centros analíticos determinada por um par primal-dual de problemas de PL e um ponto interior inviável. Além disso, estudaram essa superfície e relacionaram alguns métodos. Algoritmos $O(nL)$ iterações, além dos de Mizuno e Potra, podem ser encontrados em Mizuno, Kojima e Todd [38], Stoer [48], Bonnans e Potra [7], Miao [33] e Tseng [57]. Por outro lado, Drummond [8] estudou métodos primais-duais com inviabilidade primal e viabilidade dual, mas não tratou sobre complexidade.

Acreditamos que a dificuldade em obter $O(\sqrt{n}L)$ iterações para esses métodos

decorre da atualização dos parâmetros associados à viabilidade e à otimalidade. Isto é, todos os métodos citados decrescem esses parâmetros na mesma proporção. Todavia, observamos que às vezes devemos aumentar o parâmetro relativo à otimalidade, conforme nossas implementações em pequenos problemas testes. E, implementações estas, a propósito, que não apresentaremos aqui por serem simples mesmo. Na verdade, o nosso problema teste principal foi aquele sugerido por Gonzaga, a saber: a partir do ponto inicial inviável $e \in R^2$, resolver

$$\begin{aligned} & \textit{minimizar} && x_1 \\ & \textit{sujeito a:} && x_1 - 1000x_2 = 1000 \\ & && x_1, x_2 \geq 0. \end{aligned}$$

Nossa estratégia para este trabalho é, inicialmente, aprender como controlar os parâmetros associados à viabilidade e à otimalidade observando atentamente o problema relativo à viabilidade. Tratamos disto no capítulo 4. Em seguida, construímos um algoritmo de ponto-interior-inviável com passos curtos com complexidade $O(\sqrt{n}L)$ iterações para PL. Este é o nosso resultado que será desenvolvido no último capítulo. Os capítulos 2 e 3 são básicos e consistem em estudar métodos de pontos interiores e métodos de ponto-interior-inviável, respectivamente.

É importante comunicar que todos os resultados dos capítulos 4 e 5 têm a participação direta ou indireta do meu orientador.

Notação: Denotamos R^n o espaço vetorial Euclidiano n -dimensional. A matriz I é a matriz identidade e o vetor e é o vetor de uns. As dimensões desta matriz e deste vetor serão apropriadas no contexto. A função norma denotada por $\|\cdot\|$ é a norma Euclidiana, enquanto $\|\cdot\|_1$ é a norma da soma e $\|\cdot\|_\infty$ é a norma do máximo. Dados vetores x, s, d em R^n , as letras maiúsculas X, S, D são matrizes diagonais definidas pelos respectivos vetores. Quando isto não ocorrer notificamos imediatamente. Denotamos operações componente por componente pela notação usual de números reais. Assim, dados vetores u, v com a mesma dimensão, $uv, u/v = uv^{-1}$, etc. denotam os vetores com componentes $u_j v_j, u_j / v_j$, etc.. As funções exponencial, logaritmo natural e logaritmo na base 2 são denotadas, respectivamente, por $exp(\cdot)$,

$\ln(\cdot)$ e $\log(\cdot)$. Dada uma matriz A , um vetor u com o sub-índice p , u_p , denota o vetor projetado no espaço nulo de A . Números com vírgula, por exemplo 0,25, será denotado por 0.25, isto é, substituímos a vírgula pelo ponto.

Capítulo 2

Métodos de pontos interiores

O objetivo deste capítulo é o estudo dos métodos de pontos interiores, o que será a nossa ferramenta básica para os capítulos posteriores. O que faremos aqui, então, será reescrever o que já existe na literatura de pontos interiores, de uma maneira concisa e organizada. As demonstrações serão, em princípio, todas elas referenciadas. Na introdução do artigo de Terlaky [51], encontramos uma bibliografia comentada sobre os livros existentes até o momento de pontos interiores. Um livro básico para o estudo da programação não linear necessária para o estudo de pontos interiores, citamos o livro de Martínez e Santos [29].

2.1 O problema primal-dual

Otimização é uma área da Matemática Aplicada. O problema de otimização consiste em encontrar, se possível, os minimizadores (ou os maximizadores) de uma função definida em uma determinada região.

Consideremos os números inteiros m e n tais que $n \geq m > 0$. Dados uma matriz $A \in R^{m \times n}$ e vetores $b \in R^m$ e $c \in R^n$, o problema de programação linear é o seguinte problema de otimização, usualmente denominado problema primal:

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

Seguem-se algumas definições associadas ao problema (P). A função linear $x \mapsto c^T x$ é chamada função objetivo. O conjunto $\mathcal{X} = \{x \in R^n; Ax = b, x \geq 0\}$ é chamado conjunto viável e um ponto $x \in \mathcal{X}$ é denominado ponto viável. O conjunto $\{x \in R^n; x > 0\}$ é chamado conjunto de pontos interiores e um ponto deste conjunto é denominado ponto interior. O conjunto $\mathcal{X}^0 = \{x \in \mathcal{X}; x > 0\}$ é chamado conjunto de pontos interiores viáveis e um ponto $x \in \mathcal{X}^0$ é denominado ponto interior viável. O número $v(P) = \inf\{c^T x; x \in \mathcal{X}\}$ é denominado o valor ótimo ou custo ótimo, o conjunto $\mathcal{X}(P) = \{x \in \mathcal{X}; c^T x = v(P)\}$ é chamado conjunto de soluções ótimas e um ponto $x \in \mathcal{X}(P)$ é denominado solução ótima. Se existir uma seqüência (x^k) tal que $x^k \in \mathcal{X}$ e $c^T x^k \rightarrow -\infty$, então (P) é dito ilimitado. Se \mathcal{X} é vazio, então (P) é dito inviável.

O problema de programação linear denominado o dual do problema (P), é o seguinte problema de otimização:

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

onde o vetor $y \in R^m$ é denominado variável dual e incluímos explicitamente um vetor com componentes não negativas $s \in R^n$ denominado folga dual. Definimos os conjuntos viável, de pontos interiores, de pontos interiores viáveis e de soluções ótimas para o problema (D), respectivamente, por $\mathcal{S} = \{(y, s) \in R^m \times R^n; A^T y + s = c, s \geq 0\}$, $\{(y, s) \in R^m \times R^n; s > 0\}$, $\mathcal{S}^0 = \{(y, s) \in \mathcal{S}; s > 0\}$ e $\mathcal{S}(D) = \{(y, s) \in \mathcal{S}; b^T y = v(D)\}$, onde $v(D) = \sup\{b^T y; (y, s) \in \mathcal{S}\}$ é denominado o valor ótimo ou custo ótimo do problema (D). Definimos os pontos viável, interior, interior viável e solução ótima para o problema (D), respectivamente, pelos pontos em \mathcal{S} , $\{(y, s) \in R^m \times R^n; s > 0\}$, \mathcal{S}^0 e $\mathcal{S}(D)$. A função linear $y \mapsto b^T y$ é chamada função

objetivo.

Sem perda de generalidade, assumimos $\text{rank}A = m$. Com efeito, vamos considerar inicialmente o problema primal (P). Se as linhas da matriz A são linearmente dependentes, então (P) é inviável ou as equações lineares redundantes podem ser removidas sequencialmente até a matriz resultante ter *rank-linha* completo. Por outro lado, considerando o problema dual (D), se as colunas da matriz A^T são linearmente dependentes, isto é, $A^T z = 0$ e $b^T z \neq 0$, então se (D) é viável, ele é ilimitado. Se $b^T z = 0$ para todo z , então podemos eliminar uma das colunas correspondentes a uma componente não nula de z sem afetar (D). Continuando com este processo obtemos um problema equivalente onde A tem *rank-linha* completo.

Com esta suposição sobre o *rank* de A , o problema (D) pode ser reescrito como um problema em s apenas. Vamos fazer isto lembrando inicialmente alguns resultados de Álgebra Linear.

Considere a transformação linear definida pela matriz $A \in R^{m \times n}$. Dois subespaços importantes do espaço vetorial R^n estão associados com esta transformação: o espaço nulo $\mathcal{N}(A) = \{x \in R^n; Ax = 0\}$ e seu complemento ortogonal, o espaço imagem de A^T , definido por $\mathcal{R}(A^T) = \{x \in R^n; x = A^T z, z \in R^m\}$.

Proposição 2.1.1 *Um vetor $d \in R^n$ é ortogonal a $\mathcal{R}(A^T)$ se, e somente se, d pertence a $\mathcal{N}(A)$.*

Demonstração. Fato (página 8) em Gonzaga [16]. ■

Desta proposição, qualquer vetor $d \in R^n$ pode ser unicamente decomposto (soma direta) como $d = d_p + d_{\bar{p}}$, onde $d_p \in \mathcal{N}(A)$ e $d_{\bar{p}} \in \mathcal{R}(A^T)$. Os vetores d_p e $d_{\bar{p}}$ são, respectivamente, a projeção de d no nulo de A e no seu complemento ortogonal. A projeção de d no nulo de A é o ponto no $\mathcal{N}(A)$ com menor distância Euclidiana para d . Esta é a definição mais usual de projeção:

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

Uma vez que o operador projeção é linear, podemos representá-lo por uma

matriz P_A , tal que $d_p = P_A d$. O complemento ortogonal é $d_{\tilde{p}} = \tilde{P}_A d$, onde $\tilde{P}_A = I - P_A$.

Proposição 2.1.2 *Seja $A \in R^{m \times n}$ uma matriz de rank igual a m , $m \leq n$. Então a matriz AA^T é não singular.*

Demonstração. Fato (página 9) em Gonzaga [16]. ■

Proposição 2.1.3 *Sejam $A \in R^{m \times n}$ uma matriz de rank igual a m , $m \leq n$ e $d \in R^n$ um vetor arbitrário. Então $d_p = P_A d$ e $d_{\tilde{p}} = \tilde{P}_A d$, onde $P_A = I - A^T(AA^T)^{-1}A$ e $\tilde{P}_A = I - P_A$.*

Demonstração. Fato (página 9) em Gonzaga [16]. ■

Neste ponto, a observação interessante é que o cálculo de P_A inclui uma inversão de matriz e será sempre o procedimento mais trabalhoso em todos os algoritmos que estudaremos.

Retornando ao nosso propósito, isto é, reescrever (D) como um problema em s apenas, considere G uma matriz real $(n - m) \times n$, cujas linhas formam uma base para $\mathcal{N}(A)$. Então se $A^T y + s = c$, $Gs = GA^T y + Gs = Gc$ (Proposição 2.1.1). E, se $Gs = Gc$, então $(c - s) \in \mathcal{N}(G)$, isto é, $(c - s) \in \mathcal{R}(A^T)$. Assim, $A^T y + s = c$ para algum y . Uma vez que A tem *rank* completo, y é único (Proposição 2.1.2). Daí, existe uma correspondência um-a-um entre \mathcal{S} e $\tilde{\mathcal{S}}$ e entre \mathcal{S}^0 e $\tilde{\mathcal{S}}^0$, onde $\tilde{\mathcal{S}} = \{s; (y, s) \in \mathcal{S} \text{ para algum } y\}$ e $\tilde{\mathcal{S}}^0 = \{s; (y, s) \in \mathcal{S}^0 \text{ para algum } y\}$.

Agora vamos relembrar alguns resultados básicos acerca de programação linear que serão úteis neste trabalho.

Dadas uma solução viável primal x e uma solução viável dual (y, s) , então

$$c^T x - b^T y = (A^T y + s)^T x - (Ax)^T y = x^T s \geq 0,$$

que é a propriedade de dualidade fraca. Dizemos que $c^T x - b^T y$ é o *gap* de dualidade e $x^T s$ é denominado *gap* de complementaridade. A propriedade de dualidade forte afirma que se ambos os problemas (P) e (D) têm soluções viáveis, então ambos têm

soluções ótimas $x^* \in \mathcal{X}(P)$ e $(y^*, s^*) \in \mathcal{S}(D)$ tais que $c^T x^* - b^T y^* = (x^*)^T s^* = 0$. O teorema de dualidade é enunciado assim: se o problema primal tem uma solução ótima, então o problema dual também tem uma solução ótima e os valores das funções objetivos de ambos são iguais; se o problema primal (ou o dual) é ilimitado, então o problema dual (ou o primal) é inviável; e se o problema primal (ou o dual) é inviável, então o problema dual (ou o primal) é inviável ou ilimitado.

Considere $x^* \in \mathcal{X}$ e $s^* \in \tilde{\mathcal{S}}$ soluções ótimas, respectivamente, para os problemas primal e dual. Note que para $(x^*, s^*) \geq 0$, $(x^*)^T s^* = 0$ equivale a $x^* s^* = 0$. Esta igualdade tem uma relação combinatorial, dizendo que para cada $j = 1, \dots, n$, $x_j^* = 0$ ou $s_j^* = 0$. Esta é a condição de folga complementar. O teorema de folga complementar estrita afirma que se ambos os problemas primal e dual têm soluções viáveis, então ambos têm um par de solução complementar estrita x^* e s^* , isto é, $x^* + s^* > 0$.

Uma consequência importante de folgas complementares é o próximo resultado.

Proposição 2.1.4 *Se c não está no espaço linha de A , então $c^T x > v(P)$ para todo $x \in \mathcal{X}^0$.*

Demonstração. Proposição 2.1 em Todd [55]. ■

As condições de otimalidade para o par de problemas primal (P) e dual (D) coincidem com as condições de Karush-Kuhn-Tucker, a saber: x é uma solução ótima de (P) se, e somente se, existe um par de multiplicadores de Lagrange (y, s) tal que

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

O problema primal-dual é definido assim: encontrar, se existir, uma solução para o sistema de equações e inequações (PD). O sistema não linear (PD) acima pode ser visto como um sistema linear, bastando substituir $xs = 0$ por $c^T x - b^T y = 0$.

Definimos o conjunto viável primal-dual por

$$\mathcal{F} = \{(x, s) \in \mathbb{R}^n \times \mathbb{R}^m; x \in \mathcal{X}, s \in \tilde{\mathcal{S}}\}$$

e o conjunto viável de pontos interiores primal-dual por

$$\mathcal{F}^0 = \{(x, s) \in \mathcal{F}; (x, s) > 0\}.$$

Faremos algumas hipóteses para o desenvolvimento de métodos de pontos interiores neste capítulo. O próximo resultado nos auxiliará nesta tarefa.

Teorema 2.1.5 *Suponha que \mathcal{X} é não vazio. Então as seguintes afirmações são equivalentes:*

- (a) \mathcal{S}^0 é não vazio;
- (b) Para todo $\kappa \in \mathbb{R}$, $\{x \in \mathcal{X}; c^T x \leq \kappa\}$ é limitado;
- (c) $\mathcal{X}(P)$ é não vazio e limitado; e
- (d) Para algum $\kappa \in \mathbb{R}$, $\{x \in \mathcal{X}; c^T x \leq \kappa\}$ é não vazio e limitado.

Demonstração. Teorema 2.1 em Todd [55]. ■

As hipóteses para este capítulo são as seguintes:

(H_1) \mathcal{X}^0 é não vazio; e

(H_2) $\mathcal{X}(P)$ é não vazio e limitado.

Adiantamos que faremos uma “hipótese forte” quando tratarmos dos métodos de pontos interiores adiante para este capítulo.

Observando o teorema anterior, a hipótese (H_2) é o item (c) deste teorema. Logo, \mathcal{S}^0 é não vazio pelo item (a). Então, pela hipótese (H_1) e usando dualidade forte, estamos nas condições de otimalidade para o problema (PD). De fato, para esta implicação bastar-nos-ia considerar a hipótese (H_2) e o teorema de dualidade. Todavia, a relação \mathcal{X}^0 e \mathcal{S}^0 não vazios garantirá, na próxima seção, unicidade de pontos centrais primais-duais. Os itens (b) e (d) são importantes para se viabilizar

a construção de algoritmos de pontos interiores, mostrando a existência de um caminho no interior relativo do conjunto viável primal. A propósito, o nosso próximo intento será mostrar a existência de um caminho no interior relativo do conjunto viável primal-dual.

2.2 A trajetória central

Retomando o problema primal-dual (PD), então, e munidos das hipóteses (H_1) e (H_2) , a dificuldade básica para resolvê-lo consiste no problema combinatorial de determinar qual de cada par de variáveis (x_j, s_j) , $j = 1, \dots, n$, será nula em uma solução ótima. Para contornar esta dificuldade vamos iniciar com o resultado a seguir, o qual é útil na demonstração do próximo lema e garante existência e unicidade no passo de Newton adiante.

Proposição 2.2.1 *Considere $(x, s) > 0$ e o sistema de igualdades*

$$F(x, y, s) = \begin{pmatrix} Ax \\ A^T y + s \\ xs \end{pmatrix}.$$

A matriz Jacobiana

$$J(x, y, s) = \begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{pmatrix}.$$

é não singular.

Demonstração. Lema 2 em Mizuno, Todd e Ye [37]. ■

Definimos a função barreira logarítmica (logaritmo natural), a saber:

$$x \in R_{++}^n \mapsto p(x) = - \sum_{j=1}^n \ln(x_j).$$

Para $x \in R_{++}^n$, o gradiente e a matriz hessiana da função $p(\cdot)$ são respectivamente,

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

A função $p(\cdot)$ é estritamente convexa em \mathcal{X}^0 , porque X^{-2} é definida positiva e cresce indefinidamente para algum x_j , $j = 1, \dots, n$, próximo de zero.

A função barreira logarítmica foi utilizada pela primeira vez em otimização por Frisch [12]. Ela é usada em conjunção com o custo na função penalizada, a saber: para cada $\mu \in (0, \infty)$, definimos

$$x \in R_{++}^n \mapsto f_\mu(x) = \frac{c^T x}{\mu} + p(x).$$

O método de função penalizada barreira logarítmica usada em otimização contínua não linear restrita, estudado extensamente no livro de Fiacco e McCormick [9], pode ser aplicado para o problema de programação linear. Este método foi sugerido por Gill, Murray, Saunders, Tomlin e Wright [14], onde o algoritmo de Karmarkar [21] foi analisado do ponto de vista da função penalizada barreira logarítmica.

Em programação linear, o método de função penalizada barreira logarítmica consiste em examinar a família de problemas

$$\begin{aligned} (P_\mu) \quad & \text{minimizar} \quad f_\mu(x) \\ & \text{sujeito a:} \quad x \in \mathcal{X}^0, \end{aligned}$$

onde μ é um parâmetro de penalidade positivo no custo. Avaliar a função $f_\mu(\cdot)$ só tem sentido se \mathcal{X}^0 é não vazio. Daí a necessidade da hipótese (H_1) . A necessidade da hipótese (H_2) aparece agora.

Proposição 2.2.2 *Considere a hipótese (H_1) e $\mu > 0$ dado. Então o problema (P_μ) possui solução ótima se, e somente se, o conjunto $\mathcal{X}(P)$ é não vazio e limitado.*

Demonstração. Proposição 2.1 em Monteiro e Adler [40]. ■

O próximo resultado une as hipóteses (H_1) e (H_2) para garantir unicidade para as soluções do problema (P_μ) , para todo $\mu > 0$.

Proposição 2.2.3 *Considere as hipóteses (H_1) e (H_2) . Então, para cada $\mu > 0$, o problema (P_μ) possui uma única solução $x(\mu)$.*

Demonstração. Corolário 2.1 em Monteiro e Adler [40]. ■

As condições de Karush-Kuhn-Tucker para o problema (P_μ) , para cada $\mu > 0$, são as seguintes: o vetor $x > 0$ é a solução ótima para o problema (P_μ) se, e somente se, existirem vetores $y \in R^m$ e $s \in R^n$ tais que

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

O próximo resultado é uma consequência da última proposição.

Proposição 2.2.4 *Considere as hipóteses (H_1) e (H_2) . Então, para cada $\mu > 0$, o sistema $(PD)_\mu$ possui uma única solução $(x(\mu), y(\mu), s(\mu))$.*

Demonstração. Páginas 30 e 31 em Monteiro e Adler [40]. ■

A importância da unicidade nesta proposição será formalizada no próximo lema.

Lema 2.2.5 *A função $\mu > 0 \mapsto (x(\mu), y(\mu), s(\mu))$ é uma curva diferenciável, denominada trajetória central.*

Demonstração. Página 195 em Adler e Monteiro [1]. ■

O ponto $(x, s) \in R^n \times R^n$ é o ponto central primal-dual associado a $\mu > 0$ se, e somente se, $(x, s) \in \mathcal{F}$ e $xs = \mu e$. Desta forma, o vetor x é o ponto central primal associado a $\mu > 0$ se, e somente se, $s = \mu x^{-1}$ é o ponto central dual associado a $\mu > 0$. Neste caso, o *gap* de complementaridade entre as soluções primal e dual é dado por $x^T s = n\mu$.

A trajetória central é a curva formada pelos pontos centrais primais-duais associados a $\mu > 0$.

Neste ponto devemos verificar se, quando μ tende para zero (daí $x^T s \rightarrow 0$), a trajetória central converge para o centro analítico do conjunto de soluções ótimas de \mathcal{F} . Se isto acontecer, podemos pensar em algum método que segue de alguma maneira esta trajetória.

Teorema 2.2.6 *Considere a trajetória central $\mu > 0 \mapsto (x(\mu), y(\mu), s(\mu))$.*

- (a) *Para $\mu_0 > 0$ dado, o par central $(x(\mu), s(\mu))$ é limitado, para $\mu \in (0, \mu_0]$;*
- (b) *Para $0 < \mu' < \mu$, $c^T x(\mu') < c^T x(\mu)$ e $b^T y(\mu') > b^T y(\mu)$; e*
- (c) *$(x(\mu), s(\mu))$ converge para um par de solução ótima para (P) e (D) quando μ tende para zero. Além disso, o ponto limite $x(0)$ é o centro analítico de $\mathcal{X}(P)$, definido por*

$$x(0) = \operatorname{argmin}\left\{-\sum_{j \in P^*} \ln(x_j); x \in \mathcal{X}(P)\right\}$$

e o ponto limite $s(0)$ é o centro analítico definido por

$$s(0) = \operatorname{argmin}\left\{-\sum_{j \in S^*} \ln(s_j); (y, s) \in \mathcal{S}(D), \text{ para algum } y\right\},$$

onde (P^, S^*) é a partição de complementaridade estrita do conjunto de índices $\{1, \dots, n\}$.*

Demonstração. Teorema 2.14 em Ye [67]. ■

2.3 A medida de proximidade e o passo de Newton

Iniciamos esta seção enunciando um modelo de algoritmo de trajetória central.

Algoritmo 2.3.1 *Modelo.*

Dados: $\epsilon > 0$ e $\mu^0 > 0$.

$k := 0$.

REPITA

Encontrar aproximadamente $(x(\mu^k), s(\mu^k))$.

Escolher $\mu^{k+1} < \mu^k$.

$k := k + 1$.

ATÉ QUE $\mu^k < \epsilon$.

De fato, dado $\mu > 0$, queremos encontrar o par (x, s) tal que $xs/\mu = e$. A palavra ‘aproximadamente’ neste modelo de algoritmo se refere a alguma medida de proximidade que definimos agora:

$$(x, s, \mu) \in \mathcal{F}^0 \times R_{++} \mapsto \delta(x, s, \mu) = \left\| \frac{xs}{\mu} - e \right\|.$$

A proposição a seguir mostra que se a medida de proximidade é pequena, então o *gap* de complementaridade está bem relacionado com μ para tais pontos “quase centrais”.

Proposição 2.3.2 *Se $\delta(x, s, \mu) \in (0, 1)$, então*

$$x^T s \leq (n + \delta(x, s, \mu)\sqrt{n})\mu.$$

Demonstração. Página 202 em Gonzaga [18]. ■

Métodos de pontos interiores trabalham com pontos próximos da trajetória central. Dado o parâmetro $\alpha \in (0, 1)$ (usualmente $\alpha = 0.5$), a vizinhança da trajetória central é definida por

$$\mathcal{V}(\alpha) = \{(x, s, \mu) \in \mathcal{F}^0 \times R_{++}; \delta(x, s, \mu) \leq \alpha\}.$$

Desta forma, a Proposição 2.3.2 justifica algoritmos de trajetória central, os quais geram seqüências (x^k, s^k, μ^k) tais que μ^k tende para zero e $\delta(x^k, s^k, \mu^k) \leq \alpha$, $\alpha \in (0, 1)$.

Por outro lado, cada iteração do Algoritmo 2.3.1 inicia com um ponto interior viável (x, s) e um valor de μ , e tenta se aproximar de $(x(\mu), s(\mu))$. Idealmente, gostaríamos de encontrar

$$x^+ = x + u, \quad y^+ = y + w \quad \text{e} \quad s^+ = s + v,$$

tais que $(x^+, s^+) \in \mathcal{F}$ e $x^+ s^+ = \mu e$. A direção (u, w, v) deve ser viável, isto é, $A(x + u) = Ax$ e $A^T(y + w) + s + v = A^T y + s$. Isto é equivalente a dizer que $u \in \mathcal{N}(A)$ e $v \in \mathcal{R}(A^T)$.

O passo de Newton resolve isto aproximadamente linearizando o sistema perturbado $(PD)_\mu$, a saber:

$$\begin{aligned} Au &= 0 \\ (N) \quad A^T w + v &= 0 \\ su + xv &= -xs + \mu e. \end{aligned}$$

A linearização é básica e o sistema linearizado (N) possui uma única solução, conforme Proposição 2.2.1.

Essencialmente, dizemos que a importância da disciplina pontos interiores é a garantia de eficiência do passo de Newton para alguma medida de proximidade, no sentido do próximo teorema. Antes, porém, vamos nos munir de um resultado auxiliar relevante sobre normas de vetores devido a Mizuno.

Lema 2.3.3 *Se $u, v \in R^n$ são tais que $u^T v \geq 0$, então*

$$\|uv\| \leq \frac{\|u + v\|^2}{\sqrt{8}}.$$

Demonstração. Lema 1 em Mizuno [34]. ■

Agora o teorema fundamental de pontos interiores.

Teorema 2.3.4 *Dado $(x, s, \mu) \in \mathcal{V}(\delta)$ tal que $\delta(x, s, \mu) = \delta < 1$, seja (x^+, s^+) o resultado de um passo de Newton de (x, s) . Então,*

(a) *Para qualquer $\lambda \in [0, 1]$,*

$$\delta(x + \lambda u, s + \lambda v, \mu) \leq (1 - \lambda)\delta(x, s, \mu) + \lambda^2 \delta(x^+, s^+, \mu);$$

(b)

$$\delta(x^+, s^+, \mu) \leq \frac{\delta^2}{\sqrt{8(1-\delta)}}; \quad e$$

(c) Para $\delta \leq 0.7$, $(x^+, s^+) \in \mathcal{F}^0$.

Demonstração. Para os itens (a) e (b), Teorema 4.1 em Gonzaga [18]. Para o item (c) é o seguinte: Uma vez que (u, v) é uma direção viável por construção, basta mostrar que $(x^+, s^+) > 0$. A positividade de (x^+, s^+) é demonstrada por contradição. Suponhamos sem perda de generalidade que $x_j^+ \leq 0$ para algum j , $j \in \{1, \dots, n\}$. O item (b) com $\delta \leq 0.7$ implica que $x_j^+ s_j^+ > 0$. Logo devemos ter $x_j^+ < 0$ e $s_j^+ < 0$. Segue-se que $x_j^+ = x_j + u_j < 0$ implicando em $u_j < -x_j$, que multiplicando ambos os lados desta desigualdade estrita por s_j , obtemos $s_j u_j < -x_j s_j$. Analogamente, $s_j^+ = s_j + v_j < 0$ implicando em $v_j < -s_j$, que multiplicando ambos os lados desta desigualdade estrita por x_j , obtemos $x_j v_j < -x_j s_j$. Então, uma vez que $\mu > 0$ e considerando o passo de Newton,

$$-x_j s_j < -x_j s_j + \mu = s_j u_j + x_j v_j < -2x_j s_j,$$

que é uma contradição. Isto completa a demonstração. ■

Uma curiosidade, é que a demonstração do item (c), apesar de ter tido o mesmo raciocínio conforme Teorema 2.4 em Sheng e Potra [47], me valeu um chocolate e uma nota de agradecimentos em Gonzaga [18].

A interpretação deste teorema é a seguinte: o método de Newton é muito eficiente para resolver o problema $(PD)_\mu$, $\mu > 0$, no sentido que ele reduz a medida de proximidade quadraticamente em uma região grande. E, não apenas o passo de Newton puro é eficiente, mas a proximidade cai continuamente ao longo da direção de Newton para $\delta \leq 0.7$.

Resta-nos, ainda, estudar um pouco sobre complexidade.

2.4 Complexidade

Nesta seção estudamos um pouco de complexidade para o problema primal (P) com dados inteiros, cuja finalidade é a de avaliar os métodos de pontos interiores quanto ao número de iterações. Mostramos que ao invés de calcular uma solução ótima exata do problema, é suficiente encontrar uma solução viável com custo suficientemente próximo do valor ótimo e, então, executar um procedimento de arredondamento denominado ‘algoritmo de purificação’. Além disso, introduzimos a terceira hipótese para este capítulo.

O algoritmo de purificação é similar ao método simplex, reduzindo uma variável a zero em cada iteração por uma operação de pivoteamento, tal que no máximo $(n - 1)$ pivôs são necessários para atingir um vértice.

Definimos o número L_0 como o comprimento total dos dados de entrada em representação binária, isto é, a soma do número de *bits* usados por todas as entradas de A , b e c . O significado deste número é o seguinte: nenhuma operação contendo produtos e somas de dados sem repetições pode resultar em um número maior do que 2^{L_0} .

O tamanho do problema (P) com dados inteiros é definido pelo número $L = 2L_0 + n + 1$, onde L_0 é o número de *bits* utilizados pela entrada de dados do problema e n é a dimensão do espaço.

Mostramos no próximo resultado que 2^{-L} é de fato um número muito pequeno.

Lema 2.4.1 *Seja $v(P)$ o custo de uma solução ótima de (P) . Se x é um vértice do conjunto viável, então:*

(a) *Para $j = 1, \dots, n$, $x_j = 0$ ou $x_j > 2^{-L}$; e*

(b) *$c^T x = v(P)$ ou $c^T x > v(P) + 2^{-L}$.*

Demonstração. Lema 2.1 em Gonzaga [18]. ■

Este lema implica que se encontramos um vértice com custo menor do que $v(P) + 2^{-L}$, então ele deve ser uma solução ótima. Finalmente, dado x tal que

$c^T x \leq v(P) + 2^{-L}$, onde $v(P)$ é o custo de uma solução ótima, o procedimento de purificação deve resultar em um vértice ótimo.

Um pouco da história: Em 1984, Karmarkar [21] publicou seu algoritmo baixando a complexidade em relação ao método de Khachiyan [22]. Ele obteve um limite para o número de iterações de $O(nL)$ e um número de operações aritméticas por iteração de $O(n^{2.5})$, totalizando $O(n^{3.5}L)$ operações aritméticas. A trajetória central foi inicialmente estudada por Bayer e Lagarias [5] e Megiddo [30]. Em 1986, Renegar [46] provou que o método de centros de Huard [20] escrito em termos da função barreira logarítmica é polinomial para problemas de programação linear, se o problema de minimização auxiliar é tratado pelo método de Newton (Veja Nesterov [42]). Renegar obteve um limite de complexidade de $O(\sqrt{n}L)$ iterações, mas com complexidade total igual à de Karmarkar. Em 1987, Gonzaga [15] e Vaidya [59] obtiveram simultaneamente algoritmos com a complexidade de $O(n^3L)$, onde o primeiro desenvolve um algoritmo que usa a função penalizada barreira logarítmica e, o segundo, seguindo a mesma metodologia de Renegar. Recentemente, Anstreicher [4] obteve a complexidade de $O((n^3/\ln(n))L)$.

Neste ponto introduzimos a terceira hipótese para este capítulo, a saber:

(H_3) Um ponto inicial (x^0, s^0) e um valor $\mu^0 < 2^L$ estão disponíveis tais que $\delta(x^0, s^0, \mu^0) \leq 0.5$.

Agora estamos prontos para construirmos algoritmos de pontos interiores. É o que faremos na próxima seção.

2.5 Alguns algoritmos

Para “os algoritmos de pontos interiores”, sugerimos Gonzaga [17]. Nesta seção vamos investigar dois algoritmos. O primeiro deles é o algoritmo de trajetória central de passos curtos, que apesar de ineficiente na prática, caracteriza bem a natureza dos métodos de pontos interiores, no sentido que calcula $\mu > 0$ tal que a medida de proximidade é igual a $\alpha = 0.5$ (por exemplo) e encontra por Newton

um ponto próximo do novo ponto central. Realmente, poderíamos introduzir aqui o algoritmo preditor-corretor, que é eficiente na prática em virtude de seus passos serem longos. Todavia, faremos o seu desenvolvimento em uma abordagem mais geral no próximo capítulo, ou seja, a abordagem de ponto-interior-inviável. O segundo deles, então, é o algoritmo de redução potencial (primal-dual), cuja metodologia deu origem aos métodos de pontos interiores para programação linear através do premiado Karmarkar [21]. Ambos algoritmos têm complexidade polinomial e esse é o “charme” dos métodos de pontos interiores.

2.5.1 Passos curtos

Descrevemos o algoritmo de trajetória central de passos curtos e provamos que ele resolve o problema primal-dual (PD) em $O(\sqrt{n}L)$ iterações, onde cada iteração consiste em um passo de Newton. Inicialmente vamos enunciar o algoritmo.

Algoritmo 2.5.1 *Passos curtos.*

Dados: $\epsilon > 0$, $\mu^0 \in (0, 2^L)$ e $(x^0, s^0, \mu^0) \in \mathcal{V}(0.5)$.

$k := 0$.

REPITA

Calcule a direção (u, v) por (N) com $\mu = \mu^k$.

$(x^{k+1}, s^{k+1}) := (x^k, s^k) + (u, v)$.

$\mu^{k+1} := (1 - \frac{0.2}{\sqrt{n}})\mu^k$.

$k := k + 1$.

ATÉ QUE $\mu^k < \epsilon$.

O próximo resultado é fundamental para a convergência e para a complexidade em iterações deste algoritmo.

Lema 2.5.2 *O Algoritmo 2.5.1 gera uma seqüência de iterados tais que, para todo $k = 1, 2, \dots$, $(x^k, s^k, \mu^k) \in \mathcal{V}(0.5)$.*

Demonstração. Lema 5.2 em Gonzaga [18]. ■

Finalmente, obtemos o teorema de polinomialidade em número de iterações para o Algoritmo 2.5.1 que resolve o problema primal-dual (PD).

Teorema 2.5.3 *O Algoritmo 2.5.1 com $\epsilon = 2^{-2L}$ pára em $O(\sqrt{n}L)$ iterações com um par primal-dual (\hat{x}, \hat{s}) . Uma purificação de \hat{x} resulta em um vértice ótimo primal.*

Demonstração. Teorema 5.3 em Gonzaga [18]. ■

Pelos teoremas de dualidade e de folga complementar, a partir de um vértice ótimo primal podemos obter uma solução ótima dual.

2.5.2 Redução potencial

Investigamos uma função potencial primal-dual que mede o mérito de um par $(x, s) \in \mathcal{F}^0$ para a solução do problema primal-dual (PD) em $O(\sqrt{n}L)$ passos.

É natural combinarmos uma função monótona do *gap* de complementaridade $x^T s$ com termos barreira para ambos x e s . Definimos a função potencial primal-dual, denominada função potencial de Tanabe-Todd-Ye (trabalhos independentes: Tanabe [49] e Todd e Ye [56]), a saber: para $q \geq n$,

$$(x, s) \in \mathcal{F}^0 \mapsto f_q(x, s) = q \ln(x^T s) + p(x) + p(s).$$

O desenvolvimento que se segue tem a seguinte estratégia: observamos na definição de $f_q(\cdot, \cdot)$ que

$$\exp(f_q(x, s)) = \frac{(x^T s)^q}{\prod_{j=1}^n x_j s_j}.$$

Desta forma, se obtivermos $f_q(x, s) \rightarrow -\infty$ e $q > n$, devemos mostrar que isto implica em $x^T s \rightarrow 0$. Daí, devemos garantir um decréscimo constante da função potencial $f_q(\cdot, \cdot)$ para obtermos um algoritmo de redução potencial com a complexidade desejada.

O próximo resultado afirma que a função potencial primal-dual com $q = n$, mede a proximidade da trajetória central.

Proposição 2.5.4 Para vetores positivos x e s em R^n ,

$$n \ln(x^T s) + p(x) + p(s) \geq n \ln(n),$$

valendo a igualdade se, e somente se, $xs = \mu e$ para algum $\mu > 0$.

Demonstração. Proposição 3.1 em Todd [55]. ■

Existe uma relação muito próxima entre a função potencial primal-dual e o *gap* de complementaridade, conforme veremos agora.

Proposição 2.5.5 Para vetores positivos x e s em R^n ,

$$f_q(x, s) \geq (q - n) \ln(x^T s) + n \ln(n),$$

valendo a igualdade se, e somente se, $xs = \mu e$ para algum $\mu > 0$.

Demonstração. Páginas 25 e 26 em Todd [55]. ■

Colocando nossa estratégia em prática, no que concerne a esta subseção, vamos iniciar mostrando que a função potencial primal-dual não é limitada inferiormente, usando as hipóteses (H_1) e (H_2) . Para esse enunciado estaremos usando o Teorema 2.1.5.

Teorema 2.5.6 Se \mathcal{F}^0 é não vazio e $q > n$, então f_q não é limitada inferiormente.

Demonstração. Teorema 5.1 em Todd [55]. ■

Temos a seguir o resultado fundamental para a concretização da estratégia que estamos desenvolvendo nesta subseção.

Teorema 2.5.7 Suponhamos que \mathcal{F}^0 é não vazio, $q > n$ e $f_q(x, s) \rightarrow -\infty$. Então $c^T x \rightarrow v(P)$, $b^T y \rightarrow v(D)$ e $v(P) = v(D)$. De fato,

$$c^T x - b^T y = x^T s = \exp\left(\frac{f_q(x, s)}{q - n}\right).$$

Demonstração. Teorema 5.2 em Todd [55]. ■

Usaremos a hipótese (H_3) para o próximo resultado, que é uma consequência deste último teorema.

Corolário 2.5.8 *Suponhamos que dispomos de algum par $(x, s) \in \mathcal{F}^0$ e de um algoritmo que em cada iteração substitui $(x^k, s^k) \in \mathcal{F}^0$ por $(x^{k+1}, s^{k+1}) \in \mathcal{F}^0$, tal que*

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

para alguma constante positiva β . Então, para qualquer $\epsilon > 0$,

$$(x^K)^T s^K \leq \epsilon$$

com

$$K = \frac{q - n}{\beta} \left[\frac{f_q(x^0, s^0)}{q - n} + \ln\left(\frac{1}{\epsilon}\right) \right]$$

iterações.

Demonstração. Corolário 5.1 em Todd [55]. ■

A interpretação para este corolário é a seguinte: se a função potencial de Tanabe-Todd-Ye, f_q , é reduzida em cada iteração por uma constante $\beta > 0$ para $q = n + O(\sqrt{n})$, então obtemos um algoritmo com complexidade $O(\sqrt{n} \ln(1/\epsilon))$ passos. Assim, concluímos a nossa estratégia motivadora.

Capítulo 3

Métodos de ponto-interior-inviável

Algoritmo de ponto-interior-inviável é um algoritmo de pontos interiores (pontos no primeiro octante com coordenadas estritamente positivas) tal que o ponto inicial não é restrito a ser um ponto viável. A propósito, é neste sentido que retiraremos a hipótese (H_3) do capítulo anterior. Conforme afirmamos no capítulo 1, os algoritmos de ponto-interior-inviável primais-duais são conhecidos como os algoritmos de ponto interior mais eficientes na prática.

Neste capítulo mostramos um tratamento simples de métodos de ponto-interior-inviável para programação linear. Descrevemos um algoritmo preditor-corretor (primal-dual) que segue aproximadamente a superfície de centros, conforme Mizuno, Todd e Ye, e mostramos que ele resolve o problema de programação linear primal-dual ou detecta inviabilidade para uma certa região em $O(nL)$ iterações. Se uma certa medida de proximidade é suficientemente pequena no ponto inicial, então o algoritmo termina em $O(\sqrt{n}L)$ passos. Além disso, introduzimos o algoritmo artificial denominado algoritmo homogêneo e auto-dual, o qual se apresenta como um tipo especial de métodos de ponto-interior-inviável. Sua complexidade em iterações é $O(\sqrt{n}L)$. Novamente, como no capítulo anterior, as demonstrações serão, em princípio, todas elas referenciadas.

Sobre o algoritmo preditor-corretor que apresentamos aqui, Stoer [48] obteve esse algoritmo sem detectar inviabilidade para uma certa região, obtida por Mizuno

em [50]. O mesmo algoritmo pode ser obtido por Sheng e Potra [47] e Mizuno, Jarre e Stoer [39] para programação linear.

Gostaria de ressaltar que este capítulo foi escrito baseado no artigo de Sheng e Potra [47], Mizuno [35] e em uma idéia de que um algoritmo simples no sentido do já conhecido preditor-corretor de Mizuno, Todd e Ye [36] deveria funcionar para ponto-interior-inviável. Alguns dias depois de escrever a primeira versão deste capítulo, tive acesso ao livro editado por Terlaky [50]. No capítulo 5 deste livro, escrito por Mizuno, ele apresenta o algoritmo que desenvolvemos em um contexto muito mais rico do que teremos aqui. Nesta referência conheci o artigo de Stoer [48]. De qualquer forma, devido ao meu estudo referente ao artigo de Sheng e Potra [47], valeu-me uma nota de agradecimentos no artigo deles.

3.1 O problema primal-dual e a superfície de centros

Conforme o capítulo anterior, consideremos o problema primal-dual, isto é, encontrar, se existir, uma solução para o sistema

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

Existem três dificuldades básicas para resolver o sistema (PD) . A primeira é o caso deste sistema não ter solução. A segunda, consiste no problema combinatorial de determinar qual de cada par de variáveis (x_j, s_j) , $j = 1, \dots, n$ será nula em uma solução ótima. E, em se tratando de métodos de pontos interiores, a terceira dificuldade consiste na obtenção do ponto interior viável inicial para o par primal e dual original.

O que faremos a seguir será contornar estas três dificuldades em um caminho

geral e usual.

Dados $x^0 \in R_{++}^n$ e $(y^0, s^0) \in R^m \times R_{++}^n$, definimos os resíduos primal e dual iniciais, respectivamente, por

$$r_P^0 = b - Ax^0 \text{ e } r_D^0 = c - A^T y^0 - s^0.$$

Definimos para cada $\theta \in R$, o conjunto corrente primal-dual por

$$\mathcal{F}_\theta = \{(x, y, s) \in R_+^n \times R^m \times R_+^n; Ax = b - \theta r_P^0, A^T y + s = c - \theta r_D^0\},$$

e o seu conjunto de pontos interiores

$$\mathcal{F}_\theta^0 = \{(x, y, s) \in \mathcal{F}_\theta; (x, s) > 0\}.$$

Em particular, uma vez que assumimos *rank* completo para a matriz A , o vetor y é unicamente determinado para x e s nas equações de viabilidade conforme \mathcal{F}_θ . Assim, da mesma maneira que no capítulo anterior, podemos nos referir a uma solução primal-dual simplesmente por (x, s) ao invés de (x, y, s) . Denotamos, então,

$$\tilde{\mathcal{F}}_\theta = \{(x, s) \in R_+^n \times R_+^n; (x, y, s) \in \mathcal{F}_\theta, \text{ para algum } y \in R^m\},$$

e o seu conjunto de pontos interiores

$$\tilde{\mathcal{F}}_\theta^0 = \{(x, s) \in \tilde{\mathcal{F}}_\theta; (x, s) > 0\}.$$

Mizuno, Todd e Ye [37] definem os limitantes inferior e superior, respectivamente, para o parâmetro θ , a saber:

$$\theta_l = \inf\{\theta; \mathcal{F}_\theta \neq \emptyset\}$$

e

$$\theta_u = \sup\{\theta; \mathcal{F}_\theta \neq \emptyset\}.$$

É importante observar que o problema primal-dual (PD) tem um ponto interior viável se, e só se, $\theta_l < 0$; (PD) é viável e não tem ponto interior viável se, e só se, $\theta_l = 0$; e, se (PD) é inviável, então $\theta_l > 0$.

Para $\theta_l \in (-\infty, 1)$ e $\theta_u \in (1, \infty)$, define-se o conjunto de parâmetros

$$\mathcal{T} = \{(\theta, \mu); \theta \in [\theta_l, \theta_u], \mu \geq 0\}$$

e o seu interior

$$\mathcal{T}^0 = \{(\theta, \mu); \theta \in (\theta_l, \theta_u), \mu > 0\}.$$

Mizuno, Todd e Ye [37], provaram que o sistema perturbado

$$\begin{aligned} Ax &= b - \theta r_P^0 \\ (PD)_{\theta, \mu} \quad A^T y + s &= c - \theta r_D^0 \\ xs &= \mu e \\ x, s &\geq 0, \end{aligned}$$

define uma superfície

$$\mathcal{S} = \{(x(\theta, \mu), y(\theta, \mu), s(\theta, \mu), \theta, \mu); (\theta, \mu) \in \mathcal{T}^0\},$$

onde $(x(\theta, \mu), y(\theta, \mu), s(\theta, \mu))$ resolve unicamente $(PD)_{\theta, \mu}$ com $(\theta, \mu) \in \mathcal{T}^0$.

Lema 3.1.1 *O conjunto \mathcal{S} é uma superfície (variedade bi-dimensional).*

Demonstração. Lema 2 em Mizuno, Todd e Ye [37]. ■

A escolha padrão de pontos iniciais é dada por

$$x^0 = \rho e, \quad y^0 = 0 \quad \text{e} \quad s^0 = \rho e,$$

onde

$$\rho \geq \max\{\|A^T(AA^T)^{-1}b\|_\infty, \|c\|_\infty\}.$$

Conforme Mizuno [50], suponhamos que estamos interessados em encontrar soluções ótimas x^* de (P) e (y^*, s^*) de (D) , se existirem, somente no conjunto

$$\mathcal{B}_\rho = \{(x, y, s); \|(x, s)\|_\infty \leq \rho\}.$$

Isto significa que apesar de não supormos a existência de soluções ótimas, a superfície de centros estará confinada a uma região tipo caixa que depende de ρ . O parâmetro ρ é usado nos algoritmos somente para detectar a não existência de soluções em \mathcal{B}_ρ . Logo, podemos colocar $\rho = \infty$. Se todos os dados A, b e c são inteiros, então o problema primal-dual original tem uma solução em \mathcal{B}_ρ para $\rho \leq 2^L$ ou o problema é inviável. Kojima [23] apresenta alguns lemas básicos que trazem maior flexibilidade e melhoramento para tais restrições sobre o ponto inicial; inclusive.

Neste ponto contornamos as três dificuldades iniciais para resolver o problema original (PD) da seguinte maneira: consideremos $\rho \leq 2^L$ e, a partir de um ponto inicial em \mathcal{S} , digamos $(\rho e, 0, \rho e, 1, \rho^2) \in \mathcal{S}$, devemos encontrar aproximadamente uma solução para cada sistema $(PD)_{\theta, \mu}$ fazendo $\theta \rightarrow 0$ e $\mu \rightarrow 0$, ou identificar que não existe tal solução em \mathcal{B}_ρ . A propósito, é neste sentido que retiraremos as hipóteses (H_1) e (H_2) do capítulo anterior.

3.2 A medida de proximidade e o passo de Newton

A palavra ‘aproximadamente’ no final da seção anterior se refere a alguma medida de proximidade que definimos agora. Dado $\mu > 0$ queremos encontrar o par (x, s) tal que $xs/\mu = e$. Isto porque os métodos usados devem resolver bem as igualdades lineares associadas à viabilidade no problema perturbado primal-dual. Assim, definimos a medida de proximidade por

$$(x, s, \mu) \in \tilde{\mathcal{F}}_\theta^0 \times R_{++} \mapsto \delta(x, s, \mu) = \left\| \frac{xs}{\mu} - e \right\|,$$

da mesma forma como em pontos interiores (viáveis).

Métodos de ponto-interior-inviável trabalham com pontos próximos da superfície de centros. Dados os parâmetros $\alpha \in (0, 1)$ (usualmente $\alpha = 0.25$) e $\theta \in (\theta_l, \theta_u)$, a vizinhança da superfície de centros é definida por

$$\mathcal{V}(\alpha, \theta) = \{(x, s, \mu) \in \tilde{\mathcal{F}}_\theta^0 \times R_{++}; \delta(x, s, \mu) \leq \alpha\}.$$

Note que a escolha do ponto inicial de fato o caracteriza como um ponto central, tomando $\theta^0 = 1$ e $\mu^0 = \rho^2$. Com efeito,

$$\delta(x^0, s^0, \mu^0) = \left\| \frac{x^0 s^0}{\mu^0} - e \right\| = 0.$$

Então $(x^0, y^0, s^0, \theta^0, \mu^0) \in \mathcal{S}$.

O próximo resultado será importante para a convergência do algoritmo preditor-corretor que desenvolvemos neste capítulo, devido a Mizuno, Todd e Ye.

Teorema 3.2.1 *Suponha que $\theta_l \leq 0$ (o par (P) e (D) é não vazio). Seja $((x^k, y^k, s^k, \theta^k, \mu^k))$ uma seqüência com $(x^k, s^k, \mu^k) \in \mathcal{V}(\alpha, \theta^k)$, $\alpha \in (0, 1)$, tal que $\theta^k \rightarrow 0$ e $\mu^k \rightarrow 0$. Se θ^k / μ^k é limitado, então todo ponto limite de $((x^k, y^k, s^k))$ é um ponto interior (relativo) do conjunto de soluções ótimas*

$$\{(x, y, s); Ax = b, A^T y + s = c, x^T s = 0, x \geq 0, s \geq 0\}.$$

Demonstração. Segunda parte do Teorema 4 e primeira parte do Teorema 5, ambos, de Mizuno, Todd e Ye [37]. ■

Sejam o par de parâmetros $(\theta, \mu) \in \mathcal{T}^0$ e $\alpha \in (0, 1)$, tais que o ponto $(x, s, \mu) \in \mathcal{V}(\alpha, \theta)$. Considere $\nu \in R$ e $\gamma \geq 0$ associados aos parâmetros θ e μ , respectivamente. Idealmente, gostaríamos de encontrar

$$x^+ = x + u, \quad y^+ = y + w \quad \text{e} \quad s^+ = s + v,$$

tais que $(x^+, s^+) \in \tilde{\mathcal{F}}_{\nu\theta}$ e $x^+ s^+ = \gamma \mu e$.

O passo de Newton resolve isto aproximadamente linearizando o sistema perturbado $(PD)_{\theta, \mu}$, a saber:

$$\begin{aligned}
Au &= (1 - \nu)\theta r_P^0 \\
(N) \quad A^T w + v &= (1 - \nu)\theta r_D^0 \\
su + xv &= -xs + \gamma\mu e,
\end{aligned}$$

o qual possui uma única solução, conforme Proposição 2.2.1.

O sistema (N) é geral para métodos de ponto-interior-inviável primal-dual para programação linear.

O lema a seguir foi usado por Mizuno [35], para valores particulares dos parâmetros ν e γ .

Lema 3.2.2 *Considere o sistema (N). Então:*

$$(a) \quad D^{-1}u = -(1 - \nu)\theta \tilde{P}_{AD} D^{-1}(x^0 - \hat{x}) + (1 - \nu)\theta P_{AD} D(s^0 - \hat{s}) - P_{AD}(XS)^{-\frac{1}{2}}(xs - \gamma\mu e);$$

$$(b) \quad w = -(1 - \nu)\theta(y^0 - \hat{y}) - (AD^2 A^T)^{-1} AD[(1 - \nu)\theta D^{-1}(x^0 - \hat{x}) + (1 - \nu)\theta D(s^0 - \hat{s}) - (XS)^{-\frac{1}{2}}(xs - \gamma\mu e)]; \text{ e}$$

$$(c) \quad Dv = (1 - \nu)\theta \tilde{P}_{AD} D^{-1}(x^0 - \hat{x}) - (1 - \nu)\theta P_{AD} D(s^0 - \hat{s}) - \tilde{P}_{AD}(XS)^{-\frac{1}{2}}(xs - \gamma\mu e),$$

onde $D = X^{\frac{1}{2}} S^{-\frac{1}{2}}$, $\tilde{P}_{AD} = DA^T(AD^2 A^T)^{-1} AD$, $P_{AD} = I - \tilde{P}_{AD}$ e $(\hat{x}, \hat{y}, \hat{s})$ é uma solução de $Ax = b$ e $A^T y + s = c$.

Demonstração. Lema 5.3.3 em Mizuno [50]. ■

Agora estamos prontos para construirmos algoritmos de ponto-interior-inviável. É o que faremos na próxima seção.

3.3 Alguns algoritmos

Iniciamos esta seção enunciando um algoritmo de ponto-interior-inviável para programação linear. O restante desta seção concerne ao desenvolvimento dos resultados acerca desse algoritmo. Esses resultados acompanham os métodos de ponto-interior-inviável em geral. Finalmente, estudamos o algoritmo homogêneo e

auto-dual para programação linear devido Ye, Todd e Mizuno [71], o qual é a base para os artigos de Xu e Ye [60], Xu, Hung e Ye [61] e Tütüncü [58]. Uma curiosidade, é a motivação para a homogeneidade e auto-dualidade na página 94 do livro de Maculan e Pereira [27].

3.3.1 Preditor-corretor

Algoritmo 3.3.1 *Preditor-corretor.*

Dados: $\epsilon = 2^{-L}$, $\rho \in [1, 2^L]$, $\alpha := 0.25$ e $(x^0, y^0, s^0, \theta^0, \mu^0) := (\rho e, 0, \rho e, 1, \rho^2)$.

$k := 0$.

REPITA

$x := x^k$, $y := y^k$, $s := s^k$, $\theta := \theta^k$, $\mu := \mu^k$.

Passo preditor:

Resolva (N) *com* $\nu = 0$ *e* $\gamma = 0$ *para* (u, w, v) .

$\bar{\lambda} := \max\{\lambda \in [0, 1]; (x, y, s) + \lambda(u, w, v) \in \mathcal{V}(2\alpha, (1 - \lambda)\theta)\}$.

$(\bar{x}, \bar{y}, \bar{s}) := (x, y, s) + \bar{\lambda}(u, w, v)$.

$\theta^{k+1} := (1 - \bar{\lambda})\theta$ e $\mu^{k+1} := (1 - \bar{\lambda})\mu$.

Passo corretor:

Resolva (N) *com* $\nu = 1$ *e* $\gamma = 1$ *para* $(\bar{u}, \bar{w}, \bar{v})$.

$(x^{k+1}, y^{k+1}, s^{k+1}) := (\bar{x}, \bar{y}, \bar{s}) + (\bar{u}, \bar{w}, \bar{v})$.

$k := k + 1$.

ATÉ QUE $(\mu^k < \epsilon)$ *ou* $(\|(x^k, s^k)\|_1 > \frac{3}{\theta^k \rho} (x^k)^T s^k$ e $\theta^k > 0)$.

A estratégia para demonstrarmos a complexidade em iterações do Algoritmo 3.3.1 é a mesma para algoritmos de pontos interiores (viáveis) preditor-corretor. Isto é, uma vez que atualizamos $\mu^{k+1} = (1 - \bar{\lambda})\mu^k$, devemos estabelecer um limitante inferior para qualquer $\bar{\lambda}$ gerado pelo algoritmo. De fato, $\bar{\lambda}$ depende de $\|uv/\mu^k\|$, onde u e v são direções preditoras. O que diferirá no nosso caso (algoritmo de ponto-interior-inviável), é que no passo preditor $u^T v \in R$, enquanto que em pontos interiores (viáveis) $u^T v \geq 0$ (como é o caso de problemas de complementaridade linear com a hipótese de monotonicidade e, igual a zero, em programação linear).

Logo, $\|uv/\mu^k\|$ deve ser avaliada de outra maneira, como veremos adiante.

O próximo resultado é fundamental para a convergência e para a complexidade em iterações do Algoritmo 3.3.1. A demonstração usa o Lema 2.3.3. Para pontos interiores (viáveis), esse é o resultado de Mizuno, Todd e Ye [36].

Lema 3.3.2 *Para qualquer inteiro $k \geq 0$, o Algoritmo 3.3.1 define uma terna*

$$(x^k, s^k, \mu^k) \in \mathcal{V}(\alpha, \theta^k),$$

com

$$\mu^k = \frac{(x^k)^T s^k}{n}$$

e os correspondentes resíduos satisfazem

$$r_P^k = \theta^k r_P^0 \quad e \quad r_D^k = \theta^k r_D^0,$$

onde

$$\theta^0 = 1 \quad e \quad \theta^k = \prod_{i=0}^{k-1} (1 - \bar{\lambda}_i).$$

Demonstração. Teorema 2.4 em Sheng e Potra [47]. ■

O maior comprimento de passo satisfazendo $\mathcal{V}(2\alpha, (1 - \lambda)\theta)$, $\lambda \in [0, 1]$, é fornecido no próximo resultado, devido a Ye, Tapia e Zhang [69]. Veja também Ye, Güler, Tapia e Zhang [70].

Lema 3.3.3 *Considere o passo predictor do Algoritmo 3.3.1. Então, para cada $k \geq 0$,*

$$\bar{\lambda} \geq \lambda^+ = \frac{2}{1 + \sqrt{1 + 4\|uv/\mu^k\|/\alpha}}. \tag{3.1}$$

Demonstração. Lema 3.3 em Sheng e Potra [47]. ■

A seguir demonstramos o resultado que determinará a complexidade em iterações do Algoritmo 3.3.1. Ele estabelece um limitante superior para o valor de

$\|uv/\mu^k\|$ que aparece no cálculo de λ^+ do lema anterior. A demonstração usa o Lema 3.2.2 para $\nu = \gamma = 0$ e um raciocínio análogo ao apresentado na página 115 em Mizuno [35].

Lema 3.3.4 *Seja o par (u, v) obtido na k -ésima iteração no passo preditor do Algoritmo 3.3.1. Então,*

$$\left\| \frac{uv}{\mu^k} \right\| = O(n^2).$$

Demonstração. Denotemos no passo preditor do algoritmo, $\mu = \mu^k$, $\theta = \theta^k$, $x = x^k$ e $s = s^k$. No passo preditor, (u, v) é obtido para $\nu = \gamma = 0$. Usando o Lema 3.2.2 e pelo fato de que as matrizes \tilde{P}_{AD} e P_{AD} são matrizes de projeção ortogonal,

$$\|d^{-1}u\| \leq \theta \|d^{-1}(x^0 - \hat{x})\| + \theta \|d(s^0 - \hat{s})\| + \sqrt{x^T s}$$

e

$$\|dv\| \leq \theta \|d^{-1}(x^0 - \hat{x})\| + \theta \|d(s^0 - \hat{s})\| + \sqrt{x^T s}.$$

Pela semelhança do lado direito destas duas desigualdades, basta desenvolver uma delas. Desenvolveremos a primeira. Pela definição de ρ , podemos assumir $\|(\hat{x}, \hat{s})\|_\infty \leq \rho$. Então, para $x^0 = s^0 = \rho e$,

$$0 \leq x^0 - \hat{x} \leq 2\rho e \quad \text{e} \quad 0 \leq s^0 - \hat{s} \leq 2\rho e.$$

implicando

$$\begin{aligned} \|d^{-1}u\| &\leq \theta 2\rho \|d^{-1}\| + \theta 2\rho \|d\| + \sqrt{x^T s} \\ &\leq 2\theta\rho \|\sqrt{1/xs}\|_\infty (\|s\|_1 + \|x\|_1) + \sqrt{x^T s}, \end{aligned}$$

onde verificando a definição de d a última desigualdade decorre da propriedade de normas de vetores $\|az\| \leq \|a\|_\infty \|z\| \leq \|a\|_\infty \|z\|_1$, para todos $a, z \in R^n$. Agora, de $\delta(x, s, \mu) \leq \alpha$ e tomando $\mu = x^T s/n$ obtemos, para $j = 1, \dots, n$,

$$\frac{1}{x_j s_j} \leq \frac{1}{(1-\alpha)\mu} = \frac{n}{(1-\alpha)x^T s}$$

e, considerando que o algoritmo ainda não parou,

$$\|x\|_1 + \|s\|_1 = \|(x, s)\|_1 \leq \frac{3}{\theta\rho} x^T s,$$

se $\theta > 0$. De fato, como veremos adiante, o resultado valerá para $\theta = 0$. Segue-se que

$$\|d^{-1}u\| \leq 6\sqrt{\frac{n}{(1-\alpha)x^T s}} x^T s + \sqrt{x^T s}.$$

Então,

$$\|d^{-1}u\| \leq \left(\frac{6\sqrt{n}}{\sqrt{(1-\alpha)}} + 1\right)\sqrt{x^T s}.$$

Neste ponto, note que $\|d^{-1}u\| \leq \sqrt{x^T s} \leq \left(\frac{6\sqrt{n}}{\sqrt{(1-\alpha)}} + 1\right)\sqrt{x^T s}$, se $\theta = 0$. Da mesma forma,

$$\|dv\| \leq \left(\frac{6\sqrt{n}}{\sqrt{(1-\alpha)}} + 1\right)\sqrt{x^T s}.$$

Usando a mudança de escala, $d^{-1}u = \bar{u}$ e $dv = \bar{v}$ tal que $uv = \bar{u}\bar{v}$, temos

$$\left\|\frac{uv}{\mu}\right\| = \left\|\frac{\bar{u}\bar{v}}{\mu}\right\| \leq \left\|\frac{\bar{u}}{\mu}\right\| \|\bar{v}\|.$$

Portanto, tomando $\mu = \frac{x^T s}{n}$ e as definições de \bar{u} e \bar{v} ,

$$\left\|\frac{uv}{\mu}\right\| \leq \left(\frac{6\sqrt{n}}{\sqrt{(1-\alpha)}} + 1\right)^2 n = O(n^2).$$

Isto finaliza a demonstração. ■

Uma expressão útil para o limitante superior de $\|uv/\mu^k\|$, $k \geq 0$, é a seguinte: sabemos pela demonstração do lema anterior que

$$\begin{aligned}
\left\| \frac{uv}{\mu^k} \right\| &\leq n \left(\frac{6\sqrt{n}}{\sqrt{1-\alpha}} + 1 \right)^2 \\
&= n \left(\frac{36n + 12\sqrt{1-\alpha}\sqrt{n} + (1-\alpha)}{1-\alpha} \right) \\
&\leq n \left(\frac{36n + 12n + 1}{1-\alpha} \right).
\end{aligned}$$

Logo,

$$\left\| \frac{uv}{\mu^k} \right\| \leq \frac{49n^2}{\alpha}. \quad (3.2)$$

A utilidade de (3.2) será estabelecida agora. Isto é, usando (3.2) na desigualdade (3.1) do Lema 3.3.3 vem, para $\alpha = 0.25$,

$$\begin{aligned}
\bar{\lambda} &\geq \frac{2}{1 + \sqrt{1 + \frac{196n^2}{\alpha^2}}} \\
&\geq \frac{2}{1 + 112n} \\
&\geq \frac{0.009}{n}.
\end{aligned}$$

Daí, em toda iteração do Algoritmo 3.3.1,

$$1 - \bar{\lambda} \leq 1 - \frac{0.009}{n}. \quad (3.3)$$

Agora, devemos mostrar que a seqüência gerada pelo Algoritmo 3.3.1 converge.

Teorema 3.3.5 *A seqüência $((x^k, y^k, s^k))$ gerada pelo Algoritmo 3.3.1 sem critério de parada é limitada e aproxima-se de uma solução complementar estrita se, e somente se, $\theta_l \leq 0$.*

Demonstração. Corolário 4 em Mizuno, Todd e Ye [37]. ■

No critério de parada do Algoritmo 3.3.1 detectamos inviabilidade para uma certa região definida na norma $\|(\cdot, \cdot)\|_1$ e que depende do parâmetro ρ . Este resultado é devido a Mizuno [35].

Lema 3.3.6 *Se*

$$\|(x^k, s^k)\|_1 > \frac{3}{\theta^k \rho} (x^k)^T s^k$$

vale na k -ésima iteração do Algoritmo 3.3.1 então não existem soluções ótimas x^ de (P) e (y^*, s^*) de (D) tais que $\|(x^*, s^*)\|_\infty \leq \rho$.*

Demonstração. Lema 3.3 em Mizuno [35]. ■

Enfim, estamos prontos para demonstrar o resultado principal para o algoritmo preditor-corretor que estamos desenvolvendo, o qual afirma que o Algoritmo 3.3.1 termina em $O(nL)$ passos.

Teorema 3.3.7 *O Algoritmo 3.3.1 termina em $O(nL)$ iterações. Além disso, se o algoritmo pára com*

$$\|(x^k, s^k)\|_1 > \frac{3}{\theta^k \rho} (x^k)^T s^k$$

e $\theta^k > 0$, então não existe solução ótima primal-dual (x^*, y^*, s^*) tal que $\|(x^*, s^*)\|_\infty \leq \rho$.

Demonstração. Teorema 4.1 em Mizuno [35]. ■

Finalmente, vamos mostrar conforme Sheng e Potra [47] e Mizuno [35], que se uma certa medida de viabilidade é suficientemente pequena no ponto inicial, então o algoritmo termina em $O(\sqrt{n}L)$ passos.

Teorema 3.3.8 *Considere $\tau > 0$. Para um dado ponto inicial (x^0, y^0, s^0) com $(x^0, s^0, (x^0)^T s^0/n) \in \mathcal{V}(0.25, 1)$, se existir a terna $(\hat{x}, \hat{y}, \hat{s})$ satisfazendo $A\hat{x} = b$ e $A^T \hat{y} + \hat{s} = c$, tal que*

$$\|(x^0)^{-1}(x^0 - \hat{x}), (s^0)^{-1}(s^0 - \hat{s})\|_\infty \leq \min\{0.5, \frac{\tau}{\sqrt{n}}\},$$

então o Algoritmo 3.3.1 termina em $O(\sqrt{n}L)$ iterações na mesma situação do Teorema 3.3.7.

Demonstração. Teorema 5.7.4 em Mizuno [50]. ■

Se o Algoritmo 3.3.1 pára com $\mu^k < \epsilon$, então uma purificação de x^k resulta em um vértice primal ótimo. Veja a análise desenvolvida na seção 4 do capítulo anterior. Daí, pela dualidade (forte), podemos obter uma solução ótima dual.

3.3.2 Homogêneo e auto-dual

O método de ponto-interior-inviável homogêneo e auto-dual para programação linear, consiste em construir um problema de programação linear artificial homogêneo e auto-dual relacionado com o problema primal-dual (PD). Aqui, a homogeneidade do problema significa que todo o lado direito das restrições é igual a zero, exceto para uma delas, freqüentemente chamada de restrição de normalização. Por outro lado, a auto-dualidade do problema significa que o dual é equivalente ao primal.

Ye, Todd e Mizuno (YTM) apresentam o problema de programação linear artificial homogêneo e auto-dual relacionado com o problema (PD), a saber: dados $x^0 > 0$, $s^0 > 0$ e y^0 ,

$$\begin{aligned}
 (HLP) \quad & \text{minimizar} && ((x^0)^T s^0 + 1)\theta \\
 & \text{sujeito a:} && Ax - b\tau + r_P^0 \theta = 0 \\
 & && -A^T y + c\tau - r_D^0 \theta \geq 0 \\
 & && b^T y - c^T x + \bar{z}\theta \geq 0 \\
 & && -(r_P^0)^T y + (r_D^0)^T x - \bar{z}\tau = -(x^0)^T s^0 - 1 \\
 & && x \geq 0, \tau \geq 0,
 \end{aligned}$$

onde

$$\bar{z} = c^T x^0 + 1 - b^T y^0.$$

Aqui \bar{z} representa a “inviabilidade” do *gap* de complementaridade. Note que fazendo τ uma variável (homogeneizando!), adiciona-se a variável dual requerida para a terceira restrição. Então, para atingir viabilidade para $x = x^0$ e $(y, s) = (y^0, s^0)$, adiciona-se a variável artificial θ com os coeficientes apropriados e, então, a última restrição de igualdade é adicionada para atingir auto-dualidade. Ou seja, os coeficientes do lado esquerdo das restrições em (HLP) formam uma matriz anti-simétrica.

Denote por

$$s = -A^T y + c\tau - r_D^0 \theta \quad \text{e} \quad \kappa = b^T y - c^T x + \bar{z}\theta.$$

Note que combinando as restrições de (HLP) , podemos escrever a última restrição de igualdade como

$$(s^0)^T x + (x^0)^T s + \tau + \kappa - ((x^0)^T s^0 + 1)\theta = (x^0)^T s^0 + 1,$$

que serve como uma restrição de normalização para (HLP) .

Denote por \mathcal{F}_h o conjunto de todos os pontos $(y, x, \tau, \theta, s, \kappa)$ que são viáveis para (HLP) . Denote por \mathcal{F}_h^0 o conjunto de pontos viáveis estritamente com $(x, \tau, s, \kappa) > 0$ em \mathcal{F}_h . O teorema a seguir define um caminho em (HLP) , segundo YTM.

Teorema 3.3.9 (i) *Para qualquer $\mu > 0$, existe um único ponto $(y, x, \tau, \theta, s, \kappa) \in \mathcal{F}_h^0$ tal que*

$$\begin{pmatrix} xs \\ \tau\kappa \end{pmatrix} = \mu e; \quad e$$

(ii) *Considere $(dy, dx, d\tau, d\theta, ds, d\kappa)$ no espaço nulo Q da matriz de restrições de (HLP) após adicionar variáveis excedentes s e κ , isto é,*

$$\begin{aligned} Adx - bd\tau + r_P^0 d\theta &= 0 \\ -A^T dy + cd\tau - r_D^0 d\theta - ds &= 0 \\ b^T dy - c^T dx + \bar{z}d\theta - d\kappa &= 0 \\ -(r_P^0)^T dy + (r_D^0)^T dx - \bar{z}d\tau &= 0. \end{aligned}$$

Então

$$dx^T ds + d\tau d\kappa = 0.$$

Demonstração. Teorema 5 de Ye, Todd e Mizuno. ■

O Teorema 3.3.9 define um caminho em (HLP)

$$\mathcal{C}_{YTM} = \{(y, x, \tau, \theta, s, \kappa) \in \mathcal{F}_h^0; \begin{pmatrix} xs \\ \tau\kappa \end{pmatrix} = \frac{x^T s + \tau\kappa}{n+1}\},$$

denominado caminho auto-central de (HLP).

A seguir vamos estudar o algoritmo homogêneo e auto-dual preditor-corretor.

Para $\beta \in (0, 1)$ dado, YTM definem uma vizinhança do caminho \mathcal{C}_{YTM} dada por

$$\mathcal{V}(\beta) = \{(y, x, \tau, \theta, s, \kappa) \in \mathcal{F}_h^0; \|\begin{pmatrix} xs \\ \tau\kappa \end{pmatrix} - \mu e\| \leq \beta\mu, \text{ onde } \mu = \frac{x^T s + \tau\kappa}{n+1}\}.$$

Algoritmo 3.3.10 *Homogêneo e auto-dual.*

Dados: $\epsilon > 0$, $\beta := 0.25$, $\mu^0 := 1$ e $(y^0, x^0, \tau^0, \theta^0, s^0, \kappa^0) := (0, e, 1, 1, e, 1) \in \mathcal{V}(\beta)$.

$k := 0$.

REPITA

$y := y^k$, $x := x^k$, $\tau := \tau^k$, $\theta := \theta^k$, $s := s^k$, $\kappa := \kappa^k$, $\mu := \mu^k$.

Passo preditor:

Calcule $(dy, dx, d\tau, d\theta, ds, d\kappa) \in Q$ tal que

$$\begin{pmatrix} xds + sdx \\ \tau d\kappa + \kappa d\tau \end{pmatrix} = \gamma\mu e - \begin{pmatrix} xs \\ \tau\kappa \end{pmatrix},$$

para $\gamma = 0$.

Calcule

$\bar{\alpha} := \max\{\alpha; (y + \alpha dy, x + \alpha dx, \tau + \alpha d\tau, \theta + \alpha d\theta, s + \alpha ds, \kappa + \alpha d\kappa) \in \mathcal{V}(2\beta)\}$.

$(\bar{y}, \bar{x}, \bar{\tau}, \bar{\theta}, \bar{s}, \bar{\kappa}) := (y + \bar{\alpha} dy, x + \bar{\alpha} dx, \tau + \bar{\alpha} d\tau, \theta + \bar{\alpha} d\theta, s + \bar{\alpha} ds, \kappa + \bar{\alpha} d\kappa)$.

Passo corretor:

Calcule $(\bar{d}y, \bar{d}x, \bar{d}\tau, \bar{d}\theta, \bar{d}s, \bar{d}\kappa) \in Q$ tal que

$$\begin{pmatrix} \bar{x}\bar{d}s + \bar{s}\bar{d}x \\ \bar{\tau}\bar{d}\kappa + \bar{\kappa}\bar{d}\tau \end{pmatrix} = \gamma\mu e - \begin{pmatrix} \bar{x}\bar{s} \\ \bar{\tau}\bar{\kappa} \end{pmatrix},$$

para $\gamma = 1$.

$$(y^{k+1}, x^{k+1}, \tau^{k+1}, \theta^{k+1}, s^{k+1}, \kappa^{k+1}) := (\bar{y} + \bar{d}y, \bar{x} + \bar{d}x, \bar{\tau} + \bar{d}\tau, \bar{\theta} + \bar{d}\theta, \bar{s} + \bar{d}s, \bar{\kappa} + \bar{d}\kappa).$$

$$\text{Atualiza parâmetro: } \mu^{k+1} := \frac{(x^{k+1})^T s^{k+1} + \tau^{k+1} \kappa^{k+1}}{n+1}.$$

$$k := k + 1.$$

ATÉ QUE “critério de parada”.

YTM usam o critério de parada descrito em Ye [66], e Mehrotra e Ye [31], para o algoritmo preditor-corretor.

Note que tanto as direções quanto o tamanho do passo dependem de τ, θ, κ , além dos usuais y, x, s . Esta dinâmica por assim dizer, dificulta estabelecermos uma relação definitiva para a redução da complexidade em número de iterações dos demais métodos de ponto-interior-inviável.

Uma solução ótima $(y^*, x^*, \tau^*, \theta^*, s^*, \kappa^*) \in \mathcal{F}_h$ tal que

$$\begin{pmatrix} x^* + s^* \\ \tau^* + \kappa^* \end{pmatrix} > 0$$

é denominada solução auto-complementar estrita.

YTM demonstram que se o Algoritmo 3.3.10 pára pelo critério de parada com uma solução auto-complementar estrita $(y^*, x^*, \tau^* = 0, \theta^* = 0, s^*, \kappa^* > 0)$ para (HLP), então $c^T x^* - b^T y^* < 0$. Isto é, pelo menos um dos $c^T x^*$ e $-b^T y^*$ é estritamente menor do que zero. Se $c^T x^* < 0$ então (D) é inviável; se $-b^T y^* < 0$ então (P) é inviável; e se ambos $c^T x^* < 0$ e $-b^T y^* < 0$ então ambos (P) e (D) são inviáveis. Todavia, este resultado não fecha o problema de detecção de inviabilidade, uma vez que os próprios YTM executam um exemplo em que $c^T x^* > 0$ e $-b^T y^* < 0$, obtendo inviabilidade de ambos os problemas (P) e (D).

Aqui vamos resumir a desempenho do Algoritmo 3.3.10, segundo YTM, nos dois resultados seguintes. Considere os problemas primal (P) e dual (D), conforme capítulo anterior.

Teorema 3.3.11 *Sejam os problemas (P) e (D) enunciados com dados inteiros. Seja L o número de bits dos dados de entrada. Então pela construção de (HLP), os dados permanecem inteiros e seu comprimento é $O(L)$. Portanto, o Algoritmo 3.3.10 gera uma solução auto-complementar estrita para (HLP) em $O(\sqrt{n}L)$ iterações.*

Demonstração. Teorema 6 em Ye, Todd e Mizuno [71]. ■

Teorema 3.3.12 *Em $O(\sqrt{n}L)$ iterações, o Algoritmo 3.3.10 gera soluções ótimas para os problemas (P) e (D) ou indica que os problemas (P) e (D) são inviáveis.*

Demonstração. Corolário 7 em Ye, Todd e Mizuno [71]. ■

Capítulo 4

Um algoritmo conceitual primal

Neste capítulo vamos construir e estudar um algoritmo ‘conceitual’ primal. O termo conceitual refere-se ao uso de um oráculo, o qual fornece o novo ponto central. O algoritmo conceitual primal fornece uma estratégia para manipularmos os parâmetros associados à viabilidade e otimalidade e indica um caminho para o estudo da complexidade em iterações. Isto deve nos preparar para o próximo capítulo, onde especificamos o oráculo e desenvolvemos o algoritmo.

Informamos que a notação para este capítulo e o próximo será completamente independente dos capítulos 2 e 3. Em particular, os sub-índices ‘1’ e ‘2’ são, respectivamente, referidos à viabilidade e à otimalidade. Esta notação advém do contexto de métodos combinando fases 1 (viabilidade) e 2 (otimalidade).

4.1 O problema

Consideremos os números inteiros m e n tais que $n > m > 0$. Dados uma matriz $\tilde{A} \in R^{m \times (n-1)}$ e vetores $b \in R^m$ e $\tilde{c} \in R^{n-1}$, denominamos de problema original o seguinte problema de programação linear primal:

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

Assumimos $\text{rank} \tilde{A} = m$, sem perda de generalidade como já vimos no segundo capítulo.

Segue-se algumas definições associadas ao problema original (\tilde{P}). O conjunto $\tilde{\mathcal{X}} = \{\tilde{x} \in R^{n-1}; \tilde{A}\tilde{x} = b, \tilde{x} \geq 0\}$ é o conjunto viável, o conjunto $\tilde{\mathcal{X}}^0 = \{\tilde{x} \in \tilde{\mathcal{X}}; \tilde{x} > 0\}$ é o conjunto de pontos interiores, o número $v(\tilde{P}) = \inf\{\tilde{c}^T\tilde{x}; \tilde{x} \in \tilde{\mathcal{X}}\}$ é o valor ótimo e o conjunto $\mathcal{X}(\tilde{P}) = \{\tilde{x} \in \tilde{\mathcal{X}}; \tilde{c}^T\tilde{x} = v(\tilde{P})\}$ é o conjunto de soluções ótimas.

A hipótese para este capítulo e o próximo é:

(H) $\mathcal{X}(\tilde{P})$ é não vazio e limitado.

Esta hipótese pode ser enfraquecida, o que não faremos neste trabalho.

Definimos

$$A = [\tilde{A} \quad b - \tilde{A}\tilde{x}^0], \quad c^T = (\tilde{c}^T, 0) \quad \text{e} \quad \xi^T = (0^T, 1)$$

onde $\tilde{x}^0 \geq 0$ é dado.

Observamos que o ponto $x^T = ((\tilde{x}^0)^T, 1)$ satisfaz $Ax = b$ e $x \geq 0$. Além disso, a matriz A preserva o mesmo rank de \tilde{A} , $\xi^T x = x_n$ e $c^T x = \tilde{c}^T \tilde{x}$.

Por construção, um problema equivalente ao problema (\tilde{P}), a menos de uma dimensão, é o seguinte:

$$\begin{aligned} (P_0) \quad & \text{minimizar} && c^T x \\ & \text{sujeito a:} && Ax = b \\ & && \xi^T x = 0 \\ & && x \geq 0. \end{aligned}$$

Denotamos o conjunto viável para o problema (P_0) por

$$\mathcal{X}_0 = \{x \in R^n; Ax = b, \xi^T x = 0, x \geq 0\}.$$

e o conjunto de soluções ótimas por

$$\mathcal{X}(P_0) = \{x \in \mathcal{X}_0; c^T x = v(\tilde{P})\}.$$

Observamos que o conjunto $\mathcal{X}(P_0)$ é não vazio e limitado se, e só se, $\mathcal{X}(\tilde{P})$ é não vazio e limitado.

Vamos resolver o problema (P_0) , logo (\tilde{P}) , considerando dois problemas: dado $\theta_c \in R$, definimos o problema associado à viabilidade para (P_0) por

$$\begin{aligned} (P_1) \quad & \text{minimizar} && \xi^T x \\ & \text{sujeito a:} && Ax = b \\ & && c^T x \leq \theta_c \\ & && x \geq 0. \end{aligned}$$

E, dado $\theta \geq 0$, definimos o problema associado à otimalidade para (P_0) por

$$\begin{aligned} (P_2) \quad & \text{minimizar} && c^T x \\ & \text{sujeito a:} && Ax = b \\ & && \xi^T x \leq \theta \\ & && x \geq 0. \end{aligned}$$

Os problemas duais para (P_1) e (P_2) são, respectivamente,

$$\begin{aligned} (D_1) \quad & \text{maximizar} && b^T y - \theta_c t \\ & \text{sujeito a:} && A^T y - ct + s = \xi \\ & && t \geq 0, s \geq 0, \end{aligned}$$

e

$$\begin{aligned} (D_2) \quad & \text{maximizar} && b^T \omega - \theta l \\ & \text{sujeito a:} && A^T \omega - \xi l + r = c \\ & && l \geq 0, r \geq 0. \end{aligned}$$

Para cada $\theta_c \in R$ e $\theta \geq 0$, denotamos

$$\mathcal{X}_1(\theta_c) = \{x \in R^n; Ax = b, c^T x \leq \theta_c, x \geq 0\},$$

$$\mathcal{X}_1^0(\theta_c) = \{x \in \mathcal{X}_1(\theta_c); c^T x < \theta_c, x > 0\},$$

$$\mathcal{X}_2(\theta) = \{x \in R^n; Ax = b, \xi^T x \leq \theta, x \geq 0\},$$

$$\mathcal{X}_2^0(\theta) = \{x \in \mathcal{X}_2(\theta); \xi^T x < \theta, x > 0\}.$$

Consideremos \tilde{x}^0 o vetor de uns em R^{n-1} dado. Observamos que dado $\theta_c > c^T e$, $x = e \in \mathcal{X}_1^0(\theta_c)$. Por outro lado, dado $\theta > 1$, $x = e \in \mathcal{X}_2^0(\theta)$.

Denotamos a função barreira logarítmica, estritamente convexa, por

$$z > 0 \mapsto p(z) = - \sum_{j=1}^n \ln z_j.$$

Dados $\mu_\xi > 0$ e $\theta_c \in R$, definimos a função penalizada com peso n , estritamente convexa,

$$x \in \mathcal{X}_1^0(\theta_c) \mapsto f_1(x) = \frac{\xi^T x}{\mu_\xi} - n \ln(\theta_c - c^T x) + p(x).$$

Ainda, dados $\mu > 0$ e $\theta > 0$, definimos a função penalizada com peso n , estritamente convexa,

$$x \in \mathcal{X}_2^0(\theta) \mapsto f_2(x) = \frac{c^T x}{\mu} - n \ln(\theta - \xi^T x) + p(x).$$

Os gradientes das funções penalizadas f_1 e f_2 são, respectivamente,

$$x \in \mathcal{X}_1^0(\theta_c) \mapsto \nabla f_1(x) = \frac{\xi}{\mu_\xi} + \frac{n}{\theta_c - c^T x} c - x^{-1}$$

e

$$x \in \mathcal{X}_2^0(\theta) \mapsto \nabla f_2(x) = \frac{c}{\mu} + \frac{n}{\theta - \xi^T x} \xi - x^{-1}.$$

O ponto central $x(\theta_c, \mu_\xi)$ associado a $\mu_\xi > 0$ e $\theta_c \in R$ é

$$x(\theta_c, \mu_\xi) = \operatorname{argmin}\{f_1(x); x \in \mathcal{X}_1^0(\theta_c)\}.$$

O ponto central $x(\theta, \mu)$ associado a $\mu > 0$ e $\theta > 0$ é

$$x(\theta, \mu) = \operatorname{argmin}\{f_2(x); x \in \mathcal{X}_2^0(\theta)\}.$$

Agora vamos introduzir um limitante inferior para o parâmetro θ_c , que será importante para certificar o problema original (\tilde{P}) inviável. Assim, consideremos $\theta_0 \geq 1$ dado. Definimos o problema

$$\begin{aligned} (P_2)_{\theta_0} \quad & \text{minimizar} && c^T x \\ & \text{sujeito a:} && Ax = b \\ & && \xi^T x \leq \theta_0 \\ & && x \geq 0. \end{aligned}$$

Usando o Teorema 2.1.5, uma vez que $e \in \mathcal{X}_2^0(\theta)$ para $\theta = 1$, a hipótese (H) é equivalente ao conjunto de nível

$$\{x \in R_+^n; Ax = b, c^T x = v(\tilde{P}), \xi^T x \leq \bar{\theta}\}$$

limitado para todo $\bar{\theta} \geq 0$. Então, podemos definir o limitante para θ_c ,

$$v(P_2)_{\theta_0} = \min\{c^T x; Ax = b, \xi^T x \leq \theta_0, x \geq 0\}.$$

Para todo $\theta > 0$, o conjunto $\mathcal{X}_2^0(\theta)$ é não vazio. Com efeito, para $\theta = 0$ existe $\hat{x} \in \mathcal{X}_2(\theta)$ e, para $\theta > 1$, $e \in \mathcal{X}_2^0(\theta)$. Pela convexidade de $\mathcal{X}_2(\theta)$ segue-se que existe um segmento que passa pelos pontos e e \hat{x} tal que, para todo $\tau \in [0, 1]$,

$$x(\tau) = \hat{x} + \tau(e - \hat{x}) \in \mathcal{X}_2(\theta),$$

como queríamos demonstrar. Ainda, considere $\theta > 0$ fixo, porém arbitrário. O conjunto de soluções ótimas para o problema (P_2) é não vazio e limitado, usando a hipótese (H) e o Teorema 2.1.5.

Da mesma forma, para todo $\theta_c \in R$, $\theta_c > v(P_2)_{\theta_0}$, o conjunto $\mathcal{X}_1^0(\theta_c)$ é não vazio e o conjunto de soluções ótimas para o problema (P_1) é não vazio e limitado.

Estas duas análises nos colocam nas condições para a existência e unicidade de pontos centrais associados ao problema (P_2), para cada $\theta > 0$ e para cada $\mu > 0$,

e associados ao problema (P_1) , para cada $\theta_c > v(P_2)_{\theta_0}$ e para cada $\mu_\xi > 0$, conforme Proposição 2.2.3.

Nossa idéia para o desenvolvimento que ora apresentamos será discutida agora. Pretendemos resolver o problema (\tilde{P}) . Como vimos a pouco, isto é equivalente a resolver o problema (P_0) , a menos de uma dimensão. Todavia, retiramos a hipótese forte que diz que um ponto interior viável inicial está disponível (A propósito, retiramos a hipótese que diz que $\tilde{\mathcal{X}}^0$ é não vazio e limitado, também.). Desta forma, passamos a ter dois problemas: encontrar um ponto viável e encontrar uma solução ótima, ambos, para o problema (P_0) . Notamos que, ambos, ponto viável e ponto ótimo, são garantidos pela hipótese (H) . Na nossa opinião, a dificuldade para resolver o problema original sem a hipótese forte está no fato de que não sabemos controlar os parâmetros associados aos problemas de viabilidade (P_1) e de otimalidade (P_2) . Na verdade, o método homogêneo e auto-dual não nos ensina este truque! Nossa estratégia, inicialmente, é aprender como controlar estes parâmetros, observando atentamente o problema associado à viabilidade. Para isto, retiramos uma segunda dificuldade que é o cálculo do passo para uma certa medida de proximidade. Faremos isto usando um oráculo que cumpre este papel, ou seja, fornece um passo para um novo ponto central. É assim que construiremos o algoritmo conceitual primal.

A seguir estamos interessados em definir o *gap* de dualidade para ambos os problemas (P_1) e (P_2) .

Lema 4.1.1 (i) $s \in \mathbb{R}^n$ é uma folga dual viável para o problema (D_1) se, e somente se, $s \geq 0$ e $s_p = \xi_p + c_p t$, para algum $t \geq 0$; e

(ii) $r \in \mathbb{R}^n$ é uma folga dual viável para o problema (D_2) se, e somente se, $r \geq 0$ e $r_p = c_p + \xi_p l$, para algum $l \geq 0$.

Demonstração. Vamos demonstrar (i). A demonstração de (ii) é análoga. Considere $s \geq 0$. Então s é uma folga dual viável se, e somente se, para algum $y \in \mathbb{R}^m$ e para algum $t \geq 0$, $\xi + ct - s = A^T y$. Mas $\xi + ct - s$ pode ser decomposto

em um único caminho como $\xi + ct - s = P_A(\xi + ct - s) + A^T y$, e segue-se da decomposição da última igualdade que $P_A(\xi + ct - s) = P_A(\xi + ct) - P_A(s) = 0$, finalizando a demonstração. ■

Lema 4.1.2 (i) *Seja $x(\theta_c, \mu_\xi)$ o ponto central associado a $\mu_\xi > 0$ e $\theta_c > v(P_2)_{\theta_0}$ em (P_1) . Então*

$$s(\theta_c, \mu_\xi) = \mu_\xi x^{-1}(\theta_c, \mu_\xi)$$

é uma folga dual viável,

$$x(\theta_c, \mu_\xi)^T s(\theta_c, \mu_\xi) = n\mu_\xi$$

e a variável dual

$$t(\theta_c, \mu_\xi) = \frac{n\mu_\xi}{\theta_c - c^T x(\theta_c, \mu_\xi)} \geq 0.$$

Além disso, o gap de dualidade associado a $x(\theta_c, \mu_\xi)$ e $s(\theta_c, \mu_\xi)$ é

$$\Delta_1 = 2n\mu_\xi;$$

e

(ii) *Seja $x(\theta, \mu)$ o ponto central associado a $\mu > 0$ e $\theta > 0$ em (P_2) . Então*

$$r(\theta, \mu) = \mu x^{-1}(\theta, \mu)$$

é uma folga dual viável,

$$x(\theta, \mu)^T r(\theta, \mu) = n\mu$$

e a variável dual

$$l(\theta, \mu) = \frac{n\mu}{\theta - \xi^T x(\theta, \mu)} \geq 0.$$

Além disso, o gap de dualidade associado a $x(\theta, \mu)$ e $s(\theta, \mu)$ é

$$\Delta_2 = 2n\mu.$$

Demonstração. Vamos demonstrar (i). A demonstração de (ii) é análoga. Dados $\mu_\xi > 0$ e $\theta_c > v(P_2)_{\theta_0}$, denotamos por conveniência $x = x(\theta_c, \mu_\xi)$, $s = s(\theta_c, \mu_\xi)$ e $t = t(\theta_c, \mu_\xi)$. Sabemos que x é o único ponto tal que $P_A(\nabla f_1(x)) = 0$, ou seja,

$$\frac{\xi_p}{\mu_\xi} + \frac{n}{\theta_c - c^T x} c_p - P_A(x^{-1}) = 0.$$

Daí,

$$P_A(\mu_\xi x^{-1}) = \xi_p + \frac{n\mu_\xi}{\theta_c - c^T x} c_p.$$

Uma vez que $x > 0$,

$$s = \mu_\xi x^{-1}$$

satisfaz as condições necessárias e suficientes para ser uma folga dual viável, conforme Lema 4.1.1. Ainda, usando o Lema 4.1.1, obtemos a variável dual

$$t = \frac{n\mu_\xi}{\theta_c - c^T x} \geq 0.$$

Segue-se, também, que

$$x^T s = x^T (\mu_\xi x^{-1}) = n\mu_\xi.$$

Finalmente, o gap de dualidade é por definição, $\Delta_1 = \xi^T x - (b^T y - \theta_c t)$. Somando em ambos os lados desta igualdade $(c^T x)t$, obtemos $\Delta_1 = x^T s + (\theta_c - c^T x)t$. Usando os resultados já demonstrados aqui, $x^T s = n\mu_\xi$ e $t = n\mu_\xi / (\theta_c - c^T x)$, segue-se que

$$\Delta_1 = 2n\mu_\xi,$$

completando a demonstração. ■

O próximo resultado é a chave para a nossa estratégia. Ele estabelece a ponte ligando os problemas (P_1) e (P_2) para pontos centrais, permitindo-nos relacionar os parâmetros μ_ξ com θ e θ_c com μ .

Lema 4.1.3 *Considere $\bar{x} \in \{x \in R^n; Ax = b, x > 0\}$. Então $\bar{x} = x(\theta_c, \mu_\xi)$ para algum $\mu_\xi > 0$ e para algum $\theta_c > c^T \bar{x}$ se, e somente se, $\bar{x} = x(\theta, \mu)$ para algum $\mu > 0$ e para algum $\theta > \xi^T \bar{x}$. Se uma das afirmações são satisfeitas, então os parâmetros estão assim relacionados:*

$$\mu_\xi = \frac{\theta - \xi^T \bar{x}}{n} \quad e \quad \mu = \frac{\theta_c - c^T \bar{x}}{n}.$$

Além disso, os “gaps” de dualidade podem ser reescritos por

$$\Delta_1 = 2(\theta - \xi^T \bar{x}) \quad e \quad \Delta_2 = 2(\theta_c - c^T \bar{x}).$$

Demonstração. Suponha que $\bar{x} = x(\theta_c, \mu_\xi)$ para $\mu_\xi > 0$ e $\theta_c > c^T \bar{x}$. Isto equivale a

$$\bar{x} = \operatorname{argmin}\{f_1(x); x \in \mathcal{X}_1^0(\theta_c)\}.$$

Pelas condições necessárias e suficientes para a existência e unicidade de \bar{x} ,

$$P_A(\nabla f_1(\bar{x})) = \frac{\xi_p}{\mu_\xi} + \frac{n}{\theta_c - c^T \bar{x}} c_p - P_A(\bar{x}^{-1}) = 0.$$

Tomando θ e μ satisfazendo

$$\mu_\xi = \frac{\theta - \xi^T \bar{x}}{n}$$

e

$$\mu = \frac{\theta_c - c^T \bar{x}}{n},$$

$$P_A(\nabla f_1(\bar{x})) = P_A(\nabla f_2(\bar{x})) = 0.$$

Isto equivale a

$$\bar{x} = \operatorname{argmin}\{f_2(x); x \in \mathcal{X}_2^0(\theta)\},$$

ou equivalentemente, $\bar{x} = x(\theta, \mu)$, concluindo as primeira e segunda partes da demonstração. Agora, usando o Lema 4.1.2, os “gaps” de dualidade são $\Delta_1 = 2n\mu_\xi$ e $\Delta_2 = 2n\mu$. Concluimos a demonstração substituindo μ_ξ e μ já obtidos em Δ_1 e Δ_2 , respectivamente. ■

Uma consequência deste lema é que $\Delta_1/2 \leq \theta$. Com efeito, $\Delta_1 = 2(\theta - x_n)$. Daí, $\Delta_1/2 \leq \theta$, uma vez que $x_n \geq 0$.

4.2 O algoritmo conceitual

Agora estamos prontos para enunciar um algoritmo conceitual de ponto-interior-inviável primal para o problema original (\tilde{P}). Em seguida estudaremos a complexidade.

Iniciamos esta seção enunciando o algoritmo conceitual primal. Observamos que na inicialização deste algoritmo envolve um procedimento de centralização. Além disso, em princípio escolhemos o valor do parâmetro β no intervalo $(0, 1)$.

Algoritmo 4.2.1 *Conceitual.*

Dados: $\theta^0 > 0$, $\mu^0 > 0$, $x^0 = x(\theta^0, \mu^0)$, $\beta \in (0, 1)$ e $L \in \mathbb{N}$.

$k := 0$.

REPITA

$\theta := \theta^k$, $\mu := \mu^k$, $x := x(\theta, \mu)$.

Se $x_n \leq \frac{2}{3}\theta$

Então

$\theta^{k+1} := \beta\theta + (1 - \beta)x_n$, $\mu^{k+1} := \beta\mu$.

Senão

Se $x_n < 2^{-L}$, então *PARE*.

$\theta^{k+1} := \theta$, $\mu^{k+1} := \frac{1}{\beta}\mu$.

Oráculo: $x^{k+1} := x(\theta^{k+1}, \mu^{k+1})$.

$k := k + 1$.

ATÉ QUE $(\theta^k < 2^{-L} \text{ e } \mu^k < 2^{-L})$ ou $\mu^k > 2^L$.

O teste neste algoritmo é crucial. Vamos interpretá-lo agora mesmo. Este é um teste essencialmente para o problema de viabilidade. No nosso caso, para o problema (P_1) . Uma vez que o algoritmo usa um oráculo, estamos nas condições do Lema 4.1.3. Seja $x = x(\theta_c, \mu_\xi) = x(\theta, \mu)$ um ponto central arbitrário dado. Pela hipótese (H) , existe $\theta_c > v(P_2)_{\theta_0}$ tal que o valor da função objetivo de (P_1) é zero, porque $\xi^T x = 0$ em (P_0) . Então, o teste consiste em avaliar o valor da função objetivo do problema dual (D_1) . Assim,

$$\xi^T x - \Delta_1 = x_n - 2(\theta - x_n) \leq 0$$

equivale a $x_n \leq \frac{2}{3}\theta$. Caso contrário, $\xi^T x - \Delta_1 > 0$ equivale a $x_n \in (\frac{2}{3}\theta, \theta)$, já que $x_n < \theta$ sempre.

Agora vamos interpretar a primeira linha no item ‘senão’ do algoritmo conceitual. Na verdade, resolver o problema (P_0) , logo (\tilde{P}) , significa encontrar x tal que $c^T x \leq v(\tilde{P}) + 2^{-L}$ e $\theta < 2^{-L}$ ou certificar que o problema original é inviável. Assim, se $\theta < 2^{-L}$ e estamos no item ‘senão’ do algoritmo, então o problema está resolvido. Vamos formalizar isto demonstrando que se o algoritmo executa o item ‘senão’, para alguma iteração k , então $c^T x^k < \theta_c^k \leq v(\tilde{P})$.

Teorema 4.2.2 *Suponha que a hipótese (H) é satisfeita. Para toda iteração k , $k \geq 0$, do algoritmo conceitual, se $x_n^k > (2/3)\theta^k$, então $\theta_c^k = n\mu^k + c^T x^k \leq v(\tilde{P})$, onde $v(\tilde{P})$ é o valor ótimo do problema original (\tilde{P}) .*

Demonstração. Pela hipótese (H) , seja $v(\tilde{P})$ o valor ótimo para o problema (\tilde{P}) . Consideremos a k -ésima iteração do algoritmo. Novamente pela hipótese (H) , o problema (P_1) possui uma solução ótima \bar{x} para $\theta_c^k = n\mu^k + c^T x^k$, onde μ^k e x^k são gerados pelo algoritmo. E, também, existe uma solução para o problema (P_0)

denotada por \hat{x} . Suponhamos que $x_n^k > (2/3)\theta^k$. Então, usando dualidade e o Lema 4.1.3, existe uma solução dual viável (y, t) tal que

$$b^T y - \theta_c^k t = x_n^k - 2(\theta^k - x_n^k) > 0.$$

Logo, pela dualidade fraca, $0 = \xi^T \hat{x} < \xi^T \bar{x}$. Se $\theta_c^k > v(\tilde{P})$, então \hat{x} é viável para (P_1) e $0 = \xi^T \hat{x}$, contradizendo $0 = \xi^T \hat{x} < \xi^T \bar{x}$. Pela arbitrariedade de $k \geq 0$, finalizamos a demonstração. ■

Agora vamos demonstrar que os iterados $\mu^k > 0$ do algoritmo conceitual são limitados superiormente.

Teorema 4.2.3 *Suponha que a hipótese (H) é satisfeita. Então $\mu^k \leq 2^L$, para toda iteração k do algoritmo conceitual.*

Demonstração. Pela hipótese (H), seja $v(\tilde{P})$ o valor ótimo para o problema (\tilde{P}) . Ainda pela hipótese (H), o problema (P_2) admite uma solução ótima. Assim, seja \hat{v} o valor ótimo, com $\hat{v} \geq v(P_2)_{\theta_0}$. O algoritmo gera pontos centrais através do oráculo e aumenta o valor de μ quando no item ‘senão’. Assim, consideremos a k -ésima iteração do algoritmo no item ‘senão’. Seja o ponto central x^k . Temos $c^T x^k \geq \hat{v} \geq v(P_2)_{\theta_0}$. Segue-se pelo Teorema 4.2.2 que $\theta_c^k = n\mu^k + c^T x^k \leq v(\tilde{P})$, implicando em

$$\mu^k \leq \frac{v(\tilde{P}) - c^T x^k}{n} \leq \frac{v(\tilde{P}) - v(P_2)_{\theta_0}}{n} \leq 2^L.$$

Pela arbitrariedade de $k \geq 0$, finalizamos a demonstração. ■

Observamos que a negação deste teorema afirma que se no algoritmo conceitual $\mu^k > 2^L$, então o algoritmo certifica que a hipótese (H) não é satisfeita.

O restante desta seção destina-se a estudar a complexidade em iterações do algoritmo conceitual. Inicialmente vamos garantir o decréscimo do parâmetro θ quando no item ‘então’ do algoritmo.

Lema 4.2.4 *Considere uma iteração k do algoritmo conceitual, onde o item ‘então’ é executado. Então:*

$$(i) \theta^k \leq \frac{3}{2}\Delta_1^k; e$$

$$(ii) \theta^{k+1} \leq \frac{2+\beta}{3}\theta^k = (1 - \frac{\nu}{3})\theta^k, \quad \nu > 0.$$

Demonstração. Pelo Lema 4.1.3, $\Delta_1^k = 2(\theta^k - x_n^k)$. Daí,

$$\theta^k = \frac{\Delta_1^k}{2} + x_n^k \leq \frac{\Delta_1^k}{2} + \frac{2}{3}\theta^k,$$

pela entrada no item ‘então’. Isto implica em $\theta^k \leq \frac{3}{2}\Delta_1^k$, finalizando a primeira parte da demonstração. Finalmente, vamos demonstrar a parte (ii). No item ‘então’ temos

$$\theta^{k+1} = \beta\theta^k + (1 - \beta)x_n^k \leq \beta\theta^k + \frac{2}{3}(1 - \beta)\theta^k = \frac{2 + \beta}{3}\theta^k,$$

concluindo a demonstração para $\beta = 1 - \nu$, $\nu > 0$. ■

O próximo resultado garante viabilidade para uma certa tolerância para o iterado θ^k e uma certa estimativa para o contador k , no algoritmo conceitual.

Teorema 4.2.5 *Suponha que a hipótese (H) é satisfeita. Assuma que o algoritmo conceitual inicia com $\theta^0 = 1$ e $\mu^0 = 1$. Considere a iteração k do algoritmo e $\beta = 1 - \nu$, $\beta \in (0, 1)$. Se $k > (7/\nu)L$, então $\theta^k < 2^{-L}$.*

Demonstração. Consideremos uma iteração k do algoritmo. Tomemos $k = i + j$, onde i é o número de vezes que o algoritmo executa o item ‘então’ e j é o número relativo à execução do item ‘senão’. Suponhamos que $\theta^k \geq 2^{-L}$. Pelo Lema 4.2.4, pelo item ‘senão’ e usando recorrência,

$$\frac{\theta^k}{\theta^0} \leq \left(\frac{2 + \beta}{3}\right)^i.$$

Tomando o logaritmo na base 2 e usando $\beta = 1 - \nu$,

$$\log\left(\frac{\theta^k}{\theta^0}\right) \leq i \log\left(1 - \frac{\nu}{3}\right) \leq -i\frac{\nu}{3},$$

onde a última desigualdade decorre da propriedade $\log(1 - a) \leq -a$ para todo $a \in (0, 1)$. Pela hipótese de inicialização de θ^0 e pela nossa suposição de que $\theta^k \geq 2^{-L}$,

$$i \leq \frac{3}{\nu} \log\left(\frac{\theta^0}{\theta^k}\right) \leq \frac{3}{\nu} L.$$

Agora vamos considerar μ^k . Temos que

$$\mu^k = \left(\frac{1}{\beta}\right)^{j-i} \mu^0.$$

Pelo Teorema 4.2.3, $\mu^k \leq 2^L$ e pela hipótese de inicialização de μ^0 , $\mu^0 = 1$. Então, tomando o logaritmo,

$$j - i \leq \frac{L}{\log\left(\frac{1}{\beta}\right)} \leq \frac{L}{\nu},$$

onde esta última desigualdade decorre de $\beta = 1 - \nu$ e da mesma propriedade de logaritmo aplicada anteriormente. Segue-se que

$$k = i + j \leq 2i + \frac{L}{\nu} \leq \frac{7}{\nu} L.$$

Isto finaliza a demonstração. ■

O próximo teorema é o resultado de complexidade em iterações para o algoritmo conceitual.

Teorema 4.2.6 *Suponha que a hipótese (H) é satisfeita. Assuma que o algoritmo conceitual inicia com $\theta^0 = 1$ e $\mu^0 = 1$. Considere $\beta = 1 - \nu$, $\beta \in (0, 1)$. Então o algoritmo pára em $O\left(\frac{1}{\nu}L\right)$ passos.*

Demonstração. Consideremos uma iteração k do algoritmo. Suponhamos que $k > k_1$, onde $k_1 = \frac{7}{\nu}L$. Então, pelo Teorema 4.2.5, $\theta^k < 2^{-L}$ e, usando o algoritmo, o item ‘senão’ não ocorre mais. Logo, basta-nos considerar o item ‘então’ avaliado para μ^{k_2} tal que

$$\mu^{k_1+k_2} = \mu^k = \beta^{k_2} \mu^{k_1}.$$

Então, tomando o logaritmo, usando $\beta = 1 - \nu$ e a propriedade $\log(1 - a) \leq -a$ para todo $a \in (0, 1)$,

$$\log\left(\frac{\mu^{k_1+k_2}}{\mu^{k_1}}\right) = k_2 \log(1 - \nu) \leq -\nu k_2.$$

Daí,

$$k_2 \leq \frac{1}{\nu} \log\left(\frac{\mu^{k_1}}{\mu^{k_1+k_2}}\right) \leq \frac{2}{\nu}L,$$

onde a última desigualdade decorre do fato de que o algoritmo ainda não parou, isto é, $\mu^{k_1+k_2} \geq 2^{-L}$ e $\mu^{k_1} < 2^L$. Portanto, $k = k_1 + k_2 \leq \frac{2}{\nu}L$. Isto finaliza a demonstração. ■

Finalmente, o algoritmo conceitual tem a complexidade $O(\sqrt{n}L)$ iterações para a escolha de $\beta = 1 - \nu$, onde $\nu = 0.1/\sqrt{n}$. No próximo capítulo faremos esta escolha para β .

Capítulo 5

Um algoritmo de ponto-interior-inviável

Este capítulo conclui o nosso trabalho. Aqui desenvolvemos um algoritmo de ponto-interior-inviável que resolve o problema de programação linear em $O(\sqrt{n}L)$ iterações para passos curtos com uma abordagem primal-dual, tal que permitimos inviabilidade primal, mas mantemos viabilidade dual.

Como comentamos no capítulo 1, na inicialização deste algoritmo é necessário um procedimento de centralização, que é feito em geral nos algoritmos de pontos interiores (viáveis). Esse procedimento de centralização tem complexidade $O(\sqrt{n}L)$ iterações. No nosso caso, a idéia dessa inicialização é a seguinte: dado um ponto $x > 0$ para um problema primal, construímos um novo problema com uma dimensão a mais contendo o conjunto viável. Daí, sob a hipótese (H) que garante a existência de ponto central para esse procedimento de centralização, calculamos o ponto inicial (aproximadamente) central. Usando dualidade, obtemos o ponto (aproximadamente) central dual inicial, tal que ele é dual viável pela construção.

Por outro lado, algoritmos de passos curtos não são práticos, mas como estudamos no capítulo 2, eles são esclarecedores para o estudo de complexidade.

Consideremos o par de problemas

$$\begin{aligned}
(P_0) \quad & \text{minimizar} && c^T x \\
& \text{sujeito a:} && Ax = b \\
& && \xi^T x = 0 \\
& && x \geq 0
\end{aligned}$$

e o seu dual

$$\begin{aligned}
(D_0) \quad & \text{maximizar} && b^T y \\
& \text{sujeito a:} && A^T y - \xi t + s = c \\
& && s \geq 0.
\end{aligned}$$

O nosso objetivo é resolvê-lo com a escolha pretendida para o parâmetro $\beta = 1 - \nu$, onde $\nu = 0.1/\sqrt{n}$, conforme afirmamos no final do capítulo 4.

5.1 Sobre a notação

Reescrevemos os problemas (P_1) e (P_2) do capítulo anterior assim: dados $\theta_c > v(P_2)_{\theta_0}$ e $\theta > 0$,

$$\begin{aligned}
(P_1) \quad & \text{minimizar} && \xi^T \check{x} \\
& \text{sujeito a:} && A\check{x} = b \\
& && c^T \check{x} + \check{x}_{n+1} = \theta_c \\
& && \check{x} \geq 0, \check{x}_{n+1} \geq 0
\end{aligned}$$

e

$$\begin{aligned}
(P_2) \quad & \text{minimizar} && c^T x \\
& \text{sujeito a:} && Ax = b \\
& && \xi^T x + x_{n+1} = \theta \\
& && x \geq 0, x_{n+1} \geq 0.
\end{aligned}$$

E, os seus respectivos problemas duais,

$$\begin{aligned}
(D_1) \quad & \text{maximizar} \quad b^T \check{y} - \theta_c \check{t} \\
& \text{sujeito a:} \quad A^T \check{y} - c \check{t} + \check{s} \quad = \xi \\
& \quad \quad \quad -\check{t} + \check{s}_{n+1} \quad = 0 \\
& \quad \quad \quad \check{s} \geq 0, \check{s}_{n+1} \geq 0
\end{aligned}$$

e

$$\begin{aligned}
(D_2) \quad & \text{maximizar} \quad b^T y - \theta t \\
& \text{sujeito a:} \quad A^T y - \xi t + s \quad = c \\
& \quad \quad \quad -t + s_{n+1} \quad = 0 \\
& \quad \quad \quad s \geq 0, s_{n+1} \geq 0.
\end{aligned}$$

Para $\theta_c > v(P_2)_{\theta_0}$, denotamos

$$\mathcal{X}_{\theta_c} = \{(\check{x}, \check{x}_{n+1}) \in R_+^n \times R_+; A\check{x} = b, c^T \check{x} + \check{x}_{n+1} = \theta_c\},$$

$$\mathcal{X}_{\theta_c}^0 = \{(\check{x}, \check{x}_{n+1}) \in \mathcal{X}_{\theta_c}; (\check{x}, \check{x}_{n+1}) > 0\},$$

$$\mathcal{S}_{\theta_c} = \{(\check{y}, \check{s}, \check{s}_{n+1}) \in R^m \times R_+^n \times R_+; A^T \check{y} - c \check{t} + \check{s} = \xi, \check{t} = \check{s}_{n+1}\},$$

$$\mathcal{S}_{\theta_c}^0 = \{(\check{y}, \check{s}, \check{s}_{n+1}) \in \mathcal{S}_{\theta_c}; (\check{s}, \check{s}_{n+1}) > 0\},$$

$$\hat{\mathcal{S}}_{\theta_c} = \{(\check{s}, \check{s}_{n+1}); (\check{y}, \check{s}, \check{s}_{n+1}) \in \mathcal{S}_{\theta_c} \text{ para algum } \check{y} \in R^m\},$$

$$\hat{\mathcal{S}}_{\theta_c}^0 = \{(\check{s}, \check{s}_{n+1}) \in \hat{\mathcal{S}}_{\theta_c}; (\check{s}, \check{s}_{n+1}) > 0\},$$

$$\mathcal{F}_{\theta_c} = \{(\check{x}, \check{x}_{n+1}, \check{y}, \check{s}, \check{s}_{n+1}) \in \mathcal{X}_{\theta_c} \times \mathcal{S}_{\theta_c}\},$$

$$\mathcal{F}_{\theta_c}^0 = \{(\check{x}, \check{x}_{n+1}, \check{y}, \check{s}, \check{s}_{n+1}) \in \mathcal{X}_{\theta_c}^0 \times \mathcal{S}_{\theta_c}^0\},$$

$$\hat{\mathcal{F}}_{\theta_c} = \{(\check{x}, \check{x}_{n+1}, \check{s}, \check{s}_{n+1}) \in \mathcal{X}_{\theta_c} \times \hat{\mathcal{S}}_{\theta_c}\},$$

$$\hat{\mathcal{F}}_{\theta_c}^0 = \{(\check{x}, \check{x}_{n+1}, \check{s}, \check{s}_{n+1}) \in \mathcal{X}_{\theta_c}^0 \times \hat{\mathcal{S}}_{\theta_c}^0\}$$

e para $\theta > 0$, denotamos

$$\mathcal{X}_\theta = \{(x, x_{n+1}) \in R_+^n \times R_+; Ax = b, \xi^T x + x_{n+1} = \theta\},$$

$$\mathcal{X}_\theta^0 = \{(x, x_{n+1}) \in \mathcal{X}_\theta; (x, x_{n+1}) > 0\},$$

$$\mathcal{S}_\theta = \{(y, s, s_{n+1}) \in R^m \times R_+^n \times R_+; A^T y - \xi t + s = c, t = s_{n+1}\},$$

$$\mathcal{S}_\theta^0 = \{(y, s, s_{n+1}) \in \mathcal{S}_\theta; (s, s_{n+1}) > 0\},$$

$$\hat{\mathcal{S}}_\theta = \{(s, s_{n+1}); (y, s, s_{n+1}) \in \mathcal{S}_\theta \text{ para algum } y \in R^m\},$$

$$\hat{\mathcal{S}}_\theta^0 = \{(s, s_{n+1}) \in \hat{\mathcal{S}}_\theta; (s, s_{n+1}) > 0\},$$

$$\mathcal{F}_\theta = \{(x, x_{n+1}, y, s, s_{n+1}) \in \mathcal{X}_\theta \times \mathcal{S}_\theta\},$$

$$\mathcal{F}_\theta^0 = \{(x, x_{n+1}, y, s, s_{n+1}) \in \mathcal{X}_\theta^0 \times \mathcal{S}_\theta^0\},$$

$$\hat{\mathcal{F}}_\theta = \{(x, x_{n+1}, s, s_{n+1}) \in \mathcal{X}_\theta \times \hat{\mathcal{S}}_\theta\},$$

$$\hat{\mathcal{F}}_\theta^0 = \{(x, x_{n+1}, s, s_{n+1}) \in \mathcal{X}_\theta^0 \times \hat{\mathcal{S}}_\theta^0\}.$$

5.2 Os problemas primais-duais

O par de problemas (P_1) e (D_1) é o problema primal-dual associado ao problema de viabilidade primal-dual, enquanto o par de problemas (P_2) e (D_2) é o problema primal-dual associado ao problema de otimalidade primal-dual. A seguir vamos buscar uma formulação que envolve a resolução de sistemas de igualdades e desigualdades para estes pares de problemas.

Dados $\theta_c \in R$ e $\mu_\xi > 0$ definimos a função penalizada, com peso n ,

$$(\check{x}, \check{x}_{n+1}) \in \mathcal{X}_{\theta_c}^0 \mapsto f_1(\check{x}, \check{x}_{n+1}) = \frac{\xi^T \check{x}}{\mu_\xi} - n \ln(\check{x}_{n+1}) + p(\check{x}).$$

Seu gradiente é definido por

$$(\check{x}, \check{x}_{n+1}) \in \mathcal{X}_{\theta_c}^0 \mapsto \nabla f_1(\check{x}, \check{x}_{n+1}) = \frac{(\xi^T, 0)^T}{\mu_\xi} - \begin{pmatrix} \check{x}^{-1} \\ \frac{n}{\check{x}_{n+1}} \end{pmatrix},$$

enquanto sua matriz hessiana, definida positiva, é definida por

$$(\check{x}, \check{x}_{n+1}) \in \mathcal{X}_{\theta_c}^0 \mapsto \nabla^2 f_1(\check{x}, \check{x}_{n+1}) = \begin{pmatrix} \check{X}^{-2} & 0 \\ 0^T & \frac{n}{\check{x}_{n+1}^2} \end{pmatrix},$$

onde $\check{X} = \text{diag}(\check{x})$.

O ponto central associado a $\theta_c \in R$ e $\mu_\xi > 0$ é definido por

$$x(\theta_c, \mu_\xi) = \operatorname{argmin}\{f_1(\check{x}, \check{x}_{n+1}); (\check{x}, \check{x}_{n+1}) \in \mathcal{X}_{\theta_c}^0\}.$$

Para cada $\theta_c > v(P_2)_{\theta_0}$ e $\mu_\xi > 0$, $x(\theta_c, \mu_\xi)$ existe e é único pelo Teorema 2.1.5 e pela Proposição 2.2.3.

Considere $\theta_c > v(P_2)_{\theta_0}$ e $\mu_\xi > 0$. As condições de otimalidade para o problema que define o ponto central $(\check{x}, \check{x}_{n+1}) \in \mathcal{X}_{\theta_c}^0$ são as seguintes: existem $\bar{y} \in R^m$ e $\bar{t} \in R$ tais que

$$\nabla f_1(\check{x}, \check{x}_{n+1}) - \begin{pmatrix} A^T \\ 0^T \end{pmatrix} \bar{y} - \begin{pmatrix} c \\ 1 \end{pmatrix} \bar{t} = 0.$$

Definindo

$$\check{s} = \mu_\xi \check{x}^{-1} > 0 \text{ e } \check{s}_{n+1} = \frac{n\mu_\xi}{\check{x}_{n+1}} > 0,$$

$\check{y} = \mu_\xi \bar{y}$ e $\check{t} = -\mu_\xi \bar{t}$, obtemos o sistema primal-dual perturbado

$$(PD)_1 \quad \begin{aligned} A\check{x} &= b \\ c^T \check{x} + \check{x}_{n+1} &= \theta_c \\ A^T \check{y} - c\check{t} + \check{s} &= \xi \\ -\check{t} + \check{s}_{n+1} &= 0 \\ \check{x}\check{s} &= \mu_\xi e \\ \check{x}_{n+1}\check{s}_{n+1} &= n\mu_\xi \\ \check{x}, \check{s} &\geq 0, \check{x}_{n+1}, \check{s}_{n+1} \geq 0. \end{aligned}$$

Observamos que o *gap* de dualidade associado a $(\check{x}, \check{x}_{n+1})$ e $(\check{y}, \check{t}, \check{s}, \check{s}_{n+1})$ satisfazendo o sistema primal-dual perturbado $(PD)_1$ é definido por

$$\Delta_1 = \xi^T \check{x} - (b^T \check{y} - \theta_c \check{t}) = \check{x}^T \check{s} + \check{x}_{n+1} \check{s}_{n+1} = 2n\mu_\xi.$$

Dados $\theta > 0$ e $\mu > 0$ definimos a função penalizada, com peso n ,

$$(x, x_{n+1}) \in \mathcal{X}_\theta^0 \mapsto f_2(x, x_{n+1}) = \frac{c^T x}{\mu} - n \ln(x_{n+1}) + p(x).$$

Seu gradiente é definido por

$$(x, x_{n+1}) \in \mathcal{X}_\theta^0 \mapsto \nabla f_2(x, x_{n+1}) = \frac{(c^T, 0)^T}{\mu} - \begin{pmatrix} x^{-1} \\ \frac{n}{x_{n+1}} \end{pmatrix},$$

enquanto sua matriz hessiana, definida positiva, é definida por

$$(x, x_{n+1}) \in \mathcal{X}_\theta^0 \mapsto \nabla^2 f_2(x, x_{n+1}) = \begin{pmatrix} X^{-2} & 0 \\ 0^T & \frac{n}{x_{n+1}^2} \end{pmatrix},$$

onde $X = \text{diag}(x)$.

O ponto central associado a $\theta > 0$ e $\mu > 0$ é definido por

$$x(\theta, \mu) = \operatorname{argmin}\{f_2(x, x_{n+1}); (x, x_{n+1}) \in \mathcal{X}_\theta^0\}.$$

Para cada $\theta > 0$ e $\mu > 0$, $x(\theta, \mu)$ existe e é único pelo Teorema 2.1.5 e pela Proposição 2.2.3.

Considere $\theta > 0$ e $\mu > 0$. As condições de otimalidade para o problema que define o ponto central $(x, x_{n+1}) \in \mathcal{X}_\theta^0$ são as seguintes: existem $\tilde{y} \in R^m$ e $\tilde{t} \in R$ tais que

$$\nabla f_2(x, x_{n+1}) - \begin{pmatrix} A^T \\ 0^T \end{pmatrix} \tilde{y} - \begin{pmatrix} \xi \\ 1 \end{pmatrix} \tilde{t} = 0.$$

Definindo

$$s = \mu x^{-1} > 0 \text{ e } s_{n+1} = \frac{n\mu}{x_{n+1}} > 0,$$

$y = \mu\tilde{y}$ e $t = -\mu\tilde{t}$, obtemos o sistema primal-dual perturbado

$$\begin{aligned} Ax &= b \\ \xi^T x + x_{n+1} &= \theta \\ A^T y - \xi t + s &= c \\ (PD)_2 \quad -t + s_{n+1} &= 0 \\ xs &= \mu e \\ x_{n+1} s_{n+1} &= n\mu \\ x, s \geq 0, x_{n+1}, s_{n+1} &\geq 0. \end{aligned}$$

Observamos que o *gap* de dualidade associado a (x, x_{n+1}) e (y, t, s, s_{n+1}) satisfazendo o sistema primal-dual perturbado $(PD)_2$ é definido por

$$\Delta_2 = c^T x - (b^T y - \theta t) = x^T s + x_{n+1} s_{n+1} = 2n\mu.$$

5.3 As medidas de proximidade

Aqui formalizamos a teoria de proximidade com a finalidade de desenvolver o oráculo do algoritmo conceitual do capítulo anterior.

Dados $\theta_c > v(P_2)_{\theta_0}$ e $\mu_\xi > 0$ e considerando o sistema $(PD)_1$, gostaríamos de encontrar, em princípio, $(\check{x}, \check{x}_{n+1}, \check{s}, \check{s}_{n+1})$ tal que $\check{x}\check{s}/\mu_\xi = e$ e $\check{x}_{n+1}\check{s}_{n+1}/\mu_\xi = n$. Assim, definimos a medida de proximidade,

$$(\check{x}, \check{x}_{n+1}, \check{s}, \check{s}_{n+1}, \mu_\xi) \in \hat{\mathcal{F}}_{\theta_c}^0 \times R_{++} \mapsto \delta_{\theta_c}(\check{x}, \check{x}_{n+1}, \check{s}, \check{s}_{n+1}, \mu_\xi) = \delta_{\theta_c},$$

onde

$$\delta_{\theta_c} = \left\| \frac{1}{\mu_\xi} \begin{pmatrix} \check{x}\check{s} \\ \frac{\check{x}_{n+1}\check{s}_{n+1}}{\sqrt{n}} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\|.$$

Da mesma forma, dados $\theta > 0$ e $\mu > 0$ e considerando o sistema $(PD)_2$, gostaríamos de encontrar, em princípio, (x, x_{n+1}, s, s_{n+1}) tal que $xs/\mu = e$ e $x_{n+1}s_{n+1}/\mu = n$. Assim, definimos a medida de proximidade,

$$(x, x_{n+1}, s, s_{n+1}, \mu) \in \hat{\mathcal{F}}_{\theta}^0 \times R_{++} \mapsto \delta_{\theta}(x, x_{n+1}, s, s_{n+1}, \mu) = \delta_{\theta},$$

onde

$$\delta_{\theta} = \left\| \frac{1}{\mu} \begin{pmatrix} xs \\ \frac{x_{n+1}s_{n+1}}{\sqrt{n}} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\|.$$

Consideremos $\alpha \in (0, 1)$. Dados $\theta_c > v(P_2)_{\theta_0}$ e $\mu_\xi > 0$, dizemos que o ponto $(\check{x}, \check{x}_{n+1}, \check{s}, \check{s}_{n+1}) \in \hat{\mathcal{F}}_{\theta_c}^0$ é aproximadamente central ou α -central se, e somente se, $\delta_{\theta_c} \leq \alpha$. Por outro lado, dados $\theta > 0$ e $\mu > 0$, dizemos que o ponto $(x, x_{n+1}, s, s_{n+1}) \in \hat{\mathcal{F}}_{\theta}^0$ é aproximadamente central ou α -central se, e somente se, $\delta_{\theta} \leq \alpha$.

A proposição a seguir mostra que se as medidas de proximidade são pequenas, então os “gaps” de complementaridade estão bem relacionados com μ_ξ para pontos aproximadamente centrais em $\hat{\mathcal{F}}_{\theta_c}^0$ e com μ para pontos aproximadamente centrais em $\hat{\mathcal{F}}_{\theta}^0$.

Proposição 5.3.1 Considere $\alpha \in (0, 1)$.

(i) Se $\delta_{\theta_c}(\check{x}, \check{x}_{n+1}, \check{s}, \check{s}_{n+1}, \mu_\xi) \leq \alpha$, então

$$\check{x}^T \check{s} + \check{x}_{n+1} \check{s}_{n+1} \leq 2(n + \sqrt{n}\alpha)\mu_\xi; \text{ e}$$

(ii) Se $\delta_\theta(x, x_{n+1}, s, s_{n+1}, \mu) \leq \alpha$, então

$$x^T s + x_{n+1} s_{n+1} \leq 2(n + \sqrt{n}\alpha)\mu.$$

Demonstração. Vamos demonstrar (ii). A demonstração de (i) é análoga. Assumimos que $\delta_\theta = \delta_\theta(x, x_{n+1}, s, s_{n+1}, \mu) \leq \alpha$. Então,

$$\frac{1}{\mu} \begin{pmatrix} xs \\ \frac{x_{n+1}s_{n+1}}{\sqrt{n}} \end{pmatrix} = \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} + \begin{pmatrix} z \\ z_0 \end{pmatrix},$$

onde $\|(z^T, z_0)^T\| = \delta_\theta$. Multiplicando a última coordenada na igualdade de vetores acima por \sqrt{n} , pré-multiplicando por $(e^T, 1)$,

$$x^T s + x_{n+1} s_{n+1} \leq (2n + e^T z + \sqrt{n} z_0)\mu.$$

Usando a desigualdade de Cauchy-Schwarz e $\|(z^T, z_0)^T\| = \delta_\theta \leq \alpha$,

$$x^T s + x_{n+1} s_{n+1} \leq 2(n + \sqrt{n}\alpha)\mu,$$

concluindo a demonstração. ■

A propósito, lembramos da afirmação de Hiriart-Urruty e Lemaréchal [19], página 394, quanto à escrita do nome Schwarz: *no "t" please*.

Consideremos $\alpha = 0.25$. Uma consequência desta proposição que usaremos adiante é a seguinte:

$$\Delta_1 \leq 3n\mu_\xi.$$

Isto é imediato usando a proposição e o valor de α .

Agora vamos estabelecer a ponte entre os pares de problemas (P_1) , (D_1) e (P_2) , (D_2) para pontos aproximadamente centrais.

Lema 5.3.2 *Considere $\alpha = 0.25$. Dados $\theta > 0$ e $\mu > 0$, seja $(x, x_{n+1}, s, s_{n+1}) \in \hat{\mathcal{F}}_\theta^0$ um ponto aproximadamente central. Então, $(\check{x}, \check{x}_{n+1}, \check{s}, \check{s}_{n+1}) \in \hat{\mathcal{F}}_{\theta_c}^0$ é aproximadamente central com $\delta_{\theta_c} \leq 2\alpha$ para*

$$\theta_c = c^T x + n\mu, \quad \mu_\xi = \frac{x_{n+1}}{n},$$

$$\check{x} = x, \quad \check{x}_{n+1} = n\mu,$$

$$\check{s} = \frac{s}{s_{n+1}} \quad \text{e} \quad \check{s}_{n+1} = \frac{1}{s_{n+1}}.$$

Demonstração. Denotamos

$$\delta_\theta = \delta_\theta(x, x_{n+1}, s, s_{n+1}, \mu)$$

e

$$\delta_{\theta_c} = \delta_{\theta_c}(\check{x}, \check{x}_{n+1}, \check{s}, \check{s}_{n+1}, \mu_\xi).$$

Supomos que

$$\delta_\theta = \left\| \frac{1}{\mu} \begin{pmatrix} xs \\ \frac{x_{n+1}s_{n+1}}{\sqrt{n}} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\| \leq \alpha.$$

Vamos demonstrar que $\delta_{\theta_c} \leq 2\alpha$. Com efeito,

$$\delta_{\theta_c} = \left\| \frac{1}{\mu_\xi} \begin{pmatrix} \check{x}\check{s} \\ \frac{\check{x}_{n+1}\check{s}_{n+1}}{\sqrt{n}} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\|$$

e substituindo

$$\mu_\xi = x_{n+1}/n, \quad \check{x} = x, \quad \check{x}_{n+1} = n\mu, \quad \check{s} = s/s_{n+1} \quad \text{e} \quad \check{s}_{n+1} = 1/s_{n+1}$$

vem,

$$\delta_{\theta_c} = \left\| \frac{n\mu}{x_{n+1}s_{n+1}} \begin{pmatrix} \frac{xs}{\mu} \\ \sqrt{n} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\|.$$

Temos

$$\frac{1}{\mu} \begin{pmatrix} xs \\ \frac{x_{n+1}s_{n+1}}{\sqrt{n}} \end{pmatrix} = \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} - \begin{pmatrix} z \\ z_0 \end{pmatrix}.$$

onde $\|(z^T, z_0)^T\| = \delta_\theta \leq \alpha$. Assim, $x_{n+1}s_{n+1}/\sqrt{n}\mu = \sqrt{n} - z_0$ implicando em

$$\frac{n\mu}{x_{n+1}s_{n+1}} = \frac{\sqrt{n}}{\sqrt{n} - z_0}.$$

Tomando

$$\frac{\sqrt{n}}{\sqrt{n} - z_0} = 1 + \gamma, \quad \gamma = \frac{z_0}{\sqrt{n} - z_0} \in R.$$

Para z_0 pequeno, isto é, $|z_0| \leq \alpha$, obtemos

$$\gamma \leq \frac{z_0}{\sqrt{n}} \frac{\sqrt{n}}{\sqrt{n} - \alpha} \leq \frac{1.15z_0}{\sqrt{n}}.$$

Então, substituindo

$$\frac{n\mu}{x_{n+1}s_{n+1}} = 1 + \gamma, \quad \frac{xs}{\mu} = e - z \text{ e } \gamma = \frac{1.15z_0}{\sqrt{n}}$$

(Pegamos o pior caso para γ !) em δ_{θ_c} , obtemos

$$\begin{aligned} \delta_{\theta_c} &= \left\| (1 + \gamma) \begin{pmatrix} e - z \\ \sqrt{n} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\| \\ &= \left\| \begin{pmatrix} -z + \gamma(e - z) \\ \gamma\sqrt{n} \end{pmatrix} \right\| \\ &\leq \left\| \begin{pmatrix} -z \\ z_0 \end{pmatrix} \right\| + \left\| \begin{pmatrix} 0 \\ 0.15z_0 \end{pmatrix} \right\| + 1.15 \frac{|z_0|}{\sqrt{n}} \left\| \begin{pmatrix} e - z \\ 0 \end{pmatrix} \right\| \\ &\leq \alpha + 0.15\alpha + 1.15 \frac{\alpha}{\sqrt{n}} (\sqrt{n} + \alpha) \\ &\leq 2\alpha, \end{aligned}$$

onde as primeira e segunda desigualdades são devido às desigualdades triangular e $\|(z^T, z_0)^T\| \leq \alpha$. Isto finaliza a demonstração. ■

Observamos neste lema que definindo

$$\check{y} = \frac{y}{s_{n+1}} \text{ e } \check{t} = \frac{1}{s_{n+1}},$$

obtemos

$$\xi = A^T \check{y} - c \check{t} + \check{s} = A^T \frac{y}{s_{n+1}} - \frac{c}{s_{n+1}} + \frac{s}{s_{n+1}}$$

implicando em

$$A^T y - \xi t + s = c,$$

uma vez que $s_{n+1} = t$ se verifica trivialmente a partir de $\check{t} = \check{s}_{n+1}$. Isto significa que, a partir do par de problemas (P_2) e (D_2) , podemos usar o par de problemas (P_1) e (D_1) para os nossos objetivos como fizemos quando do capítulo anterior para pontos centrais. Notemos que não temos disponível os pontos centrais, mas sabemos o que significa tê-los. Assim, tentaremos fazer um bom uso desta maravilha.

Agora vamos verificar o teste que é crucial para o algoritmo conceitual, isto é, ‘se’

$$\check{x}_n - \Delta_1 = b^T \check{y} - \theta_c \check{t} \leq 0.$$

Sabemos que para pontos aproximadamente centrais, $\check{x}_n = x_n$, $\check{y} = \frac{y}{s_{n+1}}$ e $\check{t} = \frac{1}{s_{n+1}}$. Daí,

$$b^T y \leq \theta_c$$

é o nosso novo teste considerando pontos aproximadamente centrais.

5.4 O passo de Newton

Consideremos o problema $(PD)_2$. Dados $\theta > 0$, $(x, x_{n+1}, s, s_{n+1}) \in \hat{\mathcal{F}}_\theta^0$ e $\mu > 0$, idealmente gostaríamos de encontrar

$$x^+ = x + u, \quad x_{n+1}^+ = x_{n+1} + u_{n+1},$$

$$y^+ = y + w, \quad t^+ = t + \Delta t, \quad s^+ = s + v \text{ e } s_{n+1}^+ = s_{n+1} + v_{n+1},$$

tais que $(x^+, x_{n+1}^+, s^+, s_{n+1}^+) \in \hat{\mathcal{F}}_\theta^0$, $x^+ s^+ = \mu e$ e $x_{n+1}^+ s_{n+1}^+ = n\mu$.

O passo de Newton resolve isto aproximadamente linearizando o sistema primal-dual perturbado $(PD)_2$, da seguinte maneira:

$$(N_2) \quad \begin{array}{rcl} Au & = & 0 \\ \xi^T u + u_{n+1} & = & 0 \\ A^T w + v - \xi v_{n+1} & = & 0 \\ su & + xv & = -xs + \mu e \\ s_{n+1} u_{n+1} & + x_{n+1} v_{n+1} & = -x_{n+1} s_{n+1} + n\mu, \end{array}$$

onde $\Delta t = v_{n+1}$.

Como afirmamos no capítulo 2, essencialmente dizemos que a importância da disciplina pontos interiores é a garantia de eficiência do passo de Newton para alguma medida de proximidade. Neste sentido e, munidos do Lema 2.3.3, porque temos aqui $u^T v + u_{n+1} v_{n+1} = 0$, segue-se o próximo resultado, cuja demonstração leva o mesmo raciocínio do Teorema 4.2 em Gonzaga [18].

Teorema 5.4.1 *Considere $\nu = 0.1/\sqrt{n}$. Dados $\theta > 0$, $\mu > 0$ e $(x, x_{n+1}, s, s_{n+1}) \in \hat{\mathcal{F}}_\theta^0$ tais que $\delta_\theta(x, x_{n+1}, s, s_{n+1}, \mu) = \delta_\theta < 1$ e seja $(x^+, x_{n+1}^+, s^+, s_{n+1}^+)$ o resultado do passo de Newton a partir de (x, x_{n+1}, s, s_{n+1}) . Então,*

$$\delta_\theta(x^+, x_{n+1}^+, s^+, s_{n+1}^+, \mu) \leq \frac{\delta_\theta^2}{\sqrt{8}(1 - \delta_\theta)}$$

e $(x^+, x_{n+1}^+, s^+, s_{n+1}^+) \in \hat{\mathcal{F}}_\theta^0$ para $\delta_\theta \leq 0.7$.

Demonstração. Iniciamos fazendo uma mudança de escala para obtermos uma forma mais simples para as equações de Newton. Seja $D \in R^{(n+1) \times (n+1)}$ uma matriz diagonal e definida positiva e vamos fazer uma mudança de variáveis nas equações de Newton colocando

$$\begin{pmatrix} x \\ x_{n+1} \end{pmatrix} = D \begin{pmatrix} \bar{x} \\ \bar{x}_{n+1} \end{pmatrix}, \quad \begin{pmatrix} u \\ u_{n+1} \end{pmatrix} = D \begin{pmatrix} \bar{u} \\ \bar{u}_{n+1} \end{pmatrix},$$

$$\begin{pmatrix} s \\ s_{n+1} \end{pmatrix} = D^{-1} \begin{pmatrix} \bar{s} \\ \bar{s}_{n+1} \end{pmatrix}, \quad \begin{pmatrix} v \\ v_{n+1} \end{pmatrix} = D^{-1} \begin{pmatrix} \bar{v} \\ \bar{v}_{n+1} \end{pmatrix}.$$

Notemos que

$$\begin{pmatrix} \bar{x}\bar{s} \\ \bar{x}_{n+1}\bar{s}_{n+1} \end{pmatrix} = \begin{pmatrix} xs \\ x_{n+1}s_{n+1} \end{pmatrix} \text{ e } \begin{pmatrix} \bar{u}\bar{v} \\ \bar{u}_{n+1}\bar{v}_{n+1} \end{pmatrix} = \begin{pmatrix} uv \\ u_{n+1}v_{n+1} \end{pmatrix}.$$

O sistema (N_2) torna-se

$$\begin{pmatrix} \bar{u} \\ \bar{u}_{n+1} \end{pmatrix} \in \mathcal{N} \left(\begin{bmatrix} A \\ \xi^T \end{bmatrix} D \right), \quad \begin{pmatrix} \bar{v} \\ \bar{v}_{n+1} \end{pmatrix} \in \mathcal{R} \left(D \begin{bmatrix} A \\ \xi^T \end{bmatrix}^T \right)$$

$$\bar{s}\bar{u} + \bar{x}\bar{v} = -\bar{x}\bar{s} + \mu e,$$

$$\bar{s}_{n+1}\bar{u}_{n+1} + \bar{x}_{n+1}\bar{v}_{n+1} = -\bar{x}_{n+1}\bar{s}_{n+1} + n\mu.$$

Escolhemos

$$D = \begin{pmatrix} \sqrt{\frac{x}{s}} & 0 \\ 0 & \sqrt{\frac{x_{n+1}}{s_{n+1}}} \end{pmatrix}.$$

Segue-se que

$$\begin{aligned}
\begin{pmatrix} \bar{x} \\ \bar{x}_{n+1} \end{pmatrix} &= D^{-1} \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} \\
&= \begin{pmatrix} \sqrt{x s} \\ \sqrt{x_{n+1} s_{n+1}} \end{pmatrix} \\
&= D \begin{pmatrix} s \\ s_{n+1} \end{pmatrix} \\
&= \begin{pmatrix} \bar{s} \\ \bar{s}_{n+1} \end{pmatrix},
\end{aligned}$$

e as duas últimas igualdades no sistema (N_2) podem ser escritas como

$$\bar{u} + \bar{v} = -\frac{\mu}{\bar{x}} \left(\frac{\bar{x}\bar{s}}{\mu} - e \right) \quad (5.1)$$

$$\bar{u}_{n+1} + \bar{v}_{n+1} = -\frac{\sqrt{n}\mu}{\bar{x}_{n+1}} \left(\frac{\bar{x}_{n+1}\bar{s}_{n+1}}{\sqrt{n}\mu} - \sqrt{n} \right). \quad (5.2)$$

Sabemos que

$$\left\| \frac{1}{\mu} \begin{pmatrix} \bar{x}\bar{x} \\ \frac{\bar{x}_{n+1}\bar{x}_{n+1}}{\sqrt{n}} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\| = \left\| \frac{1}{\mu} \begin{pmatrix} x s \\ \frac{x_{n+1}s_{n+1}}{\sqrt{n}} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\| = \delta_\theta$$

e, daí, para cada componente $i = 1, \dots, n$,

$$\frac{\mu}{\bar{x}_i^2} \leq \frac{1}{1 - \delta_\theta}$$

e, para $i = n + 1$,

$$\frac{\sqrt{n}\mu}{\bar{x}_{n+1}^2} \leq \frac{1}{\sqrt{n} - \delta_\theta}$$

Usando (5.1), (5.2) e desigualdade de normas e, em seguida, usando a definição de $\|\cdot\|_\infty$, estas duas últimas desigualdades e o fato de que $\sqrt{n}/(\sqrt{n} - \delta_\theta) \leq 1/(1 - \delta_\theta)$,

$$\left\| \begin{pmatrix} \bar{u} \\ \bar{u}_{n+1} \end{pmatrix} + \begin{pmatrix} \bar{v} \\ \bar{v}_{n+1} \end{pmatrix} \right\|^2 \leq \left\| \begin{pmatrix} \frac{\mu}{\bar{x}} \\ \frac{\sqrt{n}\mu}{\bar{x}_{n+1}} \end{pmatrix} \right\|_\infty^2 \left\| \frac{1}{\mu} \begin{pmatrix} \bar{x}\bar{s} \\ \frac{\bar{x}_{n+1}\bar{s}_{n+1}}{\sqrt{n}} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\|^2$$

$$\leq \mu \frac{\delta_\theta^2}{1 - \delta_\theta}. \quad (5.3)$$

Para $\lambda \in [0, 1]$ e usando (N_2) , temos que

$$\begin{pmatrix} (x + \lambda u)(s + \lambda v) \\ (x_{n+1} + \lambda u_{n+1})(s_{n+1} + \lambda v_{n+1}) \end{pmatrix} = \begin{pmatrix} (1 - \lambda)xs + \lambda\mu e + \lambda^2 uv \\ (1 - \lambda)x_{n+1}s_{n+1} + \lambda n\mu + \lambda^2 u_{n+1}v_{n+1} \end{pmatrix}.$$

Segue-se que

$$\begin{aligned} & \frac{1}{\mu} \begin{pmatrix} (x + \lambda u)(s + \lambda v) \\ \frac{(x_{n+1} + \lambda u_{n+1})(s_{n+1} + \lambda v_{n+1})}{\sqrt{n}} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} = \\ (1 - \lambda) & \left[\frac{1}{\mu} \begin{pmatrix} xs \\ \frac{x_{n+1}s_{n+1}}{\sqrt{n}} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right] + \lambda^2 \frac{1}{\mu} \begin{pmatrix} uv \\ \frac{u_{n+1}v_{n+1}}{\sqrt{n}} \end{pmatrix}. \end{aligned}$$

Em particular, para $\lambda = 1$,

$$\begin{aligned} \delta_\theta(x^+, x_{n+1}^+, s^+, s_{n+1}^+, \mu) &= \left\| \frac{1}{\mu} \begin{pmatrix} x^+ s^+ \\ \frac{x_{n+1}^+ s_{n+1}^+}{\sqrt{n}} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\| \\ &= \left\| \frac{1}{\mu} \begin{pmatrix} uv \\ \frac{u_{n+1}v_{n+1}}{\sqrt{n}} \end{pmatrix} \right\| \\ &\leq \left\| \frac{1}{\mu} \begin{pmatrix} uv \\ u_{n+1}v_{n+1} \end{pmatrix} \right\|. \end{aligned} \quad (5.4)$$

Uma vez que $u^T v + u_{n+1}v_{n+1} = \bar{u}^T \bar{v} + \bar{u}_{n+1}\bar{v}_{n+1} = 0$ podemos usar o Lema 2.3.3, obtendo

$$\begin{aligned} \left\| \begin{pmatrix} uv \\ u_{n+1}v_{n+1} \end{pmatrix} \right\| &= \left\| \begin{pmatrix} \bar{u}\bar{v} \\ \bar{u}_{n+1}\bar{v}_{n+1} \end{pmatrix} \right\| \\ &\leq \frac{\left\| \begin{pmatrix} \bar{u} + \bar{v} \\ \bar{u}_{n+1} + \bar{v}_{n+1} \end{pmatrix} \right\|^2}{\sqrt{8}}. \end{aligned}$$

Concluimos com (5.4), esta última desigualdade e (5.3) que

$$\delta_\theta(x^+, x_{n+1}^+, s^+, s_{n+1}^+, \mu) \leq \frac{\delta_\theta^2}{\sqrt{8}(1 - \delta_\theta)}.$$

Isto finaliza a primeira parte. A demonstração da segunda parte é análoga àquela do Teorema 2.3.4, item (c). Isto completa a demonstração. ■

5.5 O algoritmo de passos curtos

Neste ponto estamos prontos para enunciar o algoritmo de ponto-interior-inviável primal-dual para passos curtos com viabilidade dual e demonstrar os resultados acerca da sua complexidade em iterações. Enfatizamos que na inicialização deste algoritmo necessita-se de um procedimento de centralização, conforme comentamos no início deste capítulo. E, também, relembramos que verificamos o teste para este algoritmo no final da seção 5.3. Finalmente, escolhemos $\beta = 1 - \nu$, $\nu = 0.1/\sqrt{n}$.

Algoritmo 5.5.1 Passos curtos.

Dados: $\alpha = 0.25$, $\theta^0 > 0$, $\mu^0 > 0$ e $(x^0, x_{n+1}^0, y^0, s^0, s_{n+1}^0) \in \mathcal{F}_{\theta^0}^0$ tais que

$$\delta_{\theta^0}(x^0, x_{n+1}^0, s^0, s_{n+1}^0, \mu^0) \leq \alpha, \nu = \frac{0.1}{\sqrt{n}} \text{ e } L \in \mathbb{N}.$$

$k := 0$.

REPITA

$$(x, x_{n+1}, y, s, s_{n+1}) := (x^k, x_{n+1}^k, y^k, s^k, s_{n+1}^k), (\theta, \mu) := (\theta^k, \mu^k).$$

$$\theta_c := c^T x + n\mu.$$

$$\text{Se } b^T y \leq \theta_c$$

Então

$$\theta^{k+1} := (1 - \nu)\theta + \nu x_n, \mu^{k+1} := (1 - \nu)\mu.$$

$$x_{n+1} := \theta^{k+1} - x_n.$$

Senão

$$\text{Se } x_n < 2^{-L}, \text{ então PARE.}$$

$$\theta^{k+1} := \theta, \mu^{k+1} := \frac{1}{1-\nu}\mu.$$

Passo de Newton:

Calcule $(u, u_{n+1}, w, v, v_{n+1})$ por (N_2) com $\mu = \mu^{k+1}$.

$$(x^{k+1}, x_{n+1}^{k+1}, y^{k+1}, s^{k+1}, s_{n+1}^{k+1}) := (x, x_{n+1}, y, s, s_{n+1}) + (u, u_{n+1}, w, v, v_{n+1}).$$

$$k := k + 1.$$

ATÉ QUE $(\theta^k < 2^{-L} \text{ e } \mu^k < 2^{-L})$ ou $\mu^k > 2^L$.

Uma observação interessante neste algoritmo é que quando da atribuição para o parâmetro θ_c , o ponto

$$(\check{x}, \check{x}_{n+1}, \check{y}, \check{t}, \check{s}, \check{s}_{n+1}) = (x, n\mu, \frac{y}{s_{n+1}}, \frac{1}{s_{n+1}}, \frac{s}{s_{n+1}}, \frac{1}{s_{n+1}})$$

para $\mu_\xi = x_{n+1}/n$, é tal que $(\check{x}, \check{x}_{n+1}, \check{s}, \check{s}_{n+1})$ é um ponto aproximadamente central para o problema (P_1) , conforme Lema 5.3.2.

O item ‘então’ do algoritmo toma

$$\theta^+ = x_n + (1 - \nu)x_{n+1}$$

e

$$\begin{pmatrix} \hat{x} \\ \hat{x}_{n+1} \end{pmatrix} = \begin{pmatrix} x \\ \theta^+ - x_n \end{pmatrix} = \begin{pmatrix} x \\ (1 - \nu)x_{n+1} \end{pmatrix},$$

obtendo viabilidade, isto é,

$$\begin{bmatrix} \check{A} & 0 \\ \check{\xi}^T & 1 \end{bmatrix} \begin{pmatrix} \hat{x} \\ \hat{x}_{n+1} \end{pmatrix} = \begin{pmatrix} b \\ \theta^+ \end{pmatrix}.$$

Isto justifica o cálculo do passo de Newton por (N_2) , o que não acontece com os métodos de ponto-interior-inviável usuais, conforme capítulo 3.

A interpretação para a primeira linha no item ‘senão’ do algoritmo de passos curtos é a mesma quando do algoritmo conceitual. Isto é, resolver o problema (P_0) , logo (\tilde{P}) , significa encontrar x tal que $c^T x \leq v(\tilde{P}) + 2^{-L}$ e $\theta < 2^{-L}$ ou certificar que o problema original é inviável. Assim, se $\theta < 2^{-L}$ e estamos no item ‘senão’ do algoritmo, então o problema original (primal) está resolvido. Vamos formalizar isto agora.

Teorema 5.5.2 *Suponha que a hipótese (H) é satisfeita. Para toda iteração k , $k \geq 0$, do algoritmo de passos curtos, se $b^T y^k > \theta_c$, então $\theta_c = n\mu^k + c^T x^k \leq v(\tilde{P})$, onde $v(\tilde{P})$ é o valor ótimo do problema original (\tilde{P}).*

Demonstração. Basta substituir na demonstração do Teorema 4.2.2 a desigualdade $x_n^k > \frac{2}{3}\theta^k$ por $b^T y^k > \theta_c$. Desta forma, não necessitamos do Lema 4.1.3 para pontos centrais. ■

Agora vamos demonstrar que os iterados $\mu^k > 0$ do algoritmo de passos curtos são limitados superiormente.

Teorema 5.5.3 *Suponha que a hipótese (H) é satisfeita. Então $\mu^k \leq 2^L$, para toda iteração k do algoritmo de passos curtos.*

Demonstração. Análoga à demonstração do Teorema 4.2.3, observando que os pontos centrais para o uso do Teorema 4.2.2 podem ser substituídos aqui pelos pontos aproximadamente centrais gerados pelo algoritmo de passos curtos, garantidos pelos Lemas 5.3.2 e 5.5.2. ■

Da mesma forma como no algoritmo conceitual, observamos que a negação deste teorema afirma que se no algoritmo de passos curtos $\mu^k > 2^L$, então o algoritmo certifica que a hipótese (H) não é satisfeita.

O restante desta seção destina-se a demonstrar a complexidade em iterações do algoritmo de passos curtos. Inicialmente vamos garantir o decréscimo do parâmetro θ quando no item ‘então’ do Algoritmo 5.5.1. Uma curiosidade é comparar o item (ii) desse próximo lema com o mesmo item no Lema 4.2.4.

Lema 5.5.4 *Considere uma iteração k do algoritmo de passos curtos, onde o item ‘então’ é executado. Então:*

$$(i) \theta^k \leq 4x_{n+1}^k; \text{ e}$$

$$(ii) \theta^{k+1} \leq (1 - \frac{\nu}{4})\theta^k.$$

Demonstração. Consideremos a k -ésima iteração do algoritmo tal que o item ‘então’ é executado, onde omitiremos o índice k . Temos que

$$\theta = x_n + x_{n+1} \leq \Delta_1 + x_{n+1} \leq 3n\mu_\xi + x_{n+1} = 4x_{n+1},$$

onde as igualdades e desigualdades seguem-se, respectivamente, pelas viabilidade primal, entrada no item ‘então’ do algoritmo, Proposição 5.3.1 e definição de μ_ξ . Isto conclui a parte (i). Finalmente vamos demonstrar a parte (ii). No item ‘então’ temos,

$$\theta^{k+1} = \theta - \nu(\theta - x_n) = \theta - \nu x_{n+1} = \theta - \nu n\mu_\xi,$$

uma vez que $n\mu_\xi = x_{n+1}$. Pelo item(i), $\theta \leq 4x_{n+1} = 4n\mu_\xi$ implicando em $\theta/4 \leq n\mu_\xi$. Então,

$$\theta^{k+1} = \theta - \nu n\mu_\xi \leq \theta - \frac{\nu\theta}{4} = \theta\left(1 - \frac{\nu}{4}\right).$$

Isto conclui a demonstração. ■

O próximo resultado é fundamental para a convergência e para a complexidade em iterações do algoritmo de passos curtos.

Lema 5.5.5 *Para toda iteração k , $k \geq 0$, do algoritmo de passos curtos,*

$$\delta_{\theta^k}(x^k, x_{n+1}^k, s^k, s_{n+1}^k, \mu^k) = \delta_{\theta^k} \leq \alpha \text{ e } (x^k, x_{n+1}^k, s^k, s_{n+1}^k) \in \hat{\mathcal{F}}_{\theta^k}^0.$$

Demonstração. Por indução. O resultado é verdadeiro para $k = 0$. Assumimos que para algum $k = 0, 1, \dots$, $(x^k, x_{n+1}^k, s^k, s_{n+1}^k) \in \hat{\mathcal{F}}_{\theta^k}^0$ e $\delta_{\theta^k} \leq \alpha$. Devemos mostrar que nesta iteração do algoritmo de passos curtos, $(x^+, x_{n+1}^+, s^+, s_{n+1}^+) \in \hat{\mathcal{F}}_{\theta^{k+1}}^0$ com $\delta_{\theta^{k+1}}(x^+, x_{n+1}^+, s^+, s_{n+1}^+, \mu^{k+1}) = \delta_{\theta^{k+1}} \leq \alpha$. Do Teorema 5.4.1, com $\theta = \theta^k$ e após o passo de Newton com $\mu = \mu^k$, $(x^+, x_{n+1}^+, s^+, s_{n+1}^+) \in \hat{\mathcal{F}}_{\theta^k}^0$ e

$$\begin{aligned} \delta_{\theta^k}(x^+, x_{n+1}^+, s^+, s_{n+1}^+, \mu) &\leq \frac{\delta_{\theta^k}^2}{\sqrt{8(1-\delta_{\theta^k})}} \\ &\leq \frac{\alpha^2}{\sqrt{8(1-\alpha)}} \\ &< 0.03. \end{aligned}$$

Resta mostrar que

$$\delta_{\theta}^{k+1} = \left\| \frac{1}{\mu^{k+1}} \begin{pmatrix} x^+ s^+ \\ (1-\nu) \frac{x_{n+1}^+ s_{n+1}^+}{\sqrt{n}} \end{pmatrix} - (1-\nu) \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\| \leq \alpha.$$

Inicialmente vamos considerar o item ‘então’ do algoritmo. Assim, temos que

$$\begin{aligned} & \left\| \frac{1}{\mu} \begin{pmatrix} x^+ s^+ \\ (1-\nu) \frac{x_{n+1}^+ s_{n+1}^+}{\sqrt{n}} \end{pmatrix} - (1-\nu) \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\| \\ & \leq \left\| \frac{1}{\mu} \begin{pmatrix} x^+ s^+ \\ \frac{x_{n+1}^+ s_{n+1}^+}{\sqrt{n}} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\| + \nu \left\| \frac{1}{\mu} \begin{pmatrix} 0 \\ \frac{x_{n+1}^+ s_{n+1}^+}{\sqrt{n}} \end{pmatrix} - \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\| \\ & \leq \delta_{\theta}(x^+, x_{n+1}^+, s^+, s_{n+1}^+, \mu) + \nu \left\| \begin{pmatrix} e \\ 1 \end{pmatrix} \right\| \\ & \leq 0.03 + \frac{0.1}{\sqrt{n}} \sqrt{n+1} \\ & \leq 0.18 \\ & \leq 0.25(1-\nu), \end{aligned}$$

onde a primeira desigualdade é devido à desigualdade triangular, as segunda e terceira pelo Teorema 5.4.1 e $1-\nu = 1 - \frac{0.1}{\sqrt{n}} \geq 0.9$ justifica a última. Então, dividindo por $(1-\nu)$, $\delta_{\theta^{k+1}} \leq 0.25$. Finalmente, vamos considerar o item ‘senão’ do algoritmo. Assim, considerando

$$\frac{1}{\nu} = 1 + \gamma, \quad \gamma = \frac{\nu}{1-\nu},$$

e usando o mesmo raciocínio quando da demonstração considerando o item ‘então’, segue-se que

$$\begin{aligned}
& \left\| \frac{1}{\mu} \begin{pmatrix} x^+ s^+ \\ \frac{x_{n+1}^+ s_{n+1}^+}{\sqrt{n}} \end{pmatrix} - \frac{1}{(1-\nu)} \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\| \\
&= \left\| \frac{1}{\mu} \begin{pmatrix} x^+ s^+ \\ \frac{x_{n+1}^+ s_{n+1}^+}{\sqrt{n}} \end{pmatrix} - (1+\gamma) \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\| \\
&\leq \delta_\theta(x^+, x_{n+1}^+, s^+, s_{n+1}^+, \mu) + |\gamma| \left\| \begin{pmatrix} e \\ \sqrt{n} \end{pmatrix} \right\| \\
&\leq \frac{0.03}{1-\nu} + \frac{\nu}{1-\nu} \sqrt{2n} \\
&\leq \frac{0.25}{1-\nu}.
\end{aligned}$$

Então, multiplicando por $(1 - \nu)$, $\delta_{\theta^{k+1}} \leq 0.25$. Isto completa a demonstração. ■

O próximo resultado garante viabilidade para uma certa tolerância para o iterado θ^k e uma certa estimativa para o contador k , no algoritmo de passos curtos.

Teorema 5.5.6 *Suponha que a hipótese (H) é satisfeita. Assuma que o algoritmo de passos curtos inicia com $\theta^0 = 1$ e $\mu^0 = 1$. Considere a iteração k do algoritmo e $\beta = 1 - \nu$, $\beta \in (0, 1)$. Se $k > (9/\nu)L$, então $\theta^k < 2^{-L}$.*

Demonstração. Consideremos uma iteração k do algoritmo. Tomemos $k = i + j$, onde i é o número de vezes que o algoritmo executa o item ‘então’ e j é o número relativo à execução do item ‘senão’. Suponhamos que $\theta^k \geq 2^{-L}$. Pelo Lema 5.5.4, pelo item ‘senão’ e usando recorrência,

$$\frac{\theta^k}{\theta^0} \leq \left(1 - \frac{\nu}{4}\right)^i.$$

Por outro lado,

$$\frac{\mu^k}{\mu^0} = (1 - \nu)^{j-i}.$$

Pelo Teorema 5.5.3, $\mu^k \leq 2^L$. A demonstração se segue análoga à do Teorema 4.2.5.

■

Finalmente, o próximo teorema é o resultado da complexidade em iterações para o nosso algoritmo para aquela (lembra!) escolha de $\beta = 1 - \nu$, $\nu = 0.1/\sqrt{n}$.

Teorema 5.5.7 *O algoritmo de passos curtos pára em, no máximo, $O(\sqrt{n}L)$ iterações.*

Demonstração. Análoga à demonstração do Teorema 4.2.6, substituindo o uso do Teorema 4.2.5 pelo Teorema 5.5.6. ■

Referências Bibliográficas

- [1] I. Adler e R. C. Monteiro *Limiting of the affine scaling continuous trajectories for linear programming problems*. Contemporary Mathematics, Vol. 114, 1990.
- [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*. Mathematical Programming 52 (1991) 429-439.
- [4] K. M. Anstreicher *Linear programming in $O((n^3/\ln(n))L)$ operations*. Para aparecer em SIAM Journal on Optimization.
- [5] D. Bayer e J. C. Lagarias *The nonlinear geometry of linear programming, I. Affine and projective scaling trajectories, II. Legendre transform coordinates and central trajectories*. Trans. Amer. Math. Soc. 314 (1989), 499-581.
- [6] S. C. Billups e M. C. Ferris *Convergence of an infeasible interior-point algorithm from arbitrary positive starting points*. SIAM Journal on Optimization Vol. 6, N. 2, pp. 316-325, 1996.
- [7] J. F. Bonnans e F. A. Potra *On the convergence of the iteration sequence of infeasible path following algorithms for linear complementarity problems*. Mathematics of Operations Research, Vol. 22, N. 2, 1997.
- [8] L. M. G. Drummond *Classical and generalized central paths with algorithmic applications in linear programming*. Tese de doutorado pelo IMPA/CNPQ, defendida sob a orientação do professor Alfredo N. Iusem em 1997.

- [9] A. Fiacco e G. McCormick *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley, New York, 1968.
- [10] R. M. Freund *Polynomial-time algorithms for linear programming based only on primal scaling and projected gradients of a potential function*. *Mathematical Programming* 51 (1991) 203-222.
- [11] R. M. Freund *A potential-function reduction algorithm for solving a linear program directly from an infeasible 'warm start'*. *Mathematical Programming* 52 (1991) 441-466.
- [12] K. R. Frisch *The logarithmic potential method of convex programming*. Memorandum, University Institute of Economics, Oslo, Norway, May 1955.
- [13] G. de Ghellinck e J. P. Vial *A polynomial Newton method for linear programming*. *Algorithmica* 1 (1986) 425-453.
- [14] P. E. Gill, W. Murray, M. A. Saunders, J. A. Tomlin e M. H. Wright *On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method*. *Mathematical Programming* 36 (1986) 183-209.
- [15] C. C. Gonzaga *An algorithm for solving linear programming problems in $O(n^3L)$ operations*. Editado por N. Megiddo, *Progress in Mathematical Programming: Interior Point and Related Methods*, Springer Verlag, New York, Cap. 1, 1989.
- [16] C. C. Gonzaga *Algoritmos de pontos interiores para programação linear*. 17^o Colóquio Brasileiro de Matemática. Instituto de Matemática Pura e Aplicada do CNPq, 1989.
- [17] C. C. Gonzaga *Path-following methods for linear programming*. *SIAM Review*, Vol. 34, N. 2, pp. 167-224, 1992.
- [18] C. C. Gonzaga *On the complexity of linear programming*. *Resenhas IME-USP* 1995, Vol. 2, N. 2, 97-207.

- [19] J.-B. Hiriart-Urruty e C. Lemaréchal *Convex Analysis and Minimization Algorithms I*. Springer-Verlag, 2ª edição, 1996.
- [20] P. Huard *Resolution of mathematical programming with nonlinear constraints by the method of centers in Nonlinear Programming*. Editado por J. Abadie, North-Holland, Amsterdam, 1967.
- [21] N. Karmarkar *A new polynomial time algorithm for linear programming*. *Combinatorica* 4 (1984) 373-395.
- [22] L. G. Khachiyan *A polynomial algorithm for linear programming*. *Soviet Mathematics Doklady*, 20:191-194, 1979.
- [23] M. Kojima *Basic lemmas in polynomial-time infeasible-interior-point methods for linear programs*. *Annals of Operations Research* 62 (1996) 1-28.
- [24] M. Kojima, S. Mizuno e A. Yoshise *A little theorem of the big \mathcal{M} in interior point algorithms*. *Mathematical Programming* 59 (1993) 361-375.
- [25] M. Kojima, N. Megiddo e S. Mizuno *A primal-dual infeasible-interior-point for linear programming*. *Mathematical Programming* 61 (1993) 263-280.
- [26] I. J. Lustig, R. E. Marsten e D. F. Shanno *Computational experience with a primal-dual interior point method for linear programming*. *Linear Algebra and its Applications* 152 (1991) 191-222.
- [27] N. Maculan e M. V. F. Pereira *Programação Linear*. Atlas, 1980.
- [28] R. E. Marsten, R. Subramanian, M. Saltzman, I. J. Lustig, R. E. Marsten e D. F. Shanno *Interior point methods for linear programming: Just call Newton, Lagrange, and Fiacco and McCormick!*. *Interfaces* 20 (1990) 105-116.
- [29] J. M. Martínez e S. A. Santos *Métodos Computacionais de Otimização*. 20º Colóquio Brasileiro de Matemática. Instituto de Matemática Pura e Aplicada do CNPQ, 1995.

- [30] N. Megiddo *Pathways to the optimal set in linear programming*. Editado por N. Megiddo, Progress in Mathematical Programming: Interior Point and Related Methods, Springer Verlag, New York, Cap. 8, 1989.
- [31] S. Mehrotra e Y. Ye *Finding an interior point in the optimal face of linear programs*. Mathematical Programming 62 (1993) 497-515.
- [32] M. A. F. Menezes *Algoritmos de pontos interiores para programação linear combinando fases 1 e fase 2*. Tese de Mestrado pela COPPE/UFRJ, defendida sob a orientação do professor Clóvis C. Gonzaga em 1991.
- [33] J. Miao *Two infeasible-interior-point predictor-corrector algorithms for linear programming*. SIAM Journal on Optimization Vol. 6, N. 3, pp. 587-599, 1996.
- [34] S. Mizuno *A new polynomial time method for linear complementarity problem*. Mathematical Programming 56 (1992) 31-43.
- [35] S. Mizuno *Polynomiality of infeasible-interior-point algorithms for linear programming*. Mathematical Programming 67 (1994) 109-119.
- [36] S. Mizuno, M. J. Todd e Y. Ye *On adaptive-step primal-dual interior-point algorithms for linear programming*. Mathematics of Operations Research Vol. 18, N. 4, 1993.
- [37] S. Mizuno, M. J. Todd e Y. Ye *A surface of analytic centers and primal-dual infeasible-interior-point algorithms for linear programming*. Mathematics of Operations Research Vol. 20, N. 1, 1995.
- [38] S. Mizuno, M. Kojima e M. J. Todd *Infeasible-interior-point primal-dual potential-reduction algorithms for linear programming*. SIAM Journal on Optimization Vol. 5, N. 1, pp. 52-67, 1995.
- [39] S. Mizuno, F. Jarre e J. Stoer *A unified approach to infeasible-interior-point algorithms via geometrical linear complementarity problems*. Applied Mathematics and Optimization 33:315-341 (1996).

- [40] R. D. C. Monteiro e I. Adler *Interior path following primal-dual algorithms. Part I: linear programming* Mathematical Programming 44 (1989) 27-41.
- [41] R. D. C. Monteiro e S. Wright *A superlinear infeasible-interior-point affine scaling algorithm for LCP*. SIAM Journal on Optimization Vol. 6, N. 1, pp. 1-18, 1996.
- [42] Y. Nesterov *An old and new approach to nonlinear programming*. Mathematical Programming 79 (1997) 285-297.
- [43] F. Potra *An infeasible-interior-point predictor-corrector algorithm for linear programming*. SIAM Journal on Optimization Vol. 6, N. 1, pp. 19-32, 1996.
- [44] F. Potra *A quadratically convergent predictor-corrector method for solving linear programs from infeasible starting points*. Mathematical Programming 67 (1994) 383-406.
- [45] F. A. Potra *An $O(nL)$ infeasible-interior-point algorithm for LCP with quadratic convergence*. Report N. 50, Department of Mathematics, Iowa University (Iowa City, Iowa, 1994).
- [46] J. Renegar *A polynomial-time algorithm, based on Newton's method, for linear programming*. Mathematical Programming 40 (1988) 59-93.
- [47] R. Sheng e F. A. Potra *A quadratically convergent infeasible-interior-point algorithm for LCP with polynomial complexity*. SIAM Journal on Optimization Vol. 7, N. 2, pp. 304-317, 1997.
- [48] J. Stoer *The complexity of an infeasible interior-point path-following method for the solution of linear programs*. Optimization Methods and Software, 1994, Vol. 3, pp. 1-12.
- [49] K. Tanabe *Centered Newton method for mathematical programming*. In: M. Iri and K. Yajima, eds., System Modelling and Optimization: Proceedings of the 13th IFIP Conference, Tokyo, Japan, Aug./Sept. 1987, Vol. 113 of Lecture

Notes in Control and Information Sciences (Springer-Verlag, Berlin, 1988) pp. 197-206.

- [50] T. Terlaky (editor) *Interior Point Methods of Mathematical Programming*. Applied Optimization Vol. 5, KLUWER Academic Publishers, 1996.
- [51] T. Terlaky *An easy way to teach interior point methods*. ISSN 0922-5641, Reports of the Faculty of Technical Mathematics and Informatics 98-24, Delft, 1998.
- [52] M. J. Todd *On Anstreicher's combined phase I-phase II projective algorithm for linear programming*. Mathematical Programming 55 (1992) 1-15.
- [53] M. J. Todd *Combined phase I and phase II in a potential reduction algorithm for linear programming*. Mathematical Programming 59 (1993) 133-150.
- [54] M. J. Todd e Y. Wang *A combined phase I-phase II projective methods for linear programming*. Technical Report 877, Cornell University, Ithaca, New York, 1989.
- [55] M. J. Todd *Potential-reduction methods in mathematical programming*. Mathematical Programming 76 (1997) 3-45.
- [56] M. J. Todd e Y. Ye *A centered projective algorithm for linear programming*. Mathematics of Operations Research 15 (1990) 508-529.
- [57] P. Tseng *Simplified analysis of an $O(nL)$ -iteration infeasible predictor-corrector path-following method for monotone linear complementarity problems*. Department of Mathematics, University of Washington, Seattle, Washington, U.S.A..
- [58] R. H. Tütüncü *An infeasible-interior-point potential-reduction algorithm for linear programming*. Para aparecer em Mathematical Programming.
- [59] P. Vaidya *an algorithm for linear programming which requires $O(((m+n)n^2 + (m+n)^{1.5}n)L)$ arithmetic operations*. Mathematical Programming 47 (1990) 175-201.

- [60] X. Xu e Y. Ye *A generalized homogeneous and self-dual linear programming algorithm*. Operations Research Letters, 17:181-190, 1995.
- [61] X. Xu, P.-F. Hung e Y. Ye *A simplified homogeneous and self-dual linear programming algorithm and its implementation*. Annals of Operations Research, 62:151-171, 1996.
- [62] S. Wright *An infeasible-interior-point algorithm for linear complementarity problems*. Mathematical Programming 67 (1994) 29-51.
- [63] S. Wright *A path-following infeasible-interior-point algorithm for linear complementarity problems*. Optimization Methods and Software, 1993, Vol. 2, pp. 79-106.
- [64] S. Wright *A path-following interior-point algorithm for linear and quadratic problems*. Annals of Operations Research 62 (1996) 103-130.
- [65] S. Wright e Y. Zhang *A superquadratic infeasible-interior-point method for linear complementarity problems*. Mathematical Programming 73 (1996) 269-289.
- [66] Y. Ye *On the finite convergence of interior-point algorithms for linear programming*. Mathematical Programming 57 (1992) 325-335.
- [67] Y. Ye *Interior-Point Algorithm: Theory and Practice*, 1995.
- [68] Y. Ye *On homogeneous and self-dual algorithms for LCP*. Mathematical Programming 76 (1997) 211-221.
- [69] Y. Ye, R. A. Tapia e Y. Zhang *A superlinearly convergent $O(\sqrt{n}L)$ -iteration algorithm for linear programming*. TR91-22, Department of Mathematical Sciences, Rice University (Houston, Texas, 1991).
- [70] Y. Ye, O. Güler, R. A. Tapia e Y. Zhang *A quadratically convergent $O(\sqrt{n}L)$ -iteration algorithm for linear programming*. Mathematical Programming 59 (1993) 151-162.

- [71] Y. Ye, M. J. Todd e S. Mizuno *An $O(\sqrt{n}L)$ -iteration homogeneous and self-dual linear programming algorithm*. Mathematics of Operations Research Vol. 19, N. 1, 1994.
- [72] Y. Zhang *On the convergent of a class of infeasible interior-point methods for the horizontal linear complementarity problem*. SIAM Journal on Optimization Vol. 4, N. 1, pp. 208-227, 1994.
- [73] Y. Zhang e D. Zhang *Superlinear convergence of infeasible-interior-point methods for linear programming*. Mathematical Programming 66 (1994) 361-377.