Lilian Ferreira Berti

Iteração Continuada Aplicada ao Método de Pontos Interiores

Campinas
2012

Universidade Estadual de Campinas
Instituto de Matemática, Estatística e
Computação Científica - IMECC
Departamento de Matemática Aplicada

Lilian Ferreira Berti

Iteração Continuada Aplicada ao Método de Pontos Interiores

Dissertação de Mestrado apresentada ao Instituto de Matemática, Estatística e Computação Científica da UNICAMP para obtenção do título de Mestre em Matemática Aplicada.

Orientador: Aurelio Ribeiro Leite de Oliveira
Coorientadora: Carla Taviane Lucke da Silva Ghidini

Este exemplar corresponde à versão final da dissertação defendida por Lilian Ferreira Berti e orientada pelo Prof. Dr. Aurelio Ribeiro Leite de Oliveira.

Prof. Dr. Aurelio Ribeiro Leite de Oliveira

Profa. Dra. Carla Taviane Lucke da Silva Ghidini

Campinas, 2012

FICHA CATALOGRÁFICA ELABORADA POR
MARIA FABIANA BEZERRA MULLER - CRB8/6162
BIBLIOTECA DO INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E
COMPUTAÇÃO CIENTÍFICA - UNICAMP

Berti, Lilian Ferreira, 1988-
B462i Iteração continuada aplicada ao método de pontos interiores / Lilian Ferreira Berti. – Campinas, SP : [s.n.], 2012.

Orientador: Aurelio Ribeiro Leite de Oliveira.
Coorientador: Carla Taviane Lucke da Silva Ghidini.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Métodos de pontos interiores. 2. Programação linear. 3. Métodos iterativos (Matemática). I. Oliveira, Aurelio Ribeiro Leite de, 1962-. II. Ghidini, Carla Taviane Lucke da Silva. III. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. IV. Título.

Informações para Biblioteca Digital

Título em inglês: Continued iteration applied to interior points method
Palavras-chave em inglês:
Interior point methods
Linear programming
Iterative methods (Mathematics)
Área de concentração: Matemática Aplicada
Titulação: Mestre em Matemática Aplicada
Banca examinadora:
Aurelio Ribeiro Leite de Oliveira [Orientador]
Márcia Aparecida Gomes Ruggiero
Geraldo Gil Veiga
Data de defesa: 02-04-2012
Programa de Pós-Graduação: Matemática Aplicada

Dissertação de Mestrado defendida em 02 de abril de 2012 e aprovada

Pela Banca Examinadora composta pelos Profs. Drs.

Prof.(a). Dr(a). AURELIO RIBEIRO LEITE DE OLIVEIRA

Prof.(a). Dr(a). MARCIA APARECIDA GOMES RUGGIERO

Prof.(a). Dr(a). GERALDO GIL VEIGA

À minha família.
Dedico

"A paz é a única forma de nos sentirmos realmente humanos."
(*Albert Einstein*)

Agradecimentos

À Deus, por estar sempre abençoando minha vida e iluminando meu caminho.

Aos meus pais, Sérgio e Luci, e a minha irmã Larissa, que sempre me apoiaram incondicionalmente.

Aos professores Aurelio e Carla pela orientação, pelo apoio, pela paciência, tenho admiração e gratidão por ambos.

Ao meu namorado Yuri, pelo apoio e compreensão.

Aos meus amigos que sempre estiveram presentes nos momentos alegres e difíceis da minha vida, em especial ao meus amigos Felix e Luiz Fernando que tenho uma profunda gratidão, sempre dispostos a ajudar quando necessitei.

À Universidade Estadual de Campinas por propiciar a realização deste trabalho.

À CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pelo apoio financeiro.

Resumo

Os métodos de pontos interiores têm sido amplamente utilizados para determinar a solução de problemas de programação linear de grande porte. O método preditor corretor, dentre todas as variações de métodos de pontos interiores, é um dos que mais se destaca, devido à sua eficiência e convergência rápida. Este método, em cada iteração, necessita resolver dois sistemas lineares para determinar a direção preditora corretora. Resolver estes sistemas lineares corresponde ao passo que requer mais tempo de processamento, devendo assim ser realizada de forma eficiente. Para resolver estes sistemas lineares a abordagem mais utilizada é a fatoração de Cholesky. No entanto, realizar a fatoração de Cholesky em cada iteração tem um alto custo computacional. Dessa forma, na busca de redução de esforços, precisamente, na redução do número de iterações foi desenvolvida a iteração continuada. Iteração continuada é uma iteração subsequente, realizada após o cálculo da direção preditora corretora, onde é determinada uma nova direção sem que seja necessário realizar uma nova fatoração de Cholesky. Os resultados computacionais dos testes realizados, principalmente em problemas de médio e grande porte mostraram que esta abordagem obtém bom desempenho em comparação com o método preditor corretor.

Palavras Chaves: métodos de pontos interiores, programação linear, iteração continuada.

Abstract

Interior point methods have been widely used in the solution of large linear programming problems. The predictor corrector method, among all interior point variants, is one of mostly used due to its efficiency and convergence properties. This method needs the solution of two linear systems to determine the predictor corrector direction, in each iteration. Solving such systems corresponds to the step which requires more processing time. Therefore, it should be done efficiently. The most common approach to solve the linear systems is the Cholesky factorization, demanding in each iteration a high computacional effort. Thus, in search of effort reduction, in particular, to reduce the iterations number continued iteration was developed. The continued iteration is a subsequent iteration performed after the predictor corrector direction is computed, where a new direction is calculated without need to of Cholesky refactorization. The numerical tests show that the continued iteration performs better in comparison with the preditor corretor method.

Keywords: interior point methods, linear programming, continued iteration.

Sumário

Agradecimentos **vi**

Resumo **vii**

Abstract **viii**

Introdução **1**

1 Métodos de Pontos Interiores **4**

1.1 Programação linear 4

1.1.1 Problema de programação linear 4

1.1.2 Problema de programação linear com variáveis canalizadas 5

1.2 Métodos de pontos interiores do tipo primal dual 7

1.2.1 Método de Newton 7

1.2.2 Método primal dual afim escala 8

1.2.3 Método primal dual afim escala com variáveis canalizadas 11

1.2.4 Método preditor corretor 15

1.2.5 Método preditor corretor com variáveis canalizadas 18

2 Método Preditor Corretor com Iteração Continuada **22**

2.1 Função Lagrangiana 23

2.2 Iteração Continuada 24

2.2.1 As componentes i e j bloqueiam 25

2.2.2 Somente a componente i bloqueia 32

2.2.3 Somente a componente j bloqueia 34

2.3 Iteração Continuada com canalização 39

2.3.1 As componentes i e j bloqueiam 39

2.3.2 Somente a componente i bloqueia 45

2.3.3 Somente a componente j bloqueia 45

3 Experimentos Computacionais **49**
3.1 Implementação 49
3.2 Problemas Testes 49
3.3 Resultados Computacionais 51

4 Conclusão e Trabalhos Futuros **57**

A Perfil de Desempenho **59**

Referências Bibliográficas **61**

Introdução

A programação linear é uma técnica de otimização muito utilizada devido a sua robustez, a eficiência de seus algoritmos e ao grande número de problemas que pode ser modelado desta forma. Essa técnica tem sido aplicada em diversas áreas para determinar caminho mínimo, custo mínimo, fluxo máximo, planejamento de operação de sistema, alocação de recursos entre outros [7]. Uma característica positiva da programação linear é a facilidade de se exercer uma análise de sensibilidade na solução encontrada, mostrando a variação no comportamento da solução em função de alterações em qualquer parâmetro do problema.

George Dantzig, em 1947, foi quem modelou a forma primal padrão dos problemas de programação linear, e logo após desenvolveu o algoritmo simplex para resolução deste problema [30]. A chave deste algoritmo está no formato da região limitada pelas restrições lineares que é denominada politopo. Como as restrições lineares formam uma região convexa, a busca pela solução ótima de um problema de programação linear passou a ser limitada a pontos extremos da região do politopo. Assim, no pior caso, todos os pontos extremos são pesquisados para determinar a solução. Permitindo, desta forma, o desenvolvimento por Dantzig de um algoritmo de baixa complexidade computacional por iteração, sendo esta complexidade exponencial na teoria, porém na prática quase linear.

O primeiro algoritmo polinomial para programação linear foi desenvolvido por Khachiyan [21], utilizando elipsóides. O método das elipsóides teve um alto impacto sobre a teoria de programação linear, porém não foi uma alternativa competitiva com o método simplex, devido ao alto custo de operações.

Em 1984, Karmarkar [20] divulga resultados formais sobre um novo método com uma complexidade menor do que o método das elipsóides. Além disso, o método utiliza a noção de função potencial para garantir uma redução da distância para a otimalidade em cada iteração. Embora uma iteração do método de Karmarkar seja cara em comparação com o método simplex, a otimalidade é alcançada em um número relativamente pequeno de iterações em problemas de grande porte. Isso torna o algoritmo computacionalmente atrativo. A prova dos resultados de complexidade no pior caso foi realizada por afirmações de eficiência

sem precedentes para o novo método por Karmarkar. Assim, este método atraiu um alto interesse dos pesquisadores em otimização, que em um tempo depois divulgaram melhorias e esclarecimentos dessa teoria. Gill et al. [14] estabeleceram uma equivalência entre o método de Karmarkar e o método de barreira de Newton projetado. Megiddo [26], divulgou um trabalho em que o método de barreira logarítmica é aplicado nos problemas primal e dual simultaneamente. Kojima, Mizuno e Yoshise [23], apresentam a fundamentação teórica do método de Megiddo e fornecem os primeiros resultados teóricos de complexidade.

Kojima, Megiddo e Mizuno [22] fornecem bons resultados teóricos sobre o algoritmo primal dual, a interpretação de métodos de pontos interiores como algoritmo que segue uma trajetória central para uma solução ótima [19]. Marsten, Subramanian, Saltzman, Lustig e Shanno [25] fazem um progresso expressivo na implementação dos métodos de pontos interiores e fornecem um melhor entendimento da utilização das funções barreiras logarítmas.

Em 1990, Monteiro, Adler e Resende [28] apresentam a estrutura de algoritmo primal dual afim escala, que levou a novos algoritmos com interessantes propostas teóricas. Em seguida, Mehrotra [27] descreve um método prático para programação linear que se tornou base para a maioria das implementações existentes, o método preditor corretor.

O estudo dos métodos de pontos interiores tem sido uma das áreas mais ativas de pesquisa em otimização nas últimas décadas, devido aos grandes avanços ocorridos [18]. Entre as variações do método de pontos interiores para a solução de problemas de programação linear, o método preditor corretor destaca-se entre todos, devido a sua eficiência e rápida convergência. O método preditor corretor é um método do tipo primal dual e necessita resolver dois sistemas lineares com uma mesma matriz simétrica definida positiva em cada iteração. Para resolver estes sistemas a abordagem mais utilizada é a fatoração de Cholesky aplicada ao sistema de equações normais [15]. Dessa forma, o método preditor corretor apresenta uma etapa crítica em seu desempenho computacional: o cálculo da fatoração de Cholesky em cada iteração.

Isto motivou, inspirado no método do gradiente projetado [24], o desenvolvimento de um novo método, o método de pontos interiores preditor corretor com iteração continuada. Este novo método, tem como o objetivo reduzir o número de fatorações de Cholesky calculadas durante a aplicação do método preditor corretor.

A iteração continuada consiste em determinar uma nova direção utilizando a fatoração de Cholesky já existente no método preditor corretor. A proposta da utilização da iteração continuada é apresentada por [29], para o método dual afim escala. Na iteração continuada

desenvolvida neste trabalho a direção é calculada fixando em zero algumas de suas componentes e de forma que a alteração em relação à direção preditora corretora original calculada seja mínima. Esta estratégia permite que a fatoração de Cholesky calculada na iteração do método preditor corretor seja utilizada novamente para resolver os sistemas lineares envolvidos. A abordagem apresentada é incorporada ao método preditor corretor, dando assim origem ao método de pontos interiores preditor corretor com iteração continuada.

No Capítulo 1, apresentamos os métodos de pontos interiores primal dual e preditor corretor para variáveis com e sem canalização. No Capítulo 2, a iteração continuada é desenvolvida para a incorporação no método preditor corretor. No Capítulo 3, os resultados dos experimentos computacionais realizados com diversos problemas selecionados de diferentes coleções são apresentados. No Capítulo 4, apresentamos a conclusão e os trabalhos futuros.

Capítulo 1

Métodos de Pontos Interiores

Neste capítulo, inicialmente, definimos o problema de programação linear com e sem variáriaveis canalizadas, em seguida apresentamos os métodos de pontos interiores primal dual afim escala e preditor corretor.

1.1 Programação linear

Problemas de programação linear (PPL) têm como objetivo determinar uma solução que minimiza ou maximiza uma função sujeita a um conjunto de equações e/ou inequações lineares [5]. Os problemas de programação linear podem apresentar variáveis com canalização ou sem canalização.

1.1.1 Problema de programação linear

O problema de programação linear na forma padrão é dado por:

$$\begin{array}{ll} \min & c^t x \\ \text{s.a} & Ax = b \\ & x \geq 0 \end{array} \tag{1.1}$$

em que $A \in \mathbb{R}^{m \times n}$, posto$(A) = m,\ x \in \mathbb{R}^n,\ c \in \mathbb{R}^n$ e $b \in \mathbb{R}^m$.

O problema (1.1) é denominado problema primal, associado a ele tem-se o problema dual,

cuja forma padrão é dada por:

$$\begin{array}{ll} \max & b^t y \\ \text{s.a} & A^t y + z = c \\ & z \geq 0 \end{array} \tag{1.2}$$

em que $y \in \mathbb{R}^m$ representa o vetor de variáveis duais livres e $z \in \mathbb{R}^n$ representa as variáveis de folga duais.

Quando $x > 0$, $z > 0$ temos que x é um ponto interior no problema primal e z é um ponto interior no problema dual. Além disso, se x satisfaz $Ax = b$, x é um ponto interior factível e se z satisfaz $A^t y + z = c$, z é um ponto interior factível.

Denominamos *gap*, representado por γ, a diferença entre os valores da função objetivo primal e dual do problema, ou seja:

$$\gamma = c^t x - b^t y. \tag{1.3}$$

Observação: Quando um ponto é primal e dual factível, o *gap* é dado por $\gamma = x^t z$.

Teorema 1: Um ponto (x, y, z) é um ponto ótimo para o problema primal e dual se, e somente se, satisfaz as condições de otimalidade de primeira ordem (Karush-Kuhn-Tucker) do problema (1.1) e (1.2) que são dadas por [32]:

$$\begin{array}{lll} \text{(i) Primal factível} & & \\ & Ax - b = 0 \qquad & x \geq 0 \\ \text{(ii) Dual factível} & & \\ & A^t y + z - c = 0 \qquad & z \geq 0 \\ \text{(iii) Complementariedade} & & \\ & XZe = 0 & \end{array} \tag{1.4}$$

sendo $X = diag(x),\ Z = diag(z)$ e $e \in \mathbb{R}^n$, tal que $e = (1, 1, ..., 1)^t$.

1.1.2 Problema de programação linear com variáveis canalizadas

O problema de programação linear com variáveis canalizadas na forma padrão primal é dado por [32]:

$$\begin{array}{ll} \min & c^t x \\ \text{s.a} & Ax = b \\ & x + s = u \\ & x \geq 0,\ s \geq 0 \end{array} \tag{1.5}$$

em que $s \in \mathbb{R}^n$ representa as variáveis de folga da canalização e $u \in \mathbb{R}^n$ representa o limitante superior de x.

Caso o limite inferior da variável x seja diferente de zero, fazendo transformações de variáveis obtém-se um problema de programação linear equivalente a (1.5).

O problema dual associado na forma padrão é dado por:

$$\begin{array}{ll} \max & b^t y - u^t w \\ \text{s.a} & A^t y + z - w = c \\ & z \geq 0,\ w \geq 0 \end{array} \tag{1.6}$$

em que $w \in \mathbb{R}^n$ representa as variáveis duais associadas as variáveis de folga s.

Quando $x, s, z, w > 0$ temos que (x, s) é um ponto interior no problema primal e (z, w) é um ponto interior no problema dual. Além disso, se (x, s) satifaz as restrições do problema primal então ele é um ponto interior factível e se (z, w) satisfaz as restrições do problema dual então ele é um ponto interior factível.

Neste caso, o *gap* é dado por:

$$\gamma = c^t x - b^t y + u^t w. \tag{1.7}$$

Observação: Quando um ponto é primal e dual factível sendo o PPL com variáveis canalizadas, o *gap* é dado por $\gamma = x^t z + s^t w$.

Teorema 2: Um ponto (x, s, y, z, w) é um ponto ótimo para o problema primal e dual se, e somente se, satisfaz as condições de otimalidade de primeira ordem (Karush-Kuhn-Tucker) do problema (1.5) e (1.6) que são dadas por [32]:

(i) Primal factível

$$\begin{array}{ll} Ax - b = 0 & (x, s) \geq 0 \\ x + s - u = 0 & \end{array}$$

(ii) Dual factível

$$A^t y + z - w - c = 0 \quad (z, w) \geq 0$$

(iii) Complementariedade

$$\begin{array}{l} XZe = 0 \\ SWe = 0 \end{array}$$

(1.8)

sendo $S = diag(s)$ e $W = diag(w)$.

1.2 Métodos de pontos interiores do tipo primal dual

O método de pontos interiores do tipo primal dual é base para implementações de outros métodos de pontos interiores atuais, sendo o método preditor corretor um deles. Assim, a partir do método primal dual afim escala desenvolveremos o método preditor corretor, o qual é o foco principal neste trabalho. Os métodos de pontos interiores do tipo primal dual são desenvolvidos utilizando o Método de Newton. Uma breve descrição deste método é dada a seguir.

1.2.1 Método de Newton

Seja $F : \mathbb{R}^n \to \mathbb{R}^n$, sendo

$$F(x) = \begin{bmatrix} F_1(x) \\ F_2(x) \\ \vdots \\ F_n(x) \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

O método de Newton é um método iterativo para encontrar uma solução aproximada $x^* \in \mathbb{R}^n$, de $F(x^*) = 0$ [8].

Expandindo a função F em série de Taylor na vizinhança de um ponto x, pelos dois primeiros termos, temos:

$$F(x + d) \simeq F(x) + J(x)d,$$

em que $J(x)$ é a matriz Jacobiana no ponto x :

$$J(x) = \begin{bmatrix} \nabla F_1(x) \\ \nabla F_2(x) \\ \vdots \\ \nabla F_n(x) \end{bmatrix}.$$

Impondo a condição $F(x) + J(x)d = 0$, temos:

$$J(x)d = -F(x).$$

Dessa maneira, o método iterativo é dado por:

$$x^{k+1} = x^k + d^k.$$

O processo iterativo é aplicado até que ocorra a convergência.

Observação: A direção d, descrita acima, é conhecida como a direção de Newton.

O método de Newton sob condições adequadas envolvendo x_0, $F(x)$ e $J(x)$ produz uma sequência $\{x^k\}$ que converge a x^* com taxa quadrática. Vale ressaltar que os resultados de convergência obtidos são locais de forma que x_0 deve estar suficientemente próximo de x^*. Para mais detalhes sobre a convergência deste método veja [8].

1.2.2 Método primal dual afim escala

Os métodos de pontos interiores do tipo primal dual consistem em aplicar o método de Newton a $F(x, y, z) = 0$, sendo esta formada pelas condições de otimalidade do PPL (1.4), desconsiderando $(x, z) \geq 0$. Dessa maneira, os problemas primal e dual são resolvidos simultaneamente. Temos que:

$$F(x, y, z) = \begin{pmatrix} Ax - b \\ A^t y + z - c \\ XZe \end{pmatrix} \tag{1.9}$$

Neste trabalho usamos a notação:

$$\begin{aligned} r_p &= b - Ax, \\ r_d &= c - A^t y - z, \\ r_a &= -XZe. \end{aligned} \tag{1.10}$$

Então, determinamos a direção afim escala, utilizando o método de Newton. Assim, temos um método iterativo dado por:

$$(x^{k+1}, y^{k+1}, z^{k+1})^t = (x^k, y^k, z^k)^t + \tilde{d}^k$$

em que

$$\tilde{d}^k = -[J(x^k, y^k, z^k)]^{-1} F(x^k, y^k, z^k) \text{ e } J(x^k, y^k, z^k) = \begin{pmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z^k & 0 & X^k \end{pmatrix}.$$

Ignorando o índice k por simplicidade de notação, obtemos o seguinte sistema de equações lineares:

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{pmatrix} \begin{pmatrix} \tilde{d}x \\ \tilde{d}y \\ \tilde{d}z \end{pmatrix} = \begin{pmatrix} r_p \\ r_d \\ r_a \end{pmatrix} \tag{1.11}$$

ou seja,

$$\begin{cases} A\tilde{dx} = r_p & (1.12) \\ A^t\tilde{dy} + \tilde{dz} = r_d & (1.13) \\ Z\tilde{dx} + X\tilde{dz} = r_a. & (1.14) \end{cases}$$

A solução $(\tilde{dx}, \tilde{dy}, \tilde{dz})$ deste sistema pode ser obtida por eliminação de variáveis. Pela equação (1.14) e observando que X e Z são inversíveis, temos:

$$\tilde{dz} = X^{-1}(r_a - Z\tilde{dx}). \tag{1.15}$$

Substituindo (1.15) na equação (1.13), obtemos:

$$\begin{aligned} A^t\tilde{dy} + X^{-1}(r_a - Z\tilde{dx}) &= r_d \\ A^t\tilde{dy} - X^{-1}Z\tilde{dx} &= r_d - X^{-1}r_a. \end{aligned}$$

Definindo, $D = Z^{-1}X$, e reescrevendo a última igualdade, temos:

$$A^t\tilde{dy} - D^{-1}\tilde{dx} = r_d - X^{-1}r_a.$$

Assim,

$$\tilde{dx} = D(A^t\tilde{dy} - r_d + X^{-1}r_a). \tag{1.16}$$

Substituindo (1.16) na equação (1.12), obtemos:

$$\begin{aligned} AD(A^t\tilde{dy} - r_d + X^{-1}r_a) &= r_p \\ ADA^t\tilde{dy} &= r_p + ADr_d - ADX^{-1}r_a. \end{aligned}$$

A matriz ADA^t é simétrica e definida positiva [15], pois A tem posto completo e D é uma matriz diagonal definida positiva.

Logo,

$$\tilde{dy} = (ADA^t)^{-1}(r_p + ADr_d - ADX^{-1}r_a).$$

Desse modo, para determinar $\tilde{dy}$ basta resolver um sistema de equações lineares cuja matriz dos coeficientes é ADA^t. Para isto, utiliza-se a fatoração Cholesky [11].

Com a direção já calculada, necessitamos determinar o tamanho do passo primal e dual, de forma que:

$$\begin{aligned} x &= x + \tilde{\alpha}_p\tilde{dx} > 0 \\ z &= z + \tilde{\alpha}_d\tilde{dz} > 0 \end{aligned}$$

em que $\tilde{\alpha}_p \in (0,1]$ e $\tilde{\alpha}_d \in (0,1]$.

Caso existam $\tilde{dx}_i < 0$, em que $1 \leq i \leq n$ devemos tomar cuidado para não obter um novo ponto negativo. Desse modo, calculamos ρ_x por:

$$\rho_x = \min_i \{-\frac{x_i}{\tilde{dx}_i} | \tilde{dx}_i < 0\}.$$

Analogamente, quando $\tilde{dz}_i < 0$, em que $1 \leq i \leq n$, ρ_z é dado por:

$$\rho_z = \min_i \{-\frac{z_i}{\tilde{dz}_i} | \tilde{dz}_i < 0\}.$$

Caso não existam $\tilde{dx}_i < 0$, então não existe nenhuma variável de bloqueio, assim podemos tomar o passo de Newton completo, ou seja, $\tilde{\alpha_p} = 1$.

Da mesma forma, caso não existam $\tilde{dz}_i < 0$, então $\tilde{\alpha_d} = 1$. Assim, o tamanho do passo primal e dual são dados, respectivamente, por:

$$\begin{aligned} \tilde{\alpha}_p &= \min\{1, \tau\rho_x\} \\ \tilde{\alpha}_d &= \min\{1, \tau\rho_z\} \end{aligned} \tag{1.17}$$

em que $\tau \in (0,1)$.

Observação: Na prática, o tamanho do passo tanto primal quanto dual é no máximo 1, sendo estes de tamanhos diferentes entre si. Dessa forma, o número de iterações para o método convergir reduz em duas ou três iterações tipicamente [12].

O método primal dual afim escala é resumido à seguir:

Algoritmo 1 Método Primal Dual Afim Escala

Dados (x^0, y^0, z^0) com $x^0, z^0 > 0$ e $\tau \in (0, 1)$

1. Repita
   2. Calcule os resíduos por (1.10)
   3. Calcule $D = Z^{-1}X$
   4. Calcule a direção afim escala $\tilde{d}$
      $\tilde{dy} = (ADA^t)^{-1}(r_p + ADr_d - ADX^{-1}r_a)$
      $\tilde{dx} = D(A^t\tilde{dy} - r_d + X^{-1}r_a)$
      $\tilde{dz} = X^{-1}(r_a - Z\tilde{dx})$
   5. Calcule o tamanho dos passos primal $\tilde{\alpha_p}$ e dual $\tilde{\alpha_d}$ por (1.17)
   6. Atualize o ponto
      $x = x + \tilde{\alpha_p}\tilde{dx}$
      $y = y + \tilde{\alpha_d}\tilde{dy}$
      $z = z + \tilde{\alpha_d}\tilde{dz}$
7. até convergir

1.2.3 Método primal dual afim escala com variáveis canalizadas

Nesta seção, descrevemos o método de pontos interiores primal dual afim escala com variáveis canalizadas. Neste caso, $F(x, s, y, z, w) = 0$ sendo definida pelas condições de otimalidade (1.8). Temos que:

$$F(x, s, y, z, w) = \begin{pmatrix} Ax - b \\ x + s - u \\ A^t y + z - w - c \\ XZe \\ SWe \end{pmatrix}. \tag{1.18}$$

Neste caso, usamos a notação:

$$\begin{aligned} r_p &= b - Ax, \\ r_u &= u - x - s, \\ r_d &= c - A^t y - z + w, \\ r_a &= -XZe, \\ r_b &= -SWe. \end{aligned} \tag{1.19}$$

Para determinar a direção afim escala, utilizamos o método de Newton, cujo processo iterativo é dado por:

$$(x^{k+1}, s^{k+1}, y^{k+1}, z^{k+1}, w^{k+1})^t = (x^k, s^k, y^k, z^k, w^k)^t + \tilde{d}^k$$

em que

$$\tilde{d}^k = -[J(x^k, s^k, y^k, z^k, w^k)]^{-1} F(x^k, s^k, y^k, z^k, w^k)$$

e

$$J(x^k, s^k, y^k, z^k, w^k) = \begin{pmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z^k & 0 & 0 & X^k & 0 \\ 0 & W^k & 0 & 0 & S^k \end{pmatrix}.$$

Ignorando o índice k por simplicidade de notação, temos o seguinte sistema de equações lineares:

$$\begin{pmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{pmatrix} \begin{pmatrix} \tilde{d}x \\ \tilde{d}s \\ \tilde{d}y \\ \tilde{d}z \\ \tilde{d}w \end{pmatrix} = \begin{pmatrix} r_p \\ r_u \\ r_d \\ r_a \\ r_b \end{pmatrix}, \tag{1.20}$$

ou seja,

$$\begin{cases} A\tilde{d}x = r_p & (1.21) \\ \tilde{d}x + \tilde{d}s = r_u & (1.22) \\ A^t\tilde{d}y + \tilde{d}z - \tilde{d}w = r_d & (1.23) \\ X\tilde{d}z + Z\tilde{d}x = r_a & (1.24) \\ S\tilde{d}w + W\tilde{d}s = r_b. & (1.25) \end{cases}$$

Resolvendo o sistema pela eliminação de variáveis encontramos a solução $(\tilde{d}x, \tilde{d}s, \tilde{d}y, \tilde{d}z, \tilde{d}w)$. Pelas equações (1.24) e (1.25), temos:

$$\tilde{d}z = X^{-1}(r_a - Z\tilde{d}x), \tag{1.26}$$

$$\tilde{d}w = S^{-1}(r_b - W\tilde{d}s). \tag{1.27}$$

Pela equação (1.22), obtemos:

$$\tilde{ds} = r_u - \tilde{dx}. \tag{1.28}$$

Substituindo os valores $\tilde{ds}$, $\tilde{dz}$ e $\tilde{dw}$ em (1.23), temos :

$$\begin{aligned}
&A^t\tilde{dy} + X^{-1}(r_a - Z\tilde{dx}) - S^{-1}(r_b - W(r_u - \tilde{dx})) = r_d\\
&A^t\tilde{dy} - X^{-1}Z\tilde{dx} - S^{-1}W\tilde{dx} = r_d - X^{-1}r_a + S^{-1}r_b - S^{-1}Wr_u\\
&A^t\tilde{dy} - (X^{-1}Z + S^{-1}W)\tilde{dx} = r_d - X^{-1}r_a + S^{-1}r_b - S^{-1}Wr_u.
\end{aligned}$$

Definindo $D = Z^{-1}X + W^{-1}S$ e reescrevendo a última igualdade obtemos:

$$A^t\tilde{dy} - D^{-1}\tilde{dx} = r_d - X^{-1}r_a + S^{-1}r_b - S^{-1}Wr_u.$$

Logo,

$$\tilde{dx} = D(A^t\tilde{dy} - r_d + X^{-1}r_a - S^{-1}r_b + S^{-1}Wr_u).$$

Da equação (1.21), determinamos $\tilde{dy}$ da seguinte forma:

$$\begin{aligned}
&AD(A^t\tilde{dy} - r_d + X^{-1}r_a - S^{-1}r_b + S^{-1}Wr_u) = r_p\\
&ADA^t\tilde{dy} = r_p + AD(r_d - X^{-1}r_a + S^{-1}r_b - S^{-1}Wr_u).
\end{aligned}$$

Logo,

$$\tilde{dy} = (ADA^t)^{-1}(r_p + AD(r_d - X^{-1}r_a + S^{-1}r_b - S^{-1}Wr_u)).$$

Como visto anteriormente, a matriz ADA^t é simétrica definida positiva e, portanto, admite a fatoração de Cholesky.

Com o acréscimo de s e w devido à canalização das variáveis é necessário além de ρ_x e ρ_z os cálculos de ρ_s e ρ_w para determinamos o tamanho dos passos primal e dual, sendo:

$$\begin{aligned}
&\rho_x = \min_i\{-\frac{x_i}{\tilde{dx}_i}|\tilde{dx}_i < 0\}, \quad \rho_s = \min_i\{-\frac{s_i}{\tilde{ds}_i}|\tilde{ds}_i < 0\}\\
&\rho_z = \min_i\{-\frac{z_i}{\tilde{dz}_i}|\tilde{dz}_i < 0\} \ \text{ e } \ \rho_w = \min_i\{-\frac{w_i}{\tilde{dw}_i}|\tilde{dw}_i < 0\}.
\end{aligned}$$

Logo, o tamanho dos passos primal e dual são dados respectivamente por:

$$\begin{aligned}
\tilde{\alpha_p} &= \min\{1, \tau\rho_x, \tau\rho_s\}\\
\tilde{\alpha_d} &= \min\{1, \tau\rho_z, \tau\rho_w\}
\end{aligned} \tag{1.29}$$

em que $\tau \in (0, 1)$.

O método primal dual afim escala com variáveis canalizadas é resumido a seguir:

Algoritmo 2 Método Primal Dual Afim Escala Com Variáveis Canalizadas

Dados $(x^0, s^0, y^0, z^0, w^0)$ com $x^0, s^0, z^0, w^0 > 0$ e $\tau \in (0, 1)$

1. Repita
   2. Calcule os resíduos por (1.19)
   3. Calcule $D = Z^{-1}X + W^{-1}S$
   4. Calcule a direção afim escala $\tilde{d}$

      $\tilde{dy} = (ADA^t)^{-1}(r_p + AD(r_d - X^{-1}r_a + S^{-1}r_b - S^{-1}Wr_u)$

      $\tilde{dx} = D(A^t\tilde{dy} - r_d + X^{-1}r_a - S^{-1}r_b + S^{-1}Wr_u)$

      $\tilde{ds} = r_u - \tilde{dx}$

      $\tilde{dw} = S^{-1}(r_b - W\tilde{ds})$

      $\tilde{dz} = X^{-1}(r_a - Z\tilde{dx})$
   5. Calcule o tamanho dos passos primal $\tilde{\alpha_p}$ e dual $\tilde{\alpha_d}$ por (1.29)
   6. Atualize o ponto

      $x = x + \tilde{\alpha_p}\tilde{dx}$

      $s = s + \tilde{\alpha_p}\tilde{ds}$

      $y = y + \tilde{\alpha_d}\tilde{dy}$

      $z = z + \tilde{\alpha_d}\tilde{dz}$

      $w = w + \tilde{\alpha_d}\tilde{dw}$
7. até convergir

Observação: Nos métodos descritos anteriormente, os produtos $x_i z_i$ de r_a e $s_i w_i$ de r_b podem convergir para zero com velocidades diferentes e, assim, os métodos podem falhar ou progredir lentamente. Para evitar isto, cria-se um parâmetro (μ) nas condições de otimalidade de forma que $x_i z_i = \mu$ e $s_i w_i = \mu$, sendo este parâmetro atualizado a cada iteração e tendendo a zero à medida que o método se aproxima da solução. Com esta alteração temos o método seguidor de caminho [32].

1.2.4 Método preditor corretor

O método preditor corretor desenvolvido por Mehrotra [27] consiste em utilizar uma direção que contém três componentes: direção afim escala (direção de Newton), direção de centragem e direção de correção não linear. A direção afim escala é determinada conforme descrito anteriormente no método de pontos interiores primal dual afim escala. A direção de centragem é calculada como descrito na observação anterior para evitar que os produtos $x_i z_i$ de r_a convirjam para zero com velocidades diferentes. Neste caso, determinamos o valor μ de acordo com o progresso da direção afim escala, de modo que se ocorrer melhora suficiente, ou seja, redução do *gap*, então μ é pequeno, caso contrário, μ é uma perturbação maior.

Sejam:

$$\begin{aligned} \gamma &= x^t z, \\ \tilde{\gamma} &= (x + \tilde{\alpha}_p \tilde{dx})^t (z + \tilde{\alpha}_d \tilde{dz}). \end{aligned} \tag{1.30}$$

A proposta de Mehrotra para o cálculo de μ é [27]:

$$\mu = \begin{cases} \left(\frac{\gamma}{\sqrt{n}}\right)\left(\frac{\gamma}{n}\right), \text{ se } \gamma < 1; \\ \left(\frac{\tilde{\gamma}}{\gamma}\right)^p \left(\frac{\gamma}{n}\right), \text{ caso contrário.} \end{cases} \tag{1.31}$$

Observação: Geralmente, o valor de p é 2 ou 3 [27].

Resta somente o cálculo da direção corretora, a qual é obtida fazendo uma correção não linear. Observe que quando tomamos o tamanho do passo igual a 1, ou seja, $\tilde{\alpha_p}^k = \tilde{\alpha_d}^k = 1$, obtemos um ponto que satifaz as restrições de igualdade da factibilidade primal e dual, pois:

$$r_p^{k+1} = b - Ax^{k+1} = b - A(x^k + \tilde{dx}^k) = b - Ax^k - A\tilde{dx}^k = r_p^k - r_p^k = 0,$$

$$r_d^{k+1} = c - A^t y^{k+1} - z^{k+1} = c - A^t(y^k + \tilde{d}y^k) - (z^k + \tilde{d}z^k) =$$
$$= c - A^t y^k - z^k - A^t \tilde{d}y^k - \tilde{d}z^k = r_d^k - r_d^k = 0.$$

Na restrição de complementariedade, temos

$$r_a^{k+1} = -X^{k+1}Z^{k+1}e = -(X^k + \tilde{D}x^k)(Z^k + \tilde{D}z^k)e =$$
$$= -X^k Z^k e - X^k \tilde{D}z^k e - Z^k \tilde{D}x^k e - \tilde{D}x^k \tilde{D}z^k e =$$
$$= r_a^k - r_a^k - \tilde{D}x^k \tilde{D}z^k = -\tilde{D}x^k \tilde{D}z^k,$$

sendo $\tilde{D}x^k = diag(\tilde{d}x^k)$ e $\tilde{Dz^k} = diag(\tilde{d}z^k)$.

Fazendo esta correção não linear e introduzindo a perturbação para centragem, temos o sistema linear:

$$\begin{cases} A\bar{d}x = 0 \\ A^t\bar{d}y + \bar{d}z = 0 \\ Z\bar{d}x + X\bar{d}z = \mu e - \tilde{D}x\tilde{D}ze. \end{cases}$$

Desse modo, a direção preditora corretora (d) é dada pela soma das duas direções encontradas, ou seja, $d = \tilde{d} + \bar{d}$. Portanto, d é determinada pelo sistema:

$$\begin{cases} Adx = r_p \\ A^t dy + dz = r_d \\ Zdx + Xdz = r_s \end{cases} \tag{1.32}$$

em que

$$r_s = r_a + \mu e - \tilde{D}x\tilde{D}ze. \tag{1.33}$$

Resolvendo o sistema (1.32) de forma similar a (1.12)-(1.14), obtemos a direção preditora corretora dada por:

$$\begin{aligned} dy &= (ADA^t)^{-1}[r_p + AD(r_d - X^{-1}r_s)] \\ dx &= D(A^t dy - r_d + X^{-1}r_s) \\ dz &= X^{-1}(r_s - Zdx). \end{aligned} \tag{1.34}$$

Observe que, assim como para determinar a direção afim escala, o passo mais caro consiste em resolver um sistema linear com a mesma matriz ADA^t.

O tamanho do passo primal e dual são dados, respectivamente, por:

$$\begin{aligned} \alpha_p &= \min\{1, \tau\rho_x\}, \\ \alpha_d &= \min\{1, \tau\rho_z\} \end{aligned} \tag{1.35}$$

em que $\tau \in (0,1)$,

$$\rho_x = \min_i\{-\frac{x_i}{dx_i}|dx_i < 0\},$$
$$\rho_z = \min_i\{-\frac{z_i}{dz_i}|dz_i < 0\}.$$

O método preditor corretor é resumido a seguir:

Algoritmo 3 Método Preditor Corretor

Dados (x^0, y^0, z^0) com $x^0, z^0 > 0$ e $\tau \in (0,1)$

1. Repita
   2. Calcule os resíduos por (1.10)
   3. Calcule $D = Z^{-1}X$
   4. Calcule a direção afim escala $\tilde{d}$
      $\tilde{d}y = (ADA^t)^{-1}(r_p + ADr_d - ADX^{-1}r_a)$
      $\tilde{d}x = D(A^t\tilde{d}y - r_d + X^{-1}r_a)$
      $\tilde{d}z = X^{-1}(r_a - Z\tilde{d}x)$
   5. Calcule o tamanho dos passos passos primal $\tilde{\alpha_p}$ e dual $\tilde{\alpha_d}$ por (1.17)
   6. Calcule $\tilde{\gamma}$ por (1.30)
   7. Calcule μ por (1.31)
   8. Calcule r_s por (1.33)
   9. Calcule a direção preditora corretora d
      $dy = (ADA^t)^{-1}[r_p + AD(r_d - X^{-1}r_s)]$
      $dx = D(A^tdy - r_d + X^{-1}r_s)$
      $dz = X^{-1}(r_s - Zdx)$
   10. Calcule o tamanho dos passos primal α_p e dual α_d por (1.35)
   11. Atualize o ponto
      $x = x + \alpha_p dx$
      $y = y + \alpha_d dy$
      $z = z + \alpha_d dz$
12. até convergir

1.2.5 Método preditor corretor com variáveis canalizadas

O método de pontos interiores preditor corretor com variáveis canalizadas é desenvolvido de maneira similar ao método sem canalização.

Sejam:

$$\begin{aligned} &\gamma = x^t z + s^t w, \\ &\tilde{\gamma} = (x + \tilde{\alpha}_p \tilde{dx})^t (z + \tilde{\alpha}_d \tilde{dz}) + (s + \tilde{\alpha}_p \tilde{ds})^t (w + \tilde{\alpha}_d \tilde{dw}). \end{aligned} \tag{1.36}$$

Neste caso:

$$\mu = \begin{cases} \left(\frac{\gamma}{\sqrt{2n}}\right)\left(\frac{\gamma}{2n}\right) \text{ se } \gamma < 1; \\ \left(\frac{\tilde{\gamma}}{\gamma}\right)^p \left(\frac{\gamma}{2n}\right), \text{ caso contrário.} \end{cases} \tag{1.37}$$

Ao realizar a correção não linear quando as variáveis são canalizadas, obtemos um ponto que satifaz as restrições de igualdade da factibilidade primal e dual, devido que:

$$r_p^{k+1} = b - Ax^{k+1} = b - A(x^k + \tilde{dx}^k) = b - Ax^k - A\tilde{dx}^k = r_p^k - r_p^k = 0,$$

$$\begin{aligned} r_u^{k+1} &= u - x^{k+1} - s^{k+1} = u - (x^k + \tilde{dx}^k) - (s^k + \tilde{ds}^k) = \\ &= u - x^k - s^k - \tilde{dx}^k - \tilde{ds}^k = r_u^k - r_u^k = 0, \end{aligned}$$

$$\begin{aligned} r_d^{k+1} &= c - A^t y^{k+1} - z^{k+1} + w^{k+1} = c - A^t(y^k + \tilde{dy}^k) - (z^k + \tilde{dz}^k) + (w^k + \tilde{dw}^k) = \\ &= c - A^t y^k - z^k + w^k - A^t \tilde{dy}^k - \tilde{dz}^k + A^t \tilde{dw}^k = r_d^k - r_d^k = 0. \end{aligned}$$

Nas restrições de complementariedade, temos:

$$\begin{aligned} r_a^{k+1} &= -X^{k+1} Z^{k+1} e = -(X^k + \tilde{D}x^k)(Z^k + \tilde{D}z^k)e = \\ &= -X^k Z^k e - X^k \tilde{D}z^k e - Z^k \tilde{D}x^k e - \tilde{D}x^k \tilde{D}z^k e = \\ &= r_a^k - r_a^k - \tilde{D}x^k \tilde{D}z^k = -\tilde{D}x^k \tilde{D}z^k, \end{aligned}$$

e

$$\begin{aligned} r_b^{k+1} &= -S^{k+1} W^{k+1} e = -(S^k + \tilde{D}s^k)(W^k + \tilde{D}w^k)e = \\ &= -S^k W^k e - S^k \tilde{D}w^k e - W^k \tilde{D}s^k e - \tilde{D}s^k \tilde{D}w^k e = \\ &= r_b^k - r_b^k - \tilde{D}s^k \tilde{D}w^k = -\tilde{D}s^k \tilde{D}w^k. \end{aligned}$$

Fazendo a correção não linear e introduzindo a perturbação, obtemos o sistema:

$$\begin{cases} A\bar{d}x = 0 \\ \bar{d}x + \bar{d}s = 0 \\ A^t\bar{d}y + \bar{d}z - \bar{d}w = 0 \\ Z\bar{d}x + X\bar{d}z = \mu e - \tilde{D}x\tilde{D}ze \\ W\bar{d}s + S\bar{d}w = \mu e - \tilde{D}s\tilde{D}we. \end{cases}$$

Como visto no caso anterior, a direção preditora corretora (d) é dada pela soma das duas direções encontradas. Assim, d é determinada pelo sistema:

$$\begin{cases} Adx = r_p \\ dx + ds = r_u \\ A^tdy + dz - dw = r_d \\ Zdx + Xdz = r_s \\ Wds + Sdw = r_t \end{cases} \tag{1.38}$$

em que

$$\begin{aligned} r_s &= r_a + \mu e - \tilde{D}x\tilde{D}ze \\ r_t &= r_b + \mu e - \tilde{D}s\tilde{D}we. \end{aligned} \tag{1.39}$$

Podemos resolver o sistema (1.38) de forma similar a (1.20), obtendo a seguinte solução:

$$\begin{aligned} dy &= (ADA^t)^{-1}[r_p + AD(r_d - X^{-1}r_s)] \\ dx &= D(A^tdy - r_d + X^{-1}r_s) \\ ds &= r_u - dx \\ dz &= X^{-1}(r_s - Zdx) \\ dw &= S^{-1}(r_t - Wds). \end{aligned}$$

Neste caso, o tamanho dos passos primal e dual são dados respectivamente por:

$$\begin{aligned} \alpha_p &= \min\{1, \tau\rho_x, \tau\rho_s\}, \\ \alpha_d &= \min\{1, \tau\rho_z, \tau\rho_w\} \end{aligned} \tag{1.40}$$

em que $\tau \in (0,1)$,

$$\rho_x = \min_i\{-\frac{x_i}{dx_i}|dx_i < 0\}, \quad \rho_s = \min_i\{-\frac{s_i}{ds_i}|ds_i < 0\},$$
$$\rho_z = \min_i\{-\frac{z_i}{dz_i}|dz_i < 0\} \text{ e } \rho_w = \min_i\{-\frac{w_i}{dw_i}|dw_i < 0\}.$$

O método preditor corretor com variáveis canalizadas é resumido a seguir:

Algoritmo 4 Método Preditor Corretor Com Variáveis Canalizadas

Dados $(x^0, s^0, y^0, z^0, w^0)$ com $x^0, s^0, z^0, w^0 > 0$ e $\tau \in (0, 1)$

1. Repita
   2. Calcule os resíduos por (1.19)
   3. Calcule $D = Z^{-1}X + W^{-1}S$
   4. Calcule a direção afim escala $\tilde{d}$

      $\tilde{dy} = (ADA^t)^{-1}(r_p + AD(r_d - X^{-1}r_a + S^{-1}r_b - S^{-1}Wr_u)$

      $\tilde{dx} = D(A^t\tilde{dy} - r_d + X^{-1}r_a - S^{-1}r_b + S^{-1}Wr_u)$

      $\tilde{ds} = r_u - \tilde{dx}$

      $\tilde{dw} = S^{-1}(r_b - W\tilde{ds})$

      $\tilde{dz} = X^{-1}(r_a - Z\tilde{dx})$
   5. Calcule o tamanho dos passos primal $\tilde{\alpha_p}$ e dual $\tilde{\alpha_d}$ por (1.29)
   6. Calcule $\tilde{\gamma}$ por (1.36)
   7. Calcule μ por (1.37)
   8. Calcule r_s e r_t por (1.39)
   9. Calcule a direção preditora corretora d

      $dy = (ADA^t)^{-1}[r_p + AD(r_d - X^{-1}r_s)]$

      $dx = D(A^tdy - r_d + X^{-1}r_s)$

      $ds = r_u - dx$

      $dz = X^{-1}(r_s - Zdx)$

      $dw = S^{-1}(r_t - Wds)$
   10. Calcule o tamanho dos passos primal α_p e dual α_d por (1.40)
   11. Atualize o ponto

       $x = x + \alpha_p dx$

       $s = s + \alpha_p ds$

       $y = y + \alpha_d dy$

       $z = z + \alpha_d dz$

       $w = w + \alpha_d dy$
12. até convergir

Portanto, o método preditor corretor resolve dois sistemas lineares com a mesma matriz ADA^t em cada iteração. O esforço computacional por iteração do método com e sem variáveis canalizadas é essencialmente o mesmo.

Capítulo 2

Método Preditor Corretor com Iteração Continuada

O método preditor corretor é uma das variações do método de pontos interiores que mais se destaca, devido a sua eficiência e convergência. Neste método, que foi apresentado no Capítulo 1, é necessário resolver dois sistemas lineares em cada iteração com a matriz ADA^t. Como esta matriz é simétrica definida positiva a fatoração de Cholesky é utilizada para resolver os sistemas envolvidos, sendo este o passo que requer mais esforço computacional durante a aplicação do método. Assim, com o objetivo de reduzir esforços, mais precisamente reduzir o número total de fatoração de Cholesky realizadas, desenvolvemos a iteração continuada para este método. Utilizamos como motivação o método do gradiente projetado [24], estendendo os resultados apresentados por Oliveira, Lyra e Correia [29] para o método dual afim escala [1]. Na figura 2.1 os eixos representam as variáveis y_1 e y_2. Cada segmento de reta representa um restrição do tipo $a_i^t y + z_i \leq c_i$, onde $z_i = 0$ está exatamente sobre a reta correspondente e todos os z_i são positivos na região colorida. A proposta de [29] consiste em fixar em zero a componente de bloqueio k na direção dz, ou seja, $dz_k = 0$. Na figura 2.1, temos a trajetória da direção quando esta proposta é utilizada, de forma a obter um ponto próximo da fronteira do politopo. A estratégia pode ser utilizada enquanto a direção projetada melhora a função objetivo do problema dual.

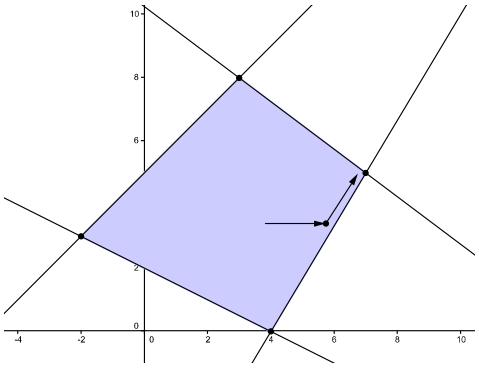

Figura 2.1: Proposta inicial no método dual afim escala

Para o desenvolvimento da iteração continuada utilizamos um problema auxiliar de programação não linear e, para resolvê-lo, utilizamos a função Lagrangiana.

2.1 Função Lagrangiana

Considere o problema:

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & h_i(x) = 0 \qquad i = 1, \ldots, m. \end{array} \tag{2.1}$$

em que $f(x) : \mathbb{R}^n \to \mathbb{R},\ h_i(x) : \mathbb{R}^n \to \mathbb{R},\ i = 1, \ldots, m.$

Por conveniência de notação, assumimos a função $h(x) : \mathbb{R}^n \to \mathbb{R}^m$ em que $h = (h_1, \ldots, h_m)$.

A função Lagrangiana $l(x, \lambda) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ associada ao problema (2.1) é dada por [2]:

$$l(x, \lambda) = f(x) + \lambda^t h(x).$$

Proposição 1 (Condições necessárias de 1ª e 2ª ordem) Seja x^* um ponto de mínimo local de f sujeito a $h(x) = 0$ e assumindo que $\nabla h_1(x^*), \ldots, \nabla h_m(x^*)$ são linearmente independentes. Então existe um único vetor $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)^t$, tal que:

$$l(x^*, \lambda^*) = \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) = 0.$$

Os elementos λ^* são chamados de multiplicadores de Lagrange.

Assumindo $f, g \in \mathcal{C}^2$, temos:

$$y^t(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*))\ y \geq 0$$

para todo $y \in M(x^*)$, em que $M(x^*) = \{y \mid \nabla h_i(x^*)^t y = 0,\ i = 1, \ldots, m\}$.

Podemos escrever esta proposição de forma compacta, sendo que:

Se x^* é um mínimo local e $\nabla h_1(x^*), \ldots, \nabla h_m(x^*)$ são linearmente independentes, então:

$$\nabla_x l(x^*, \lambda^*) = 0 \ \text{ e } \ \nabla_\lambda l(x^*, \lambda^*) = 0 \tag{2.2}$$

$$y^t L(x^*) y \geq 0\ \forall y \in M(x^*) \tag{2.3}$$

em que $L(x^*) = \nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*)$.

A condição (2.2) é conhecida como condição necessária de primeira ordem e a condição (2.3) como condição necessária de segunda ordem.

Proposição 2 (Condições suficientes de 2ª ordem) Sejam $f, g \in \mathcal{C}^2,\ x^* \in \mathbb{R}^n$ e $\lambda^* \in \mathbb{R}^m$ que satisfazem:

$$\nabla_x l(x^*, \lambda^*) = 0 \ \text{ e } \ \nabla_\lambda l(x^*, \lambda^*) = 0$$
$$y^t L(x^*) y \geq 0 \ \ \forall y \neq 0 \text{ com } \nabla h(x^*)^t y = 0.$$

Então x^* é um mínimo local estrito de f sujeito $h(x) = 0$.

Apresentamos agora o desenvolvimento da iteração continuada para aplicação no método preditor corretor com e sem variáveis canalizadas.

2.2 Iteração Continuada

A iteração continuada é desenvolvida, neste trabalho, para aplicação no método preditor corretor, com o objetivo de reduzir o número total de iterações do método de pontos interiores preditor corretor e consequentemente reduzir o número de fatorações de Cholesky. Esta iteração consiste em determinar uma nova direção, após a iteração do método preditor corretor, sem que seja necessário realizar uma nova fatoração de Cholesky. Determinamos a nova direção considerando que algumas componentes sejam nulas e de forma que a alteração em relação à direção preditora corretora seja a mínima possível. Além disso, a fatoração de

Cholesky da matriz ADA^t, já calculada na iteração do método preditor corretor é utilizada novamente para resolver os sistemas lineares envolvidos. Assim, o cálculo desta nova direção não é muito caro em comparação à realização de uma nova iteração do método preditor corretor. Esta nova direção é denominada direção preditora corretora continuada. Seu cálculo é realizado de modo análogo ao da direção preditora corretora, calculando inicialmente a direção afim escala continuada, depois a direção de centragem e a correção não linear continuada.

As componentes que devem ser nulas na nova direção são aquelas responsáveis pelo bloqueio na direção anterior. Então é realizada uma busca pelas componentes responsáveis pelo bloqueio nas direções (dx, dz), ou seja, de forma a encontrar i e j, tais que:

$$\begin{aligned} i &= \arg\min_t \{-\frac{x_t}{dx_t} | dx_t < 0\}, \\ j &= \arg\min_t \{-\frac{z_t}{dz_t} | dz_t < 0\}. \end{aligned} \tag{2.4}$$

Como pode não existir variável de bloqueio em alguma ou em ambas direções, temos alguns casos a considerar. De (2.4), temos:

1. As componentes i e j bloqueiam;
2. Somente a componente i bloqueia;
3. Somente a componente j bloqueia;
4. Nenhuma componente bloqueia.

Observação: Caso exista mais de uma componente i ou j, para o desenvolvimento da iteração continuada neste trabalho, apenas a primeira componente de bloqueio determinada será considerada.

No caso 4, como não existe nenhuma componente de bloqueio, a iteração continuada não é realizada. Assim, analisamos os três primeiros casos:

2.2.1 As componentes i e j bloqueiam

Obtidas as componentes que bloqueiam, na direção afim escala continuada $\hat{d}$ a ser calculada devemos ter:

$$\begin{aligned} \hat{d}x_i &= 0 \\ \hat{d}z_j &= 0. \end{aligned} \tag{2.5}$$

Além disso, $\hat{d}$ deve satisfazer aproximadamente o sistema abaixo:

$$\begin{cases} A\hat{d}x \simeq r_p & (2.6) \\ A^t\hat{d}y + \hat{d}z \simeq r_d & (2.7) \\ X\hat{d}z + Z\hat{d}x \simeq r_a, & (2.8) \end{cases}$$

em que $\hat{d}x, \hat{d}z \in \mathbb{R}^n$ e $\hat{d}y \in \mathbb{R}^m$.

Calculamos $\hat{d}$ aproximadamente, pois o sistema (1.12) - (1.14) é um sistema quadrado de (m+2n) equações e variáveis, e admite solução única. Além disso, estamos impondo novas condições (2.5) a este sistema de equações.

De forma similar à ideia de Dikin [9] ao desenvolver o método primal afim escala, para encontrar a solução deste sistema, utilizamos um problema auxiliar para determinar primeiramente $\hat{d}x$. Neste problema, $\hat{d}x$ é calculado de forma a minimizar a alteração na direção original dx. O problema auxiliar é construído utilizando (2.5) e (2.6). Como em (2.5) existe uma componente de $\hat{d}z$, é necessário encontrar a componente correspondente em $\hat{d}x$. Por (2.8), temos que:

$$\begin{aligned} \hat{d}x &= Z^{-1}(r_a - X\hat{d}z) \\ \hat{d}x &= Z^{-1}(-XZe - X\hat{d}z) \\ \hat{d}x &= -Xe - Z^{-1}X\hat{d}z. \end{aligned}$$

Tomando a j-ésima componente, temos:

$$\begin{aligned} \hat{d}x_j &= -x_j - x_j\frac{\hat{d}z_j}{z_j} \\ \hat{d}x_j &= -x_j(1 + \frac{\hat{d}z_j}{z_j}). \end{aligned}$$

Como $\hat{d}z_j = 0$, segue que

$$\hat{d}x_j = -x_j. \tag{2.9}$$

Observação: As componentes que serão nulas, devem ser diferentes, ou seja, $i \neq j$. Caso contrário, teríamos:

- Se $\hat{d}x_i = 0$ implicaria $\hat{d}z_i = -z_i$;
- Se $\hat{d}z_j = 0$ implicaria $\hat{d}x_j = -x_j$.

Caso $i = j$, então $x_i = z_j = 0$. Mas, isto é um absurdo, pois estamos considerando pontos interiores. Logo, a componente nula i é diferente de j. Portanto, se $i = j$ a iteração continuada não é realizada.

Considere $\bar{dx} = dx$, a última direção calculada. Assim formulamos um problema que determina $\hat{dx}$, de modo que a mudança em relação a direção $\bar{dx}$ seja a mínima possível, por:

$$\begin{array}{ll} \min & \frac{1}{2}\|D^{-\frac{1}{2}}\hat{dx} - D^{-\frac{1}{2}}\bar{dx}\|^2 \\ \text{s.a} & A\hat{dx} = r_p \\ & \hat{dx}_i = \beta_a \\ & \hat{dx}_j = \beta_b \end{array} \tag{2.10}$$

em que $D = Z^{-1}X,\ \hat{dx},\ \bar{dx} \in \mathbb{R}^n,\ \beta_a = 0,\ \beta_b = -x_j$.

Observe que na função objetivo temos a matriz $D^{-\frac{1}{2}}$, isto porque na resolução do problema obteremos uma matriz ADA^t. Desse modo, a fatoração de Cholesky já calculada é novamente utilizada para resolver os sistemas envolvidos.

Como este é um problema não linear, utilizamos a função Lagrangiana para encontrar a solução. A função Lagrangiana deste problema é dada por:

$$l(\hat{dx}, v, \alpha, \gamma) = \frac{1}{2}(D^{-\frac{1}{2}}\hat{dx} - D^{-\frac{1}{2}}\bar{dx})^t(D^{-\frac{1}{2}}\hat{dx} - D^{-\frac{1}{2}}\bar{dx}) + v^t(A\hat{dx} - r_p) + \alpha(\hat{dx}_i - \beta_a) + \gamma(\hat{dx}_j - \beta_b) \tag{2.11}$$

em que $v = (v_1, v_2, \ldots, v_m)^t \in \mathbb{R}^m$ e $\alpha, \gamma \in \mathbb{R}$.

Pela condição necessária de primeira ordem, temos $\nabla l(\hat{dx}, v, \alpha, \gamma) = 0$, ou seja,

$$\begin{cases} D^{-1}\hat{dx} - D^{-1}\bar{dx} + A^t v + \alpha e_i + \gamma e_j = 0 & (2.12) \\ A\hat{dx} - r_p = 0 & (2.13) \\ \hat{dx}_i - \beta_a = 0 & (2.14) \\ \hat{dx}_j - \beta_b = 0, & (2.15) \end{cases}$$

em que e_i e e_j são vetores canônicos, $1 \leq i \leq n,\ 1 \leq j \leq n,\ i \neq j$.

Por (2.12),

$$D^{-1}\hat{dx} = D^{-1}\bar{dx} - A^t v - \alpha e_i - \gamma e_j$$
$$\hat{dx} = \bar{dx} - DA^t v - \alpha De_i - \gamma De_j.$$

Logo,

$$\hat{dx} = \bar{dx} - DA^t v - \alpha d_{ii} e_i - \gamma d_{jj} e_j. \tag{2.16}$$

Por (2.13) e (2.16),

$$r_p = A\hat{dx} = A\bar{dx} - ADA^t v - \alpha d_{ii} A_i - \gamma d_{jj} A_j.$$

Como $r_p = A\bar{dx}$, temos:

$$ADA^t v = -\alpha d_{ii} A_i - \gamma d_{jj} A_j.$$

Visto que a matriz ADA^t é simétrica definida positiva, então:

$$v = -\alpha d_{ii}(ADA^t)^{-1} A_i - \gamma d_{jj}(ADA^t)^{-1} A_j. \tag{2.17}$$

Por (2.14) e (2.16),

$$\beta_a = \hat{dx}_i = \bar{dx}_i - d_{ii} A_i^t v - \alpha d_{ii}. \tag{2.18}$$

Substituindo o valor de v em (2.18) temos:

$$\beta_a = \bar{dx}_i - d_{ii} A_i^t(-\alpha d_{ii}(ADA^t)^{-1} A_i - \gamma d_{jj}(ADA^t)^{-1} A_j) - \alpha d_{ii}$$
$$\beta_a = \bar{dx}_i + \alpha d_{ii}^2 A_i^t(ADA^t)^{-1} A_i + \gamma d_{ii} d_{jj} A_i^t(ADA^t)^{-1} A_j - \alpha d_{ii}.$$

Logo,

$$\alpha d_{ii} - \alpha d_{ii}^2 A_i^t(ADA^t)^{-1} A_i = \bar{dx}_i + \gamma d_{ii} d_{jj} A_i^t(ADA^t)^{-1} A_j - \beta_a.$$

Portanto,

$$\alpha = \frac{\bar{dx}_i - \beta_a + \gamma d_{ii} d_{jj} A_i^t(ADA^t)^{-1} A_j}{d_{ii}(1 - d_{ii} A_i^t(ADA^t)^{-1} A_i)}. \tag{2.19}$$

Por (2.15) e (2.16),

$$\beta_b = \hat{dx}_j = \bar{dx}_j - d_{jj} A_j^t v - \gamma d_{jj}. \tag{2.20}$$

Substituindo o valor de v, dado por (2.17), em (2.20) temos:

$$\beta_b = \bar{dx}_j - d_{jj} A_j^t(-\alpha d_{ii}(ADA^t)^{-1} A_i - \gamma d_{jj}(ADA^t)^{-1} A_j) - \gamma d_{jj}.$$

Logo,

$$\beta_b = \bar{dx}_j + \alpha d_{ii} d_{jj} A_j^t(ADA^t)^{-1} A_i + \gamma d_{jj}^2 A_j^t(ADA^t)^{-1} A_j - \gamma d_{jj}. \tag{2.21}$$

Considere:

$$\begin{aligned}\theta_i &= A_i^t(ADA^t)^{-1}A_i,\\ \theta_j &= A_j^t(ADA^t)^{-1}A_j,\\ \theta_{ij} &= A_i^t(ADA^t)^{-1}A_j.\end{aligned}$$

Reescrevendo (2.21) e substituindo o valor de α, dado por (2.19) obtemos:

$$\beta_b = \bar{d}x_j + \left[\frac{\bar{d}x_i - \beta_a + \gamma d_{ii}d_{jj}\theta_{ij}}{d_{ii}(1-d_{ii}\theta_i)}\right] d_{ii}d_{jj}\theta_{ij} + \gamma d_{jj}^2\theta_j - \gamma d_{jj}$$

$$\beta_b = \bar{d}x_j + \left[\frac{(\bar{d}x_i - \beta_a)d_{jj}\theta_{ij}}{1-d_{ii}\theta_i}\right] + \left[\frac{\gamma d_{ii}d_{jj}^2\theta_{ij}^2}{1-d_{ii}\theta_i}\right] + \gamma d_{jj}^2\theta_j - \gamma d_{jj}.$$

Assim,

$$\gamma d_{jj} - \gamma\left[\frac{d_{ii}d_{jj}^2\theta_{ij}^2}{1-d_{ii}\theta_i}\right] - \gamma d_{jj}^2\theta_j = \bar{d}x_j - \beta_b + \left[\frac{(\bar{d}x_i - \beta_a)d_{jj}\theta_{ij}}{1-d_{ii}\theta_i}\right]. \tag{2.22}$$

Seja:

$$\xi = 1 - \left[\frac{d_{ii}d_{jj}\theta_{ij}^2}{1-d_{ii}\theta_i}\right] - d_{jj}\theta_j.$$

Reescrevendo a equação (2.22), temos:

$$\gamma d_{jj}\xi = \bar{d}x_j - \beta_b + \left[\frac{(\bar{d}x_i - \beta_a)d_{jj}\theta_{ij}}{1-d_{ii}\theta_i}\right].$$

Logo,

$$\gamma = \frac{\bar{d}x_j - \beta_b + \left[\frac{(\bar{d}x_i-\beta_a)d_{jj}\theta_{ij}}{1-dii\theta_i}\right]}{d_{jj}\xi}.$$

Portanto, a partir de (2.11) encontramos $\hat{d}x$, um ponto crítico para $l(\hat{d}x, \alpha, \gamma)$. Analisaremos a matriz Hessiana da função Lagrangiana neste ponto para verificar se $\hat{d}x$ é o mínimo para (2.10).

Considere:

$$\begin{aligned}f(\hat{d}x) &= \frac{1}{2}\|D^{-\frac{1}{2}}\hat{d}x - D^{-\frac{1}{2}}\bar{d}x\|^2,\\ h_1(\hat{d}x) &= \hat{d}x_i - \beta_a,\\ h_2(\hat{d}x) &= \hat{d}x_j - \beta_b,\\ g_k(\hat{d}x) &= A_k^t\hat{d}x - (r_p)_k,\end{aligned}$$

em que A_k é a coluna k da matriz A, $1 \leq k \leq m$. Então a matriz Hessiana de $l(\hat{dx}, v, \alpha, \gamma)$ é dada por:

$$L(\hat{dx}) = F(\hat{dx}) + \sum_{k=1}^{m} v_k G_k(\hat{dx}) + \alpha H_1(\hat{dx}) + \gamma H_2(\hat{dx})$$

em que $F(\hat{dx})$, $G_k(\hat{dx})$, $H_1(\hat{dx})$ e $H_2(\hat{dx})$ são matrizes Hessianas de $f(\hat{dx})$, $g_k(\hat{dx})$, $h_1(\hat{dx})$ e $h_2(\hat{dx})$, respectivamente.

Temos que:

$$\nabla f(\hat{dx}) = D \begin{pmatrix} \hat{dx}_1 - \bar{dx}_1 \\ \hat{dx}_2 - \bar{dx}_2 \\ \vdots \\ \hat{dx}_n - \bar{dx}_n \end{pmatrix}, \ \nabla g_k(\hat{dx}) = A_k, \ \nabla h_1(\hat{dx}) = e_i, \ \nabla h_2(\hat{dx}) = e_j.$$

Logo,

$$F(\hat{dx}) = D, \ G_k(\hat{dx}) \ = \ H_1(\hat{dx}) \ = \ H_2(\hat{dx}) \ = \ 0_{n \times n},$$

para todo $1 \leq k \leq m$. Assim, $L(\hat{dx}) = F(\hat{dx})$. Como $F(\hat{dx})$ é definida positiva para todo valor de $\hat{dx}$, temos que $\hat{dx}$ é um ponto de mínimo global. Portanto, a direção $\hat{dx}$ encontrada é a solução mínima do problema.

Pelas equações (2.7) e (2.8), encontramos as outras direções. Por (2.8), temos:

$$\hat{dz} = X^{-1}(r_a - Z\hat{dx}). \tag{2.23}$$

Como (2.7) é um sistema sobredeterminado, obtemos a solução de quadrados mínimos [15]. Para aproveitar a fatoração de Cholesky já existente, tomamos o novo sistema $D^{\frac{1}{2}} A^t \hat{dy} = D^{\frac{1}{2}}(r_d - \hat{dz})$. Assim, pelas equações normais:

$$\hat{dy} = (ADA^t)^{-1} AD(r_d - \hat{dz}). \tag{2.24}$$

Observação: O cálculo da direção afim escala continuada implica na resolução de três sistemas lineares com ADA^t, sendo dois sistemas no cálculo de v em (2.17) para encontrar $\hat{dx}$ e um sistema para encontrar $\hat{dy}$ em (2.24).

Calculada a direção afim escala continuada, falta realizar a centragem e a correção não linear para obtermos a direção preditora corretora continuada, de modo análogo como feito no método preditor corretor. Então, fazendo o cálculo de μ e a correção não linear, na direção d devemos ter:

$$dx_i = 0$$

$$dz_j = 0$$

Além disso, d deve satisfazer aproximadamente o sistema:

$$\begin{cases} Adx \simeq r_p & (2.25) \\ A^t dy + dz \simeq r_d & (2.26) \\ Zdx + Xdz \simeq r_s & (2.27) \end{cases}$$

em que

$$r_s = r_a + \mu e - \hat{D}_x \hat{D}_z e. \tag{2.28}$$

Este sistema é resolvido de forma similar ao sistema para determinar a direção afim escala continuada descrita anteriormente, sendo que no problema auxiliar (2.10) ocorre alteração no valor β_b devido ao fato da realização da centragem e da correção não linear.

Se $dz_j = 0$, encontramos o valor da componente dx_j correspondente pela relação (2.27).

$$\begin{aligned} dx &= Z^{-1}(r_s - Xdz) \\ dx &= Z^{-1}(r_a + \mu e - \hat{D}_x \hat{D}_z e - Xdz) \\ dx &= Z^{-1}(-XZe + (\mu e - \hat{D}_x \hat{D}_z e) - Xdz) \\ dx &= -Xe - Z^{-1}Xdz + Z^{-1}(\mu e - \hat{D}_x \hat{D}_z e). \end{aligned}$$

Tomando a j-ésima componente, temos:

$$\begin{aligned} dx_j &= -x_j - x_j \tfrac{dz_j}{z_j} + \tfrac{\mu - \hat{d}x_j \hat{d}z_j}{z_j} \\ dx_j &= -x_j(1 + \tfrac{dz_j}{z_j}) + \tfrac{\mu - \hat{d}x_j \hat{d}z_j}{z_j}. \end{aligned}$$

Como $dz_j = 0$, segue que

$$dx_j = -x_j + \frac{\mu - \hat{d}x_j \hat{d}z_j}{z_j}. \tag{2.29}$$

Portanto, o problema que deve ser utilizado é:

$$\begin{aligned} \min \quad & \tfrac{1}{2}\|D^{-\frac{1}{2}}dx - D^{-\frac{1}{2}}\bar{d}x\|^2 \\ \text{s.a} \quad & Adx = r_p \\ & dx_i = \beta_a \\ & dx_j = \beta_b \end{aligned} \tag{2.30}$$

em que $D = Z^{-1}X,\ dx,\ \bar{dx} \in \mathbb{R}^n$,

$$\begin{aligned} \beta_a &= 0 \\ \beta_b &= -x_j + \frac{\mu - \hat{d}x_j \hat{d}z_j}{z_j}. \end{aligned}$$

Este problema auxiliar é resolvido de modo análogo ao anterior. Obtendo:

$$dx = \bar{d}x - DA^t v - \alpha d_{ii} e_i - \gamma d_{jj} e_j \tag{2.31}$$

em que

$$v = -\alpha d_{ii}(ADA^t)^{-1}A_i - \gamma d_{jj}(ADA^t)^{-1}A_j,$$

$$\alpha = \frac{\bar{d}x_i - \beta_a + \gamma d_{ii} d_{jj} A_i^t (ADA^t)^{-1} A_j}{d_{ii}(1 - d_{ii} A_i^t (ADA^t)^{-1} A_i)},$$

$$\gamma = \frac{\bar{d}x_j - \beta_b + \left[\frac{(\bar{d}x_i - \beta_a) d_{jj} \theta_{ij}}{1 - dii\theta_i}\right]}{d_{jj}\xi}.$$

Pelas equações (2.26) e (2.27), encontramos as outras direções.

$$\begin{aligned} dz &= X^{-1}(r_a - Zdx), \\ dy &= (ADA^t)^{-1}AD(r_d - dz). \end{aligned} \tag{2.32}$$

Observação: Para determinar a direção d acima não é necessário resolver novamente os sistemas lineares envolvendo a matriz ADA^t no cálculo de v. Assim, é necessário resolver somente um sistema linear com a matriz ADA^t para determinar dy.

2.2.2 Somente a componente i bloqueia

Como apenas a componente i deve ser nula, então na direção afim escala continuada:

$$\hat{d}x_i = 0. \tag{2.33}$$

Adicionalmente, a direção $\hat{d}$ deve satisfazer aproximadamente o sistema:

$$\begin{cases} A\hat{d}x \simeq r_p & (2.34) \\ A^t\hat{d}y + \hat{d}z \simeq r_d & (2.35) \\ X\hat{d}z + Z\hat{d}x \simeq r_a & (2.36) \end{cases}$$

Este sistema é resolvido de maneira análoga ao caso na subseção (2.2.1). Com as equações (2.33) e (2.34) construímos um problema para determinar $\hat{d}x$, sendo dado por:

$$\begin{aligned} \min \quad & \tfrac{1}{2}\|D^{-\frac{1}{2}}\hat{d}x - D^{-\frac{1}{2}}\bar{d}x\|^2 \\ \text{s.a} \quad & A\hat{d}x = r_p \\ & \hat{d}x_i = \beta_a \end{aligned} \tag{2.37}$$

em que $D = Z^{-1}X,\ \hat{dx},\ \bar{dx} \in \mathbb{R}^n, \beta_a = 0$.

A função Lagrangiana deste problema é a seguinte:

$$l(\hat{dx}, v, \alpha) = \frac{1}{2}(D^{-\frac{1}{2}}\hat{dx} - D^{-\frac{1}{2}}\bar{dx})^t(D^{-\frac{1}{2}}\hat{dx} - D^{-\frac{1}{2}}\bar{dx}) + v^t(A\hat{dx} - r_p) + \alpha(\hat{dx}_i - \beta_a)$$

em que $v = (v_1, v_2, \ldots, v_m)^t \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$.

Pela condição necessária de primeira ordem, temos $\nabla l(\hat{dx}, v, \alpha) = 0$, ou seja,

$$\begin{cases} D^{-1}\hat{dx} - D^{-1}\bar{dx} + A^t v + \alpha e_i = 0 & (2.38) \\ A\hat{dx} - r_p = 0 & (2.39) \\ \hat{dx}_i - \beta_a = 0 & (2.40) \end{cases}$$

em que e_i é um vetor canônico, $1 \leq i \leq n$.

Por (2.38),

$$\hat{dx} = \bar{dx} - DA^t v - \alpha d_{ii} e_i. \tag{2.41}$$

Por (2.39), (2.41) e com o fato de $r_p = A\bar{dx}$, temos:

$$v = -\alpha d_{ii}(ADA^t)^{-1}A_i. \tag{2.42}$$

Por (2.40) e (2.41),

$$\beta_a = \hat{dx}_i = \bar{dx}_i - d_{ii}A_i^t v - \alpha d_{ii}. \tag{2.43}$$

Substituindo em (2.43) o valor de v dado por (2.42) obtemos:

$$\alpha = \frac{\bar{dx}_i - \beta_a}{d_{ii}(1 - d_{ii}A_i^t(ADA^t)^{-1}A_i)}. \tag{2.44}$$

Portanto, encontramos $\hat{dx}$, um ponto crítico. Analisando a matriz Hessiana da função Lagrangiana neste ponto como na subseção (2.2.1), verifica-se que $\hat{dx}$ é a solução mínima do problema auxiliar.

Pelas equações (2.35) e (2.36), encontramos as demais direções:

$$\begin{aligned} \hat{dz} &= X^{-1}(r_a - Z\hat{dx}), \\ \hat{dy} &= (ADA^t)^{-1}AD(r_d - \hat{dz}). \end{aligned} \tag{2.45}$$

Como na subseção (2.2.1), para determinar $\hat{dy}$ resolvemos (2.35) por quadrados mínimos.

Observação: Neste caso, o cálculo da direção afim escala continuada implica na resolução de dois sistemas lineares com ADA^t, sendo um sistema no cálculo de v para encontrar $\hat{dx}$ e o outro para encontrar $\hat{dy}$.

Calculada a direção afim escala continuada, realizando o cálculo da centragem e da correção não linear para obtermos a direção preditora corretora continuada, devemos ter:

$$dx_i = 0. \tag{2.46}$$

A direção d deve satisfazer aproximadamente o sistema linear:

$$\begin{cases} Adx \simeq r_p & (2.47) \\ A^t dy + dz \simeq r_d & (2.48) \\ Zdx + Xdz \simeq r_s & (2.49) \end{cases}$$

em que

$$r_s = r_a + \mu e - \hat{D}_x \hat{D}_z e. \tag{2.50}$$

Este sistema é resolvido de forma similar ao sistema para determinar a direção afim escala continuada descrito anteriormente. Mas, neste caso não é necessário resolver o problema auxiliar novamente, pois como as equações (2.46) não se altera, a direção $dx = \hat{dx}$. Pelas equações (2.49) e (2.48), temos:

$$\begin{aligned} dz &= X^{-1}(r_s - Zdx), \\ dy &= (ADA^t)^{-1} AD(r_d - dz). \end{aligned} \tag{2.51}$$

Observação: Neste etapa, é necessário resolver somente um sistema linear com a matriz ADA^t para determinar dy.

2.2.3 Somente a componente j bloqueia

Como somente a componente j dever ser nula, então na direção afim escala continuada:

$$\hat{dz}_j = 0. \tag{2.52}$$

Adicionalmente, a direção $\hat{d}$ deve satisfazer aproximadamente o sistema:

$$\begin{cases} A\hat{dx} \simeq r_p & (2.53) \\ A^t \hat{dy} + \hat{dz} \simeq r_d & (2.54) \\ X\hat{dz} + Z\hat{dx} \simeq r_a. & (2.55) \end{cases}$$

Para resolver este sistema, criamos um problema auxiliar para determinar $\hat{dx}$, de modo análogo aos problemas criados anteriormente. Neste caso, utilizamos as equações (2.52), (2.53) e (2.55).

Como $\hat{d}z_j = 0$, por (2.9) segue que

$$\hat{d}x_j = -x_j.$$

Considerando $\bar{d}x = dx$, a última direção calculada, formulamos o problema que determina $\hat{d}x$ por:

$$\begin{array}{ll} \min & \frac{1}{2}\|D^{-\frac{1}{2}}\hat{d}x - D^{-\frac{1}{2}}\bar{d}x\|^2 \\ \text{s.a} & A\hat{d}x = r_p \\ & \hat{d}x_j = \beta_b \end{array} \tag{2.56}$$

em que $D = Z^{-1}X,\ \hat{d}x,\ \bar{d}x \in \mathbb{R}^n, \beta_b = -x_j$.

Utilizaremos a função Lagrangiana para encontrar a solução deste problema, que neste caso é dada por:

$$l(\hat{d}x, v, \gamma) = \frac{1}{2}(D^{-\frac{1}{2}}\hat{d}x - D^{-\frac{1}{2}}\bar{d}x)^t(D^{-\frac{1}{2}}\hat{d}x - D^{-\frac{1}{2}}\bar{d}x) + v^t(A\hat{d}x - r_p) + \gamma(\hat{d}x_j - \beta_b)$$

em que $v = (v_1, v_2, \ldots, v_m)^t \in \mathbb{R}^m$ e $\gamma \in \mathbb{R}$.

Pela condição necessária de primeira ordem, temos $\nabla l(\hat{d}x, v, \gamma) = 0$, ou seja,

$$\begin{cases} D^{-1}\hat{d}x - D^{-1}\bar{d}x + A^t v + \gamma e_j = 0 & (2.57) \\ A\hat{d}x - r_p = 0 & (2.58) \\ \hat{d}x_j - \beta_b = 0, & (2.59) \end{cases}$$

em que e_j é um vetor canônico, $1 \leq j \leq n$.

Por (2.57),

$$\hat{d}x = \bar{d}x - DA^t v - \gamma d_{jj} e_j. \tag{2.60}$$

Por (2.58) e (2.60), visto que a matriz ADA^t é definida positiva, temos que:

$$v = -\gamma d_{jj}(ADA^t)^{-1}A_j. \tag{2.61}$$

Por (2.59) e (2.60),

$$\beta_b = \hat{d}x_j = \bar{d}x_j - d_{jj}A_j^t v - \gamma d_{jj}.$$

Substituindo o valor de v dado em (2.61), segue que

$$\gamma = \frac{\bar{d}x_j - \beta_b}{d_{jj}(1 - d_{jj}A_j^t(ADA^t)^{-1}A_j)}.$$

Portanto, encontramos $\hat{dx}$, um ponto crítico. De modo análogo aos casos anteriores, analisando a matriz Hessiana da função Lagrangiana neste ponto, verificamos que $\hat{dx}$ é a solução mínima do problema auxiliar.

Pelas equações (2.54) e (2.55), encontramos as demais direções:

$$\begin{aligned} \hat{dz} &= X^{-1}(r_a - Z\hat{dx}), \\ \hat{dy} &= (ADA^t)^{-1}AD(r_d - \hat{dz}). \end{aligned} \tag{2.62}$$

Observação: Neste caso, o cálculo da direção afim escala continuada implica na resolução de dois sistemas lineares com ADA^t, sendo um sistema no cálculo de v para encontrar $\hat{dx}$ e um sistema para encontrar $\hat{dy}$.

Calculada a direção afim escala continuada, falta realizar a centragem e a correção não linear para obtermos a direção preditora corretora continuada. Então, fazendo o cálculo de μ e a correção não linear, na direção d temos:

$$dz_j = 0.$$

Adicionalmente, a direção d, deve satisfazer aproximadamente o sistema linear:

$$\begin{cases} Adx \simeq r_p & (2.63) \\ A^t dy + dz \simeq r_d & (2.64) \\ Zdx + Xdz \simeq r_s & (2.65) \end{cases}$$

em que

$$r_s = r_a + \mu e - \hat{D}_x\hat{D}_z e. \tag{2.66}$$

Este sistema é resolvido de forma similar ao sistema para determinar a direção afim escala continuada descrito anteriormente, sendo que no problema auxiliar (2.56) ocorre alteração no valor β_b devido a realização da centragem e da correção não linear.

Como $dz_j = 0$, por (2.29) segue que

$$dx_j = -z_j + \frac{\mu - \hat{dx}_j\hat{dz}_j}{x_j}.$$

Portanto, o problema que deve se utilizado é:

$$\begin{aligned} \min \quad & \tfrac{1}{2}\|D^{-\frac{1}{2}}dx - D^{-\frac{1}{2}}\bar{dx}\|^2 \\ \text{s.a} \quad & Adx = r_p \\ & dx_j = \beta_b \end{aligned} \tag{2.67}$$

em que $D = Z^{-1}X,\ dx,\ \bar{dx} \in \mathbb{R}^n$,

$$\beta_b = -z_j + \frac{\mu - \hat{dx}_j \hat{dz}_j}{x_j}.$$

Este problema auxiliar é resolvido de modo análogo ao (2.56), de forma a obter:

$$dx = \bar{dx} - DA^t v - \gamma d_{jj} e_j \tag{2.68}$$

em que

$$\begin{aligned} v &= -\gamma d_{jj}(ADA^t)^{-1}A_j, \\ \gamma &= \frac{\bar{dx}_j - \beta_b}{d_{jj}(1 - d_{jj}A_j^t(ADA^t)^{-1}A_j)}. \end{aligned}$$

Pelas equações (2.65) e (2.64), temos:

$$\begin{aligned} dz &= X^{-1}(r_a - Zdx), \\ dy &= (ADA^t)^{-1}AD(r_d - dz). \end{aligned} \tag{2.69}$$

Para determinar a direção d acima não é necessário resolver um sistema linear com a matriz ADA^t para obter o valor de v no cálculo de dx. Assim, basta resolver apenas um sistema linear com a matriz ADA^t para determinar dy.

Observação: No cálculo das direções a única diferença entre os casos considerados está no cálculo de $\hat{dx}$ e dx, para as outras direções os cálculos são realizados da mesma forma.

Neste caso, o tamanho dos passos primal e dual possuem o mesmo tamanho. O tamanho de passo é dado por:

$$\alpha = \min\{\alpha_p, \alpha_d\} \tag{2.70}$$

em que α_p e α_d são dados por (1.35). Note que, este é o cálculo originalmente nos métodos de pontos interiores. Este cálculo aqui é adotado porque obtém resultados númericos muito superiores em comparação com α_p e α_d diferentes [31].

Definidas as novas direções, os passos (10) e (11) do Algoritmo 3 - Método Preditor Corretor são substituídos pelo Algoritmo 5.

Algoritmo 5 Iteração Continuada

1. Calcule o tamanho dos passos primal α_p e dual α_d por (1.35)
2. Atualize o ponto

   $x = x + \alpha_p dx$

   $y = y + \alpha_d dy$

   $z = z + \alpha_d dz$
3. Atualize os resíduos por (1.10)
4. Encontre as componentes de bloqueio

   $i = \arg\min_t \{-\frac{x_t}{dx_t} | dx_t < 0\}$ e $\hat{dx}_i = 0$

   $j = \arg\min_t \{-\frac{z_t}{dz_t} | dz_t < 0\}$ e $\hat{dz}_j = 0$
5. Se $i = j$ ou não existem i e j, siga para próxima iteração
6. Caso:

   6.1. i e j existem então calcule $\hat{dx}$ por (2.16)

   6.2. Somente i existe então calcule $\hat{dx}$ por (2.41)

   6.3. Somente j existe então calcule $\hat{dx}$ por(2.60)
7. Calcule $\hat{dy}$ e $\hat{dz}$ por (2.62)
8. Calcule μ por (1.31) e r_s por (1.33)
9. Caso:

   9.1 i e j existem então calcule dx por (2.31)

   9.2 Somente i existe então $dx = \hat{dx}$

   9.3 Somente j existe então calcule dx por (2.68)
10. Calcule dy e dz por (2.69)
11. Calcule o tamanho do passo α por (2.70)
12. Atualize o ponto

    $x = x + \alpha dx$

    $y = y + \alpha dy$

    $z = z + \alpha dz$.

2.3 Iteração Continuada com canalização

A abordagem da iteração continuada também pode ser desenvolvida quando as variáveis são canalizadas. Neste caso, é realizada uma busca pelas componentes de bloqueio nas direções (dx, ds, dz, dw). As componentes de bloqueio i e j nas direções dx e dz, são conforme definidas anteriormente. Assim, as componentes de bloqueio nas direções ds, dw, são definidas por:

$$k = \arg\min_t \{-\frac{s_t}{ds_t} | ds_t < 0\},$$
$$l = \arg\min_t \{-\frac{w_t}{dw_t} | dw_t < 0\}.$$

Com o aumento de duas variáveis o cálculo para determinar as componentes nulas na nova direção é realizado como descrito abaixo.

Sejam $h = \min\{-\frac{x_i}{dx_i}, -\frac{s_k}{ds_k}\}$ e $g = \min\{-\frac{z_j}{dz_j}, -\frac{w_l}{dw_l}\}$. Se $h = -\frac{x_i}{dx_i}$, então na nova direção $\hat{dx}_i = 0$, caso contrário i recebe o valor de k e $\hat{ds}_i = 0$. No caso de g, se $g = -\frac{z_j}{dz_j}$, então na nova direção $\hat{dz}_j = 0$, caso contrário j recebe o valor de l e $\hat{dw}_l = 0$. Como h e g podem não estar definidos em alguma iteração, então pode ocorrer de alguma componente de bloqueio não existir. Temos os seguintes casos a considerar:

1. As componentes i e j bloqueiam;
2. Somente a componente i bloqueia;
3. Somente a componente j bloqueia;
4. Nenhum componente bloqueia.

No último caso, a iteração continuada não é realizada. Assim, analisamos os três primeiros casos:

2.3.1 As componentes i e j bloqueiam

Com as componentes determinadas, calculamos a direção afim escala continuada de modo que:

$$\hat{dx}_i = 0 \quad \text{ou} \quad \hat{ds}_i = 0 \tag{2.71}$$
$$\hat{dz}_j = 0 \quad \text{ou} \quad \hat{dw}_j = 0. \tag{2.72}$$

Adicionalmente, a direção afim escala continuada $\hat{d}$ deve satisfazer aproximadamente o sistema:

$$\begin{cases} A\hat{dx} \simeq r_p & (2.73) \\ \hat{dx} + \hat{ds} \simeq r_u & (2.74) \\ A^t\hat{dy} + \hat{dz} - \hat{dw} \simeq r_d & (2.75) \\ X\hat{dz} + Z\hat{dx} \simeq r_a. & (2.76) \\ S\hat{dw} + W\hat{ds} \simeq r_b & (2.77) \end{cases}$$

em que $\hat{dx}, \hat{ds}, \hat{dz}, \hat{dw} \in \mathbb{R}^n$ e $\hat{dy} \in \mathbb{R}^m$.

De modo análogo a iteração continuada sem canalização, para determinar $\hat{dx}$ utilizamos o problema do tipo (2.10). Neste caso, criamos um problema com (2.71) (2.72) e (2.73). Como existem duas possibilidades para (2.71) e (2.72), os dois casos para determinar o valor de β_a e β_b, que fazem parte das restrições do problema auxiliar, são analisados.

a) Determinando o valor de β_a.

- Se $\hat{dx}_i = 0$ então $\beta_a = 0$;
- Se $\hat{ds}_i = 0$, devemos encontrar o valor da componente $\hat{dx}_i$ correspondente. Por (2.74), temos:

  $$\hat{dx} = r_u - \hat{ds}.$$

  Tomando a i-ésima componente, temos:

  $$\hat{dx}_i = (r_u)_i - \hat{ds}_i.$$

  Como $\hat{ds}_i = 0$, segue que:

  $$\hat{dx}_i = (r_u)_i.$$

  Portanto, $\beta_a = (r_u)_i$.

b) Determinando o valor de β_b.

- Se $\hat{dz}_j = 0$, como já calculado em (2.9) no caso sem variáveis canalizadas, temos $\hat{dx}_j = -x_j$. Portanto, $\beta_b = -x_j$.

- Se $\hat{dw}_j = 0$, encontramos o valor da componente $\hat{dx}_j$ correspondente através das equações (2.74) e (2.77). Por (2.77), temos:

$$\begin{aligned}\hat{ds} &= W^{-1}(r_b - S\hat{dw})\\ \hat{ds} &= W^{-1}(-SWe - S\hat{dw})\\ \hat{ds} &= -Se - W^{-1}S\hat{dw}.\end{aligned}$$

  Tomando a j-ésima componente, temos:

$$\hat{ds}_j = -s_j(1 + \frac{\hat{dw}_j}{w_j}).$$

  Como $\hat{dw}_j = 0$, segue que:

$$\hat{ds}_j = -s_j. \tag{2.78}$$

  Por (2.74):

$$\hat{dx} = r_u - \hat{ds}.$$

  Assim, tomando a j-ésima componente e substituindo o valor de $\hat{ds}_j$ encontrado em (2.78), temos:

$$\hat{dx}_j = (r_u)_j + s_j.$$

Portanto, o problema auxiliar é dado por:

$$\begin{array}{ll}\min & \frac{1}{2}\|D^{-\frac{1}{2}}\hat{dx} - D^{-\frac{1}{2}}\bar{dx}\|^2\\ \text{s.a} & A\hat{dx} = r_p\\ & \hat{dx}_i = \beta_a\\ & \hat{dx}_j = \beta_b\end{array} \tag{2.79}$$

em que $D = Z^{-1}X + W^{-1}S,\ \hat{dx},\ \bar{dx} \in \mathbb{R}^n$,

$$\beta_a = \begin{cases} 0, \text{ se } \hat{dx}_i = 0;\\ (r_u)_i, \text{ se } \hat{ds}_i = 0.\end{cases} \tag{2.80}$$

e

$$\beta_b = \begin{cases} -x_j, \text{ se } \hat{dz}_j = 0;\\ (r_u)_j + s_j, \text{ se } \hat{dw}_j = 0.\end{cases} \tag{2.81}$$

O problema (2.79) é similar a (2.10), sendo alterado apenas os valores β_a e β_b. Logo,

$$\hat{d}x = \bar{d}x - DA^t v - \alpha d_{ii} e_i - \gamma d_{jj} e_j \tag{2.82}$$

em que

$$v = -\alpha d_{ii}(ADA^t)^{-1}A_i - \gamma d_{jj}(ADA^t)^{-1}A_j,$$

$$\alpha = \frac{\bar{d}x_i - \beta_a + \gamma d_{ii} d_{jj} A_i^t (ADA^t)^{-1} A_j}{d_{ii}(1 - d_{ii} A_i^t (ADA^t)^{-1} A_i)},$$

$$\gamma = \frac{\bar{d}x_j - \beta_b + \left[\frac{(\bar{d}x_i - \beta_a) d_{jj} \theta_{ij}}{1 - dii\theta_i}\right]}{d_{jj}\xi}.$$

Pelas equações (2.74),(2.75),(2.76) e (2.77), respectivamente, temos as outras direções dadas por:

$$\begin{aligned} \hat{d}s &= r_u - \hat{d}x \\ \hat{d}z &= X^{-1}(r_a - Z\hat{d}x) \\ \hat{d}w &= S^{-1}(r_b - W\hat{d}s) \\ \hat{d}y &= (ADA^t)^{-1}AD(r_d - \hat{d}z + \hat{d}w). \end{aligned} \tag{2.83}$$

Novamente, como (2.75) é um sistema sobredeterminado utilizamos quadrados mínimos para resolvê-lo.

Fazendo o cálculo de μ e da correção não linear na direção preditora corretora continuada devemos ter:

$$dx_i = 0 \quad \text{ou} \quad ds_i = 0 \tag{2.84}$$

$$dz_j = 0 \quad \text{ou} \quad dw_j = 0, \tag{2.85}$$

Adicionalmente, a direção preditora corretora continuada d deve satisfazer aproximadamente o seguinte sistema linear:

$$\begin{cases} Adx \simeq r_p & (2.86) \\ dx + ds \simeq r_u & (2.87) \\ A^t dy + dz \simeq r_d & (2.88) \\ Xdz + Zdx \simeq r_s & (2.89) \\ Sdw + Wds \simeq r_t & (2.90) \end{cases}$$

em que

$$\begin{aligned} r_s &= r_a + \mu e - \hat{D}_x \hat{D}_z e \\ r_t &= r_b + \mu e - \hat{D}_s \hat{D}_w e. \end{aligned} \tag{2.91}$$

A resolução deste sistema é feita de forma similar ao sistema para determinar a direção afim escala continuada descrito anteriormente, sendo que no problema auxiliar (2.79) ocorre alteração no valor de β_b.

- Se $dz_j = 0$, encontramos o valor da componente dx_j correspondente pelas equações (2.87) e (2.90). Por (2.29), temos:

$$dx_j = -x_j + \frac{\mu - \hat{dx}_j \hat{dz}_j}{z_j}.$$

- Se $dw_j = 0$, encontramos o valor da componente dx_j correspondente através das equações (2.87) e (2.90). Então por (2.90), temos:

$$\begin{aligned} ds &= W^{-1}(r_t - Sdw) \\ ds &= W^{-1}(r_b + \mu e - \hat{D}_s \hat{D}_w e - Sdw) \\ ds &= W^{-1}(-SWe + \mu e - \hat{D}_s \hat{D}_w e - Sdw) \\ ds &= W^{-1}(-SWe - Sdw) + W^{-1}(\mu e - \hat{D}_s \hat{D}_w e) \\ ds &= -Se - W^{-1}Sdw + W^{-1}(\mu e - \hat{D}_s \hat{D}_w e). \end{aligned}$$

  Tomando a j-ésima componente, temos:

$$ds_j = -s_j(1 + \frac{dw_j}{w_j}) + \frac{\mu - \hat{ds}_j \hat{dw}_j}{w_j}.$$

  Como $dw_j = 0$, segue que

$$ds_j = -s_j + \frac{\mu - \hat{ds}_j \hat{dw}_j}{w_j}.$$

  Por (2.87):

$$dx = r_u - ds.$$

  Assim, tomando a j-ésima componente e substituindo o valor de ds_j encontrado, temos:

$$dx_j = (r_u)_j + s_j - \frac{\mu - \hat{ds}_j \hat{dw}_j}{w_j}.$$

Portanto, o problema auxiliar é dado por:

$$\begin{array}{ll} \min & \frac{1}{2}\|D^{-\frac{1}{2}}dx - D^{-\frac{1}{2}}\bar{d}x\|^2 \\ \text{s.a} & Adx = r_p \\ & dx_i = \beta_a \\ & dx_j = \beta_b \end{array} \tag{2.92}$$

em que $D = Z^{-1}X + W^{-1}S,\ dx,\ \bar{d}x \in \mathbb{R}^n$,

$$\beta_a = \begin{cases} 0, \text{ se } dx_i = 0; \\ (r_u)_i, \text{ se } ds_i = 0. \end{cases} \tag{2.93}$$

e

$$\beta_b = \begin{cases} -x_j + \frac{\mu - \hat{dx}_j \hat{dz}_j}{z_j}, \text{ se } dz_j = 0; \\ (r_u)_j + s_j - \frac{\mu - \hat{ds}_j \hat{dw}_j}{w_j}, \text{ se } dw_j = 0. \end{cases} \tag{2.94}$$

Como este problema é similar a (2.79), segue que:

$$dx = \bar{d}x - DA^t v - \alpha d_{ii} e_i - \gamma d_{jj} e_j \tag{2.95}$$

em que

$$v = -\alpha d_{ii}(ADA^t)^{-1}A_i - \gamma d_{jj}(ADA^t)^{-1}A_j,$$

$$\alpha = \frac{\bar{d}x_i - \beta_a + \gamma d_{ii} d_{jj} A_i^t (ADA^t)^{-1} A_j}{d_{ii}(1 - d_{ii} A_i^t (ADA^t)^{-1} A_i)},$$

$$\gamma = \frac{\bar{d}x_j - \beta_b + \left[\frac{(\bar{d}x_i - \beta_a) d_{jj} \theta_{ij}}{1 - dii\theta_i}\right]}{d_{jj}\xi}.$$

Pelas equações (2.87)-(2.90), temos as outras direções, dadas por:

$$\begin{aligned} ds &= r_u - dx \\ dz &= X^{-1}(r_a - Zdx) \\ dw &= S^{-1}(r_b - Wds) \\ dy &= (ADA^t)^{-1}AD(r_d - dz + dw). \end{aligned} \tag{2.96}$$

Utilizando novamente quadrados mínimos para determinar dy.

2.3.2 Somente a componente i bloqueia

Como somente a componente i bloqueia, então na direção afim escala continuada:

$$\hat{d}x_i = 0 \quad \text{ou} \quad \hat{d}s_i = 0. \tag{2.97}$$

Adicionalmente, esta direção deve satisfazer aproximadamente o sistema:

$$\begin{cases} A\hat{d}x \simeq r_p \\ \hat{d}x + \hat{d}s \simeq r_u \\ A^t\hat{d}y + \hat{d}z - \hat{d}w \simeq r_d \\ X\hat{d}z + Z\hat{d}x \simeq r_a \\ S\hat{d}w + W\hat{d}s \simeq r_b. \end{cases}$$

Observe que este sistema é similar ao do caso (2) para as variáveis não canalizadas na subseção (2.2.2), podendo sofrer alteração somente quando $ds_i = 0$. Então, a única mudança que pode acontecer em relação ao cálculo da direção preditora corretora continuada do caso (2) quando as variáveis não são canalizadas é no valor de β_a. Dessa forma, no problema auxiliar (2.37) β_a é dado por (2.80).

2.3.3 Somente a componente j bloqueia

Neste caso, a direção afim escala continuada é calculada de modo a obter:

$$\hat{d}z_j = 0 \quad \text{ou} \quad \hat{d}w_j = 0. \tag{2.98}$$

Adicionalmente, a direção $\hat{d}$ deve satisfazer aproximadamente o sistema:

$$\begin{cases} A\hat{d}x \simeq r_p \\ \hat{d}x + \hat{d}s \simeq r_u \\ A^t\hat{d}y + \hat{d}z - \hat{d}w \simeq r_d \\ X\hat{d}z + Z\hat{d}x \simeq r_a \\ S\hat{d}w + W\hat{d}s \simeq r_b. \end{cases}$$

Este sistema é similar ao do caso (3) na subseção (2.2.3), quando as variáveis não são canalizadas, podendo sofrer alteração somente em $dw_j = 0$. Então, a única mudança que pode acontecer em relação ao cálculo da direção preditora corretora continuada do caso (3) para as variáveis não canalizadas é no valor de β_b, dos problemas auxiliares. Sendo que em (2.56) β_b é dado por (2.81) e em (2.67) dado por (2.94).

O tamanho de passo α neste caso, é dado por:

$$\alpha = \min\{\alpha_p, \alpha_d\} \tag{2.99}$$

em que α_p e α_d são dados por (1.40).

Observações:

- A única diferença nos três casos analisados está no cálculo das direções $\hat{dx}$ e dx.
- Como no problema sem váriaveis canalizadas, basta resolver a mesma quantidade de sistemas adicionais, realizando o mesmo esforço computacional.

Nos passos(10) e (11) do Algoritmo 4 é inserido o Algoritmo 6.

Algoritmo 6 Iteração continuada com variáveis canalizadas

1. Calcule o tamanho dos passos primal α_p e dual α_d por (1.40)
2. Atualize o ponto

   $x = x + \alpha_p dx$

   $s = s + \alpha_p ds$

   $y = y + \alpha_d dy$

   $z = z + \alpha_d dz$

   $w = w + \alpha_d dw$
3. Atualize os resíduos por (1.19) e calcule γ por (1.7)
4. Calcule $h = \min(\rho_x, \rho_s, \rho_z, \rho_w)$
5. Se $h = \rho_x$ então $i = \arg\min_t \{-\frac{x_t}{dx_t} | dx_t < 0\}$ e $\hat{d}x_i = 0$
6. Se $h = \rho_s$ então $i = \arg\min_t \{-\frac{s_t}{ds_t} | ds_t < 0\}$ e $\hat{d}s_i = 0$
7. Se $h = \rho_w$ então $j = \arg\min_t \{-\frac{w_t}{dw_t} | dw_t < 0\}$ e $\hat{d}w_j = 0$
8. Se $h = \rho_z$ então $j = \arg\min_t \{-\frac{z_t}{dz_t} | dz_t < 0\}$ e $\hat{d}z_j = 0$
9. Se $i = j$ ou não existem i e j então siga para a próxima iteração
10. Caso:

    10.1 i e j existem então calcule $\hat{d}x$ por (2.82)

    10.2. Somente i existe então calcule $\hat{d}x$ por (2.41) conforme β_a dado por (2.80)

    10.3. Somente j existe então calcule $\hat{d}x$ por (2.60) conforme β_b dado por (2.81)
11. Calcule $\hat{d}s, \hat{d}y, \hat{d}z$ e $\hat{d}w$ por (2.83)
12. Calcule γ por (1.7), μ por (1.37) e atualize os resíduos por (2.91)
13. Caso

    13.1 i e j existem então calcule dx por (2.95)

    13.2 Somente i existe então $dx = \hat{d}x$

    13.3 Somente j existe então calcule dx por (2.68) conforme β_b dado por (2.81)
14. Calcule ds, dy, dz e dw por (2.96)
15. Calcule o tamanho do passo α por (2.99)
16. Atualize o ponto

    $x = x + \alpha dx$

    $s = s + \alpha ds$

    $y = y + \alpha dy$

    $z = z + \alpha dz$

    $w = w + \alpha dw$.

Portanto, com a aplicação da iteração continuada no método preditor corretor, um novo método é obtido, de modo a iteragir em dois níveis. No nível externo, é calculada uma fatoração de Cholesky para ADA^t para o cálculo da direção preditora corretora. No nível interno, são realizadas atualizações da direção através da eliminação de algumas de suas componentes, de forma que a fatoração já existente do nível externo seja utilizada nos sistemas lineares envolvidos.

Capítulo 3

Experimentos Computacionais

Experimentos preliminares em Matlab foram realizados com os métodos primal dual e preditor corretor e a incorporação da iteração continuada. Os resultados computacionais mais promissores foram obtidos com o método preditor corretor, então o estudo foi focado neste método. Neste capítulo, apresentamos os experimentos e os resultados computacionais obtidos ao aplicar a iteração continuada no método de pontos interiores preditor corretor e comparamos a abordagem proposta com o método original.

3.1 Implementação

Os experimentos computacionais foram realizados em um processador Intel Core i5, 6 GB RAM, 2.3GHZ, 750GB HD e Sistema Operacional Linux.

A iteração continuada foi implementada em linguaguem C e incorporada ao código do PCx [6] com as múltiplas correções desligada [17].

3.2 Problemas Testes

Para analisar o desempenho do PCx com a incorporação da abordagem proposta, os experimentos computacionais foram realizados com problemas de livre acesso na internet pertencentes as coleções Netlib [13], Qaplib [3], Kennington [4] e outros problemas gentilmente cedidos por Gonçalves [16]. Somente problemas cujo número de colunas é maior que cinco mil foram considerados. A Tabela 3.1 mostra na primeira coluna o nome dos problemas se-

lecionados, na segunda coluna a dimensão dos problemas (número de linhas e colunas), após o preprocessamento do PCx. Na terceira coluna, temos a densidade do fator de Cholesky, fornecida pelo PCx que realiza um preprocessamento para que este fator seja esparso, e a última coluna indica a coleção que o problema pertence.

Problema	Dimensão		Densidade	Coleção
	Linhas	Colunas		
80BAU3B	2140	11066	0,01760	Netlib
DFL001	5984	12143	0,09132	Netlib
FIT2D	25	10524	0,99680	Netlib
FIT2P	3000	13525	0,00033	Netlib
MAROS-R7	2152	7440	0,23023	Netlib
PILOT87	1971	6373	0,21863	Netlib
STOCFOR3	15362	22228	0,00144	Netlib
WOODW	708	5364	0,11839	Netlib
CRE-A	2994	6692	0,00708	Kennington
CRE-B	5336	36382	0,01728	Kennington
CRE-C	2375	5421	0,00969	Kennington
CRE-D	4102	28601	0,02497	Kennington
KEN11	10085	16740	0,00192	Kennington
KEN13	22534	36561	0,0113	Kennington
KEN18	78862	128434	0,00061	Kennington
OSA-07	1081	25030	0,04747	Kennington
OSA-14	2300	54760	0,02255	Kennington
OSA-30	4313	104337	0,01214	Kennington
OSA-60	10243	243209	0,00497	Kennington
PDS06	9156	28472	0,01395	Kennington
PDS10	15648	48780	0,01372	Kennington
PDS20	38722	106180	0,01357	Kennington
BL	5729	12462	0,01190	Gonçalves
BL2	5729	12462	0,01190	Gonçalves
CO5	4849	10787	0,01374	Gonçalves
CO9	9090	19997	0,01106	Gonçalves

CQ9	8004	19317	0,01222	Gonçalves
EX05	825	7797	0,21791	Gonçalves
EX06	832	7895	0,22610	Gonçalves
EX09	1821	18184	0,08611	Gonçalves
GE	9150	14990	0,00660	Gonçalves
NL	6665	14680	0,01223	Gonçalves
CHR25A	8149	15325	0,07689	Qaplib
CHR22B	5587	10417	0,08850	Qaplib
ELS19	4350	13186	0,40664	Qaplib
KRA30A	18059	85725	0,51777	Qaplib
KRA30B	18059	85725	0,51777	Qaplib
ROU20	7359	37640	0,76870	Qaplib
SCR15	2234	6210	0,53284	Qaplib
SCR20	5079	15980	0,51710	Qaplib
STE36A	27683	131076	0,46092	Qaplib
STE36B	27683	131076	0,46092	Qaplib
STE36C	27683	131076	0,46092	Qaplib

Tabela 3.1: Problemas Testes

Os problemas Qap são modelos para problemas quadráticos linearizados. Um gerador para este tipo de problema pode ser obtido da Netlib. Outra fonte para este tipo de problema pode ser a Qaplib.

3.3 Resultados Computacionais

No primeiro experimento realizado, a iteração continuada foi potencialmente acionada em cada iteração do método de pontos interiores. Neste experimento para todos os problemas testes selecionados o método não convergiu, uma vez que quando a sequência está próximo a solução, algumas variáveis apresentam valores muito pequeno causando instabilidade numérica.

A Tabela 3.2 mostra os resultados obtidos dos experimentos quando a iteração continuada é acionada até as 10 primeiras iterações do método de pontos interiores. Fixamos este

intervalo, devido que em alguns experimentos realizados o método não convergiu para vários problemas teste quando a iteração continuada é acionada nas últimas iterações. Além disso, pelo PCx temos problemas testes sendo resolvidos em 16 iterações. Na primeira coluna da Tabela 3.2, temos os problemas testes selecionados, na segunda os resultados obtidos pelo PCx sem a iteração continuada e na terceira coluna, os resultados obtidos pelo PCx com a iteração continuada, o qual será referido como PCx-IC. As colunas (I)-(IV) do PCx-IC mostram os resultados em que a iteração continuada potencialmente é realizada nas seguintes iterações:

(I) 1 a 5;

(II) 5 a 10;

(III)1 a 10;

(IV) 6 a 8,

Como em PCx-IC(II) os resultados obtidos foram melhores entre os três primeiros, o intervalo 5 a 10 foi refinado sendo que a iteração continuada aplicada nas iterações de 6 a 8 obteve melhor desempenho.

Nas colunas de PCx e de PCx-IC (I)-(IV), encontram-se o número total de iterações (k) realizadas pelos métodos de pontos interiores e o tempo em segundos (t) em que cada método convergiu.

Problema	PCx		PCx-IC							
			I		II		III		IV	
	k	t	k	t	k	t	k	t	k	t
80BAU3B	36	0,25	*	*	42	0,28	*	*	42	0,28
DFL001	56	45,40	49	40,09	55	44,90	49	40,09	52	42,65
FIT2D	22	0,26	21	0,24	21	0,24	20	0,25	21	0,24
FIT2P	20	0,21	18	0,18	19	0,20	19	0,20	20	0,21
MAROS-R7	16	1,68	16	1,68	16	1,68	17	1,86	16	1,68
PILOT87	33	3,85	34	4,03	32	3,79	35	4,20	32	3,79
STOCFOR3	30	0,50	30	0,50	30	0,52	28	0,42	31	0,53
WOODW	30	0,11	31	0,12	30	0,11	32	0,13	32	0,13
CRE-A	23	0,11	24	0,13	24	0,13	31	0,17	22	0,10
CRE-B	41	1,80	42	1,88	42	1,88	43	2,17	41	1,80
CRE-C	25	0,09	26	0,12	26	0,12	25	0,09	25	0,11
CRE-D	41	1,55	43	1,64	41	1,58	45	1,73	40	1,51
KEN11	20	0,29	24	0,40	21	0,37	25	0,44	20	0,29
KEN13	23	0,89	28	1,26	26	1,17	*	*	24	0,95
KEN18	29	7,32	35	8,70	32	8,14	40	10,53	28	7,19
OSA-07	22	0,28	22	0,28	23	0,32	28	0,39	21	0,26
OSA-14	25	0,80	21	0,72	25	0,80	24	0,91	22	0,79
OSA-30	24	1,82	24	1,82	23	1,77	31	2,34	21	1,68
OSA-60	33	6,98	22	5,83	21	5,62	34	7,28	21	5,62
PDS06	34	5,35	37	5,80	40	6,20	42	6,53	36	5,72
PDS10	39	28,73	43	31,58	44	31,92	48	34,55	40	29,63
PDS20	55	315,66	56	321,83	57	325,94	61	358,81	53	310,31
BL	32	0,85	38	1,15	33	0,87	40	1,23	31	0,83
BL2	37	0,95	45	1,16	35	0,94	47	1,39	34	0,90
CO5	47	0,92	*	*	47	0,92	*	*	49	0,97
CO9	*	*	*	*	*	*	*	*	52	3,33
CQ9	49	2,59	50	2,77	50	2,86	49	2,59	49	2,59
EX05	32	0,37	32	0,37	31	0,43	35	0,49	30	0,35

EX06	63	0,65	65	0,68	61	0,65	62	0,66	58	0,60
EX09	37	0,84	53	1,19	37	1,00	58	1,29	35	0,81
GE	43	1,25	44	1,35	44	1,35	44	1,35	43	1,25
NL	37	1,53	43	1,78	36	1,62	43	1,78	36	1,62
CHR25A	28	22,44	27	22,05	25	20,40	27	22,05	25	20,40
CHR22B	27	8,96	29	9,63	28	9,52	30	10,20	29	9,77
ELS19	26	83,38	27	87,35	25	80,67	27	87,35	26	83,32
KRA30A	26	8997,46	30	12122,78	32	11007,51	28	9212,05	28	11395,82
KRA30B	27	9360,69	29	9824,10	32	10557,81	31	10238,02	26	9105,47
ROU20	16	708,13	17	867,12	20	1054,32	22	1156,12	20	1050,73
SCR15	21	13,73	21	13,74	21	13,73	21	13,74	21	13,73
SCR20	20	155,41	20	158,79	19	155,57	20	157,60	19	143,39
STE36A	32	30985,52	30	28913,63	31	29906,32	31	30802,74	31	29980,73
STE36B	31	29817,93	31	29863,78	31	29763,09	31	35772,23	31	29911,80
STE36C	31	31469,55	30	28931,16	32	31399,48	31	31872,39	30	28917,35

Tabela 3.2: Comparação entre PCx e PCx-IC.

* significa que o método falhou.

Observação: O método falhou, devido a instabilidade numérica.

O desempenho ruim nas primeiras iterações está relacionado ao fato que estamos fixando em zero potencialmente duas componentes da direção, assim a trajetória da direção é muito afetada, de tal forma que em alguns casos o método não recupera-se, ocorrendo sua falha.

Na Tabela 3.2, os resultados mostram que pelo método PCx-IC(I) os problemas 80BAU3B, CO5, CO9, pelo PCx-IC(II) o problema CO9 e pelo PCx-IC(III) os problemas 80BAU3B, KEN13, CO5 e CO9 não são resolvidos. O problema CO9 é resolvido somente pelo PCx-IC(IV). O número de iterações é reduzido em 8 problemas pelo PCx-IC(I), 14 problemas pelo PCx-IC(II), 8 problemas pelo PCx-IC(III) e 22 problemas pelo PCx-IC(IV). A maior redução de iterações ocorre no problema OSA-60 tanto pelo PCX-IC(II) quanto pelo PCx-IC(IV), sendo reduzido 12 iterações. Portanto, os resultados mostram que o método PCx-IC(IV) apresenta melhores resultados. Com a redução do número de iterações o tempo computacional necessário para determinar a solução da maioria dos problemas foi reduzido.

Na Tabela 3.3, temos um resumo em porcentagem em relação ao número de iterações

dos resultados da Tabela 3.2. Observamos que todos os problemas são resolvidos pelo PCx-IC(IV), ao contrário do PCx original que não resolve um problema teste.

PCx-IC	Redução	Aumento	Falha
(I)	18,60%	53,49%	0,7%
(II)	32,55%	44,18%	0,23%
(III)	18,68%	58,14%	0,93%
(IV)	51,16%	23,25%	0%

Tabela 3.3: Síntese dos resultados obtidos com PCx-IC

Para analisar o desempenho da abordagem proposta utilizaremos, neste trabalho, a ferramenta perfil de desempenho, descrita com detalhes no apêndice.

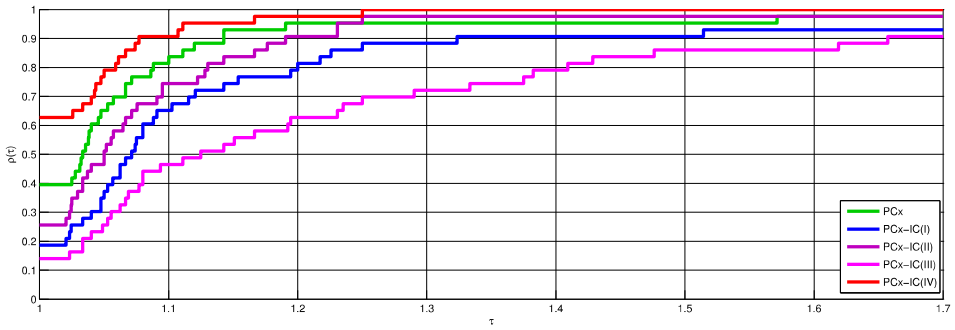

Figura 3.1: Perfil de desempenho dos métodos em relação ao número total de iterações

A Figura 3.1 apresenta o perfil de desempenho dos métodos PCx, PCx-IC(I) a (IV), utilizando como medida de desempenho o total de iterações de pontos interiores. Observamos no gráfico que o método PCx-IC(IV) é o que apresenta a maior probabilidade de resolver os problemas testes em menor tempo, sendo aproximadamente 61%. Além disso, é o único método que resolver todos os problemas testes, pois é o único a atingir $\rho(\tau) = 1$. Portanto, o método PCx-IC(IV) é o mais eficiente e robusto.

A Figura 3.2 apresenta o perfil de desempenho dos métodos PCx, PCx-IC(I) a (IV), utilizando como medida de desempenho o tempo total de resolução dos problemas. Neste gráfico, observamos que o método PCx-IC(IV) é o mais eficiente, em relação a robustez os métodos PCx e PCx-IC(IV) são os mais robustos, apresentando uma diferença muito pequena entre eles.

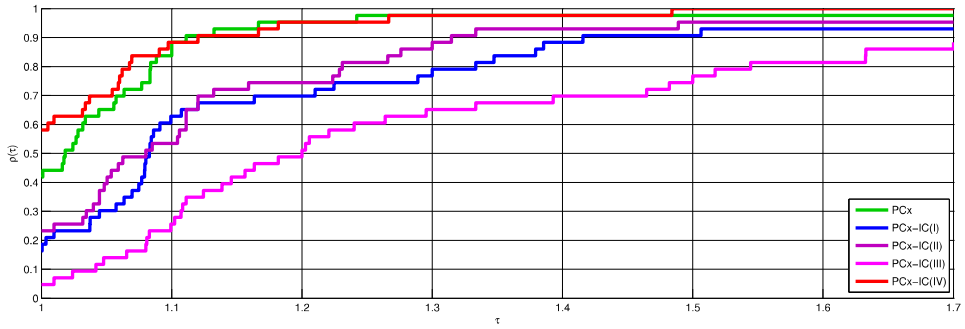

Figura 3.2: Perfil de desempenho dos métodos em relação ao tempo total de resolução dos problemas

Portanto, o método PCx-IC(IV) possui o melhor perfil de desempenho em ambas análises realizadas.

Capítulo 4

Conclusão e Trabalhos Futuros

Neste trabalho, a iteração continuada foi desenvolvida para o método de pontos interiores preditor corretor com e sem variáveis canalizadas. O objetivo principal consiste em reduzir o número total de iterações do método de pontos interiores preditor corretor, e consequentemente, o tempo total de resolução. Nesta iteração é determinada uma direção preditora corretora continuada sendo considerado que algumas de suas componentes sejam nulas, de acordo com as componentes de bloqueio na última direção calculada. Para encontrar a direção preditora corretora continuada minimizamos a alteração realizada ao fixar como nulas algumas componentes em relação à última direção determinada. A fatoração de Cholesky já calculada na iteração do método preditor corretor é utilizada novamente, para resolver os sistemas lineares envolvidos. Desta forma, evita-se o cálculo de uma nova fatoração de Cholesky. Assim, o novo método é realizado em dois níveis. No nível externo é realizado o cálculo da fatoração de Cholesky e o cálculo da direção preditora corretora tradicional. No nível interno é aplicada a iteração continuada para calcular uma nova direção utilizando a fatoração de Cholesky existente do nível externo. O esforço computacional de cada iteração continuada é dominado pela solução de no máximo quatro sistemas lineares com uma mesma matriz já fatorada.

Os resultados computacionais mostraram que quando a iteração continuada é aplicada nas iterações seis a oito o número total de iterações e o tempo total de resolução foram reduzidos na maioria dos problemas, em torno de 51%. Além disso, um problema que não era resolvido pelo PCx sem a incorporação da iteração continuada, utilizando a abordagem proposta passou a ser resolvido.

Portanto, o novo método de pontos interiores preditor corretor com iteração continuada aplicada na sexta a oitava iteração PCx-IC(IV), obteve melhor desempenho que o método original PCx.

Como trabalhos futuros a serem desenvolvidos, pretende-se buscar um critério de aceitação para a direção preditora corretora continuada calculada e projeções alternativas que podem determinar uma nova direção continuada. Além disso, realizar experimentos com a iteração continuada e outras abordagens incorporadas juntas ao método de pontos interiores preditor corretor, de modo que possa resultar em melhorias para este método.

Apêndice A

Perfil de Desempenho

O perfil de desempenho é uma ferramenta para avaliar e comparar o desempenho dos métodos em otimização, introduzido por Dolan e Moré [10]. Considerando que temos n_s métodos e n_p problemas, podemos comparar os métodos em medidas diferentes de desempenho, por exemplo, tempo de processamento, números de avaliações da função, entre outros. Considere S o conjunto de métodos e P conjunto de problemas. A taxa de desempenho é dado por:

$$r_{p,s} = \frac{t_{p,s}}{\min\{t_{p,s} : s \in S\}},$$

em que $t_{p,s}$ é o tempo computacional necessário para resolver o problema p pelo algoritmo s. Considerando o parâmetro r_m fixado, $r_m \geq r_{p,s}$ para um p escolhido. Temos que, $r_m = r_{p,s}$ se s não resolve o problema p.

A função de distribuição acumulativa ρ_s para a taxa de desempenho $r_{p,s}$, é dada por:

$$\rho_s : \mathbb{R} \longrightarrow [0, 1]$$
$$\rho_s(\tau) = \tfrac{1}{n_p} \mid B \mid,$$

em que $\mid B \mid$ é a cardinalidade de $B = \{p \in P | r_{p,s} < \tau\}$. Esta função é não decrescente e constante por partes. O valor de $\rho_s(1)$ representa a probalidade de problemas resolvidos em menor tempo pelo algoritmo s. Assim, se estamos interessados no método mais eficiente, necessitamos somente comparar os valores de $\rho_s(1)$ com todos os métodos. Agora, considerando que $r_{p,s} \in [1, r_m]$, sendo que $r_{p,s} = r_M$ somente quando algum problema p não é resolvido pelo método s. Tomando $\rho_s(r_M) = 1$, a probabilidade que o método resolve um problema é dado por:

$$\rho_s^* = \lim_{\tau \to r_M^-} \rho_s(\tau).$$

Logo, se estamos interessados somente no método mais robusto, necessitamos comparar os valores de ρ_s^* para com todos os algoritmos e escolher o de maior valor.

Referências Bibliográficas

[1] I. Adler, M. G. C. Resende, G. Veiga, and N. Karmarkar, *An implementation of Karmarkar's algorithm for linear programming*, Mathematical Programming, 44 (1989), pp. 297–335.

[2] D. P. Bertsekas, *Nonlinear Programming Second Edition*, Athena Scientific, Beltmont, Massachusetts, 1999.

[3] R. E. Burkard, S. E. Karish, and F. Rendl, *QAPLIB - A quadratic assignment problem library*, European Journal of Operational Research, 55 (1991), pp. 115–119.

[4] W. J. Carolan, J. E. Hill, J. L. Kennington, S. Niemi, and S. J. Wichman, *An empirical evaluation of the KORBX algoritmos for military airlift applications*, Operations Research, 28 (1990), pp. 240–248.

[5] V. V. Chvátal, *Linear Programming*, W. H. Freeman and Company, New York, USA, 1983.

[6] J. Czyzyk, S. Mehrotra, M. Wagner, and S. J. Wright, *PCx an interior point code for linear programming*, Optimization Methods & Software, 11-2 (1999), pp. 397–430.

[7] G. B. Dantzig, *Linear Programming and Extensions*, Princeton University Press, New-Jersey, USA, 1998.

[8] J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, PA, 1996.

[9] I. I. Dikin, *Iterative solution of problems of linear and quadratic programming*, Soviets Math. Doklady, 8 (1967), pp. 674–675.

[10] E. D. Dolan and J. J. Moré, *Benchmarking optimization software with performance profiles*, Math. Program., Ser. A, 91 (2002), pp. 201–213.

[11] I. S. Duff, A. M. Erisman, and J. K. Reid, *Direct Methods for Sparse Matrices*, Clarendon Press, Oxford, 1986.

[12] A. S. El-Bakry, R. A. Tapia, and Y. Zhang, *A study of indicators for identifying zero variables in interior-point methods*, SIAM Rev., 36 (1994), pp. 45–72.

[13] D. M. Gay, *Eletronic mail distribution of linear programming test probems*, Mathematical Programming Society Committee on Algorithms COAL Newsletter, (1985), pp. 10–12.

[14] P. E. Gill, W. Murray, M. A. Saunders, J. A. Tomlin, and M. H. Wright, *On the projected newton barrier methods for linear programming and equivalence to Karmarkar's projective method*, Mathematical Programming 36, (1986), pp. 183–209.

[15] G. H. Golub and C. F. Van Loan, *Matrix Computations Third Edition*, The Johns Hopkins University Press, Baltimore, Maryland, 1996.

[16] J. P. M. Gonçalves, R. H. Storer, and J. Gondzio, *A family of linear programming algorithms based on an algorithm by von Neumann*, Optimization Methods and Software, 24 (2009), pp. 461–478.

[17] J. Gondzio, *Multiple centrality corrections in a primal-dual method for linear programming*, Computational Optimization and Applications, 6 (1996), pp. 137–156.

[18] J. Gondzio, *Interior point methods 25 years later*, European Journal of Operational Research, 218 (2012), pp. 587–601.

[19] C. C. Gonzaga, *Path-following methods for linear programming*, SIAM Review, 34 (1992), pp. 167–224.

[20] N. Karmarkar, *A new polynomial-time algorithm for linear programming*, Combinatorica, 4 (1984), pp. 373–395.

[21] L. G. Khachiyan, *Polynomial algorithms in linear programming*, Soviet Mathematics Doklady, 20 (1979), pp. 191–194.

[22] M. Kojima, N. Megiddo, and S. Mizuno, *A primal - dual infeasible interior point algorithm for linear programming*, Mathematical Programming, 61 (1993), pp. 263–280.

[23] M. Kojima, S. Mizuno, and A. Yoshise, *A primal-dual interior point algorithm for linear programming*, Progress in Mathematical Programming:Interior- Point Algorithms and Related Methods, N. Megiddo, (1989), pp. 29–47.

[24] D. G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley, Reading, 1984.

[25] R. E. Marsten, R. Subramanian, M. J. Saltzman, I. J. Lustig, and D. F. Shanno, *Interior point methods for linear programming: just call Newton, Lagrange and Fiacco and McCornick!*, Interfaces, 20 (1990), pp. 105–116.

[26] N. Megiddo, *Pathways to the optimal set in linear programming, in Progress in Mathematical: Interior Point Algorithm and Related Methods, N. Megiddo*, ed. Springer Verlag, Berlin, Germany, 1989.

[27] S. Mehrotra, *On the implementation of a primal-dual interior point method*, SIAM Journal on Optimization, 2 (1992), pp. 575–601.

[28] R. D. C. Monteiro, I. Adler, and M. G. C. Resende, *A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension*, Mathematics of Operations Research, 15 (1990), pp. 191–214.

[29] A. R. L. Oliveira, C. Lyra, and P. B. Correia, *Implementação computacional de algoritmo polinomial de programação linear: aplicação ao planejamento de operação de sistemas hidrotérmicos*, Anais do VII Congresso Brasileiro de Automática - CBA, São José dos Campos, SP, (1988), pp. 928–929.

[30] R. S. Souza, C. T. L. Silva, and M. N. Arenales, *Métodos do tipo dual simplex para problemas de otimização linear canalizados*, Pesquisa Operacional, 25 (2005), pp. 349–382.

[31] R. J. Vanderbei, *Linear Programming: Foundations and Extensions, Third Edition*, Springer, Princeton University, New Jersey,USA, 2008.

[32] S. J. Wright, *Primal–Dual Interior–Point Methods*, SIAM Publications, SIAM, Philadelphia, PA, USA, 1997.