



UNIVERSIDADE ESTADUAL DE  
CAMPINAS

Instituto de Matemática, Estatística e  
Computação Científica

LILIAN FERREIRA BERTI

# **Redução de Iterações dos Métodos de Pontos Interiores com Iteração Continuada**

Campinas

2016

Lilian Ferreira Berti

## **Redução de Iterações dos Métodos de Pontos Interiores com Iteração Continuada**

Tese apresentada ao Instituto de Matemática,  
Estatística e Computação Científica da Uni-  
versidade Estadual de Campinas como parte  
dos requisitos exigidos para a obtenção do  
título de Doutora em Matemática Aplicada.

Orientador: Aurelio Ribeiro Leite de Oliveira

Coorientadora: Carla Tavianne Lucke da Silva Ghidini

O arquivo digital corresponde à ver-  
são final da Tese defendida pela aluna  
Lilian Ferreira Berti e orientada pelo  
Prof. Dr. Aurelio Ribeiro Leite de Oli-  
veira.

Campinas

2016

**Agência(s) de fomento e nº(s) de processo(s):** FAPESP, 2011/20623-7

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Matemática, Estatística e Computação Científica  
Ana Regina Machado - CRB 8/5467

B462r Berti, Lilian Ferreira, 1988-  
Redução de iterações dos métodos de pontos interiores com iteração  
continuada / Lilian Ferreira Berti. – Campinas, SP : [s.n.], 2016.

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

1. Métodos de pontos interiores. 2. Programação linear. 3. Métodos  
numéricos. I. Oliveira, Aurelio Ribeiro Leite de, 1962-. II. Ghidini, Carla Taviane  
Lucke da Silva, 1976-. III. Universidade Estadual de Campinas. Instituto de  
Matemática, Estatística e Computação Científica. IV. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Interior point methods iteration reduction with continued iteration

**Palavras-chave em inglês:**

Interior-point methods

Linear programming

Numerical methods

**Área de concentração:** Matemática Aplicada

**Titulação:** Doutora em Matemática Aplicada

**Banca examinadora:**

Aurelio Ribeiro Leite de Oliveira [Orientador]

Francisco de Assis Magalhães Gomes Neto

Marta Ines Velazco Fontova

Roberto Quirino do Nascimento

Jair da Silva

**Data de defesa:** 25-01-2016

**Programa de Pós-Graduação:** Matemática Aplicada

**Tese de Doutorado defendida em 25 de janeiro de 2016 e aprovada**

**Pela Banca Examinadora composta pelos Profs. Drs.**

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

**Prof(a). Dr(a). FRANCISCO DE ASSIS MAGALHÃES GOMES NETO**

**Prof(a). Dr(a). MARTA INES VELAZCO FONTOVA**

**Prof(a). Dr(a). ROBERTO QUIRINO DO NASCIMENTO**

**Prof(a). Dr(a). JAIR DA SILVA**

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no processo de vida acadêmica do aluno.

*Aos meus pais...*

# Agradecimentos

Agradeço imensamente a orientação do Prof. Aurelio e da Profa. Carla, pelo apoio, pela paciência, pelo auxílio prestado através de suas sugestões e por sempre estarem dispostos a discutirem sobre os assuntos desta tese.

À Deus, por estar sempre abençoando minha vida, me amparando nesta jornada.

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

À Yuri por sempre estar do meu lado, me ajudando a superar todos os obstáculos.

Aos professores da banca pelas sugestões que ajudaram a melhorar a redação final do manuscrito.

Aos amigos, Luiz Fernando, Luciana, Francielle, Lucas, Júnior, Gislene, Akemi, Aline, Camila e Fábio pelo apoio e pela ajuda.

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

À FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo, pelo apoio financeiro.

## Resumo

Os métodos de pontos interiores têm sido extensivamente utilizados para resolver os problemas de programação linear de grande porte. Entre todas as variações de métodos de pontos interiores, o preditor corretor com múltiplas correções de centralidade apresenta um grande destaque, devido a sua eficiência e rápida convergência. Este método necessita resolver sistemas lineares, em cada iteração, para determinar a direção de busca, correspondendo ao passo que requer mais tempo de processamento. Neste trabalho, a iteração continuada é apresentada e introduzida no método preditor corretor com múltiplas correções de centralidade, com objetivo de reduzir o número de iterações e o tempo computacional deste método para determinar a solução de problemas de programação linear. A iteração continuada consiste em determinar uma nova direção combinada com a direção de busca dos métodos de pontos interiores. Apresentamos duas novas direções continuadas e duas formas diferentes de utilizá-las, propondo um aumento no tamanho dos passos a serem dados na direção de busca, acelerando a convergência do método. Além disso, utilizamos o algoritmo de ajustamento ótimo para  $p$  coordenadas para determinar melhores pontos iniciais para o método de pontos interiores em conjunto com a iteração continuada. Experimentos computacionais foram realizados e os resultados obtidos ao incorporar a iteração continuada com o método de pontos interiores preditor corretor e as múltiplas correções de centralidade são superiores à abordagem tradicional. A utilização do algoritmo de ajustamento ótimo para  $p$  coordenadas na nova abordagem leva a resultados semelhantes.

**Palavras-chave:** métodos de pontos interiores, programação linear, métodos numéricos.

## Abstract

The interior point methods have been extensively used to solve large-scale linear programming problems. Among all variations of interior point methods, the predictor corrector with multiple centrality corrections is the method of choice due to its efficiency and fast convergence. This method requires solving linear systems, to determine the search direction corresponding to the step that requires more processing time, in each iteration. In this work, the continued iteration is presented and introduced to the predictor corrector method with multiple centrality corrections, in order to reduce the number of iterations and the computational time to determine the linear programming problems solution. The continued iteration consists of determining a new direction combined with the search direction of the interior point methods. Two new continued directions and two different ways of being used, increasing of the steps sizes taken in the search direction, speeding up the convergence of the method. In addition, we use the optimal adjustment algorithm for p coordinates to determine the best starting point for the interior point method in conjunction with the continued iteration. Computational experiments were performed and the results achieved by incorporating the continued iteration in the predictor corrector interior point method and multiple centrality corrections outperform the traditional approach. Using the optimal adjustment algorithm for p coordinates leads to similar results.

**Keywords:** interior point methods, linear programming, numerical methods.

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
<b>2</b>	<b>MÉTODOS DE PONTOS INTERIORES</b>	<b>14</b>
2.1	Problemas de programação linear	14
2.2	Método de Newton	15
2.3	Método de pontos interiores do tipo primal dual	16
2.3.1	Método de pontos interiores preditor corretor	21
2.3.2	Múltiplas Correções de Centralidade	25
<b>3</b>	<b>ITERAÇÃO CONTINUADA</b>	<b>28</b>
3.1	Componentes de Bloqueio	29
3.2	Direções Continuadas	31
3.2.1	Direção preditora corretora continuada	31
3.2.2	Direção Moderada	43
3.2.3	Direção Simples	44
3.3	Formas de utilização da direção continuada	45
3.3.1	Forma avançada	45
3.3.2	Forma atrasada	56
3.4	Critério para utilização da direção continuada	66
<b>4</b>	<b>PONTOS INICIAIS PARA O MÉTODO DE PONTOS INTERIORES</b>	<b>68</b>
4.1	Algoritmo de ajustamento ótimo para $p$ coordenadas	68
4.2	Heurística de Mehrotra	78
<b>5</b>	<b>EXPERIMENTOS COMPUTACIONAIS</b>	<b>80</b>
5.1	Código PCx	80
5.2	Implementação	80
5.3	Problemas teste	81
5.4	Resultados computacionais	82
5.4.1	Comparação AVDM, ATDM, AVDS e ATDS	83
5.4.2	Comparação PCx e ATDS	85
5.4.3	Comparação ATDS e o Algoritmo de ajustamento ótimo para $p$ coordenadas combinado com ATDS	88
<b>6</b>	<b>CONCLUSÕES</b>	<b>92</b>

<b>REFERÊNCIAS</b> . . . . .	<b>94</b>
<b>A</b> <b>FUNÇÃO LAGRANGIANA</b> . . . . .	<b>98</b>
<b>B</b> <b>PERFIL DE DESEMPENHO</b> . . . . .	<b>100</b>

# Capítulo 1

## Introdução

A programação linear é uma das áreas da pesquisa operacional mais utilizadas, devido ao grande número de problemas de otimização que podem ser modelados nesta forma. Um problema de programação linear consiste em determinar uma solução que maximize ou minimize uma função linear, conhecida como função objetivo, sujeita a um conjunto de restrições, que podem ser equações e/ou inequações lineares. Problemas modelados como problemas de programação linear são encontrados em diversas áreas, para determinar lucro máximo, custo mínimo, fluxo máximo, caminho mínimo, alocação de recursos, entre outros. Assim, o desenvolvimento de métodos eficientes e robustos para resolução destes problemas é necessário, desde a muitos anos.

O método simplex foi o primeiro método proposto para resolver problemas de programação linear, desenvolvido por Dantzig em 1947. A região limitada pelas restrições lineares do problema é denominada politopo. Como essas 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. No pior caso que todos os pontos extremos são pesquisados para determinar a solução, o algoritmo apresenta complexidade exponencial [25]. No entanto, o método proposto por Dantzig tem baixa complexidade por iteração e para os problemas reais da época convergia em um número relativo pequeno de iterações.

Em 1979, um algoritmo polinomial foi desenvolvido por Khachiyan [24], utilizando elipsoides. O método das elipsoides teve um grande impacto sobre a teoria de programação linear, porém não foi uma alternativa competitiva com o método simplex, devido aos problemas de estabilidade numérica e o comportamento igual ao pior caso para todos os problemas. Em 1984, Karmarkar [23] divulgou um artigo com resultados formais sobre um novo método com uma complexidade menor do que o método das elipsoides, utilizando a noção de função potencial para garantir uma redução da distância para a otimalidade em cada iteração. Este método, inicialmente, não teve um bom desempenho. As primeiras experiências computacionais, que indicavam que as ideias de Karmarkar

poderiam levar a um algoritmo tão ou mais eficiente que o simplex, foram obtidas com o método dual afim escala, apresentado em 1989 [2]. Embora uma iteração deste método seja cara em comparação com o método simplex, a otimalidade é alcançada em um número relativamente pequeno de iterações em problemas de grande porte. Este método atraiu um grande interesse dos pesquisadores em otimização, que desenvolveram novos estudos, compreendendo melhor essa teoria [22, 26, 27, 29, 31, 32] e dando origem a uma nova classe de métodos, os métodos de pontos interiores.

O estudo dos métodos de pontos interiores para resolução de problemas de programação linear tem sido uma das mais interessantes áreas de pesquisa em otimização nas últimas décadas, devido aos grandes avanços nesta área [20]. Códigos computacionais sofisticados baseados nesses métodos vêm se apresentando como alternativas eficientes para a solução de problemas de grande porte com estruturas genéricas [1, 6, 7, 13, 17, 35, 36, 39]. Os métodos de pontos interiores do tipo primal dual tem sido base para novas variações dessa classe de métodos. Entre todos os métodos do tipo primal dual, o preditor corretor desenvolvido por Mehrotra [31] apresenta destaque devido a sua eficiência e convergência rápida. Na maioria das implementações o esforço computacional do método preditor corretor está concentrado na resolução de dois sistemas lineares com uma mesma matriz de coeficientes em cada iteração, para determinar a direção de busca. Para resolver os sistemas envolvidos, é feita uma eliminação de variáveis de forma a obter um sistema de equações normais cuja matriz é simétrica e definida positiva. Dessa forma, a fatoração de Cholesky é utilizada para determinação da solução. As múltiplas correções de centralidade, propostas por Gondzio [20], para o método de pontos interiores preditor corretor é considerada por muitos pesquisadores uma das variações mais eficientes dos métodos de pontos interiores. Isto porque as múltiplas correções de centralidade fazem com que o método prossiga na vizinhança da trajetória central, proporcionando melhoria da convergência. Neste trabalho, apresentamos a abordagem iteração continuada e o algoritmo de ajustamento ótimo para  $p$  coordenadas desenvolvido por Silva [38], na busca de redução de iterações e, consequentemente, de redução de tempo total do método de pontos interiores.

Inicialmente, a iteração continuada foi desenvolvida para o método dual afim escala [34] e depois estendida para o método de pontos interiores preditor corretor [4]. Esta abordagem consiste em determinar uma nova direção combinada com a direção de busca do método de pontos interiores com o objetivo de acelerar a sua convergência. Neste trabalho, apresentamos a direção preditora corretora continuada [4] e novas direções continuadas: moderada e simples. As direções continuadas têm como característica um esforço menor comparado a uma iteração do método. Desta forma, pode-se obter uma redução no número total de iterações que, potencialmente, pode ser refletido na redução de tempo computacional. Duas formas de combinação da direção de busca do método com a direção continuada são apresentadas. Além disso, uma análise é realizada para encontrar critérios para a utilização desta abordagem no método de pontos interiores. A abordagem

proposta apresenta resultados computacionais superiores ao método preditor corretor com múltiplas correções de centralidade nos experimentos numéricos realizados.

O algoritmo de ajustamento ótimo para  $p$  coordenadas é uma generalização do algoritmo de ajustamento pelo par ótimo [19], que tem como característica principal a simplicidade, pois envolve apenas multiplicação de matriz por vetor e resolve um sistema linear com uma matriz definida positiva de dimensão pequena. Este algoritmo poderia ser utilizado para resolver um problema de programação linear, mas sua convergência é lenta. Como a convergência dos métodos de pontos interiores é sensível ao ponto inicial, então decidimos utilizar o algoritmo de ajustamento ótimo para  $p$  coordenadas para determinar melhores pontos iniciais, como realizado em [37].

A combinação da iteração continuada e o algoritmo de ajustamento ótimo para  $p$  coordenadas foi feita buscando a melhora da convergência dos métodos de pontos interiores.

O texto está organizado da seguinte forma: no Capítulo 2, apresentamos os métodos de pontos interiores para resolução de problemas de programação linear. No Capítulo 3, a iteração continuada e as diferentes formas de utilizá-la são desenvolvidas para a incorporação no método preditor corretor com múltiplas correções de centralidade. O Capítulo 4 aborda uma estratégia para determinar pontos iniciais melhores para o método de pontos interiores utilizando o algoritmo de ajustamento ótimo para  $p$  coordenadas. No Capítulo 5, os resultados dos experimentos computacionais realizados com diversos problemas selecionados de diferentes coleções são apresentados. Além disso, uma análise é feita comparando os resultados obtidos do método de pontos interiores com e sem a iteração continuada e o algoritmo de ajustamento ótimo para  $p$  coordenadas. No Capítulo 6, apresentamos as conclusões. Por fim, estão as referências bibliográficas.

## Capítulo 2

# Métodos de pontos interiores

Neste capítulo, apresentamos os métodos de pontos interiores do tipo primal dual para resolução de problemas de programação linear, destacando o método de preditor corretor, uma vez que este tornou-se base para a sofisticação dos métodos de pontos interiores.

### 2.1 Problemas de programação linear

Um problema de programação linear (PPL) consiste em determinar uma solução que minimiza ou maximiza uma função linear, denominada função objetivo, sujeita a um conjunto de equações e/ou inequações lineares [11]. A forma padrão para o problema de programação linear com variáveis canalizadas [3, 28] é dada por:

$$\begin{aligned} &\text{minimizar} && c^t x \\ &\text{sujeito a} && Ax = b \\ &&& x + s = u \\ &&& (x, s) \geq 0, \end{aligned} \tag{2.1}$$

em que  $A \in \mathbb{R}^{m \times n}$  é a matriz dos coeficientes de posto completo  $m$ ,  $c \in \mathbb{R}^n$  é o vetor de coeficientes da função objetivo (vetor de custo),  $x \in \mathbb{R}^n$  é o vetor das variáveis de decisão,  $s$  é um vetor coluna de dimensão  $n$  de variáveis de folga primais,  $u$  é o vetor dos limitantes superiores das variáveis de decisão e  $b \in \mathbb{R}^m$  é o vetor de termos independentes. Este problema (2.1) é denominado problema primal, ao qual tem-se associado o seguinte problema dual:

$$\begin{aligned} &\text{maximizar} && b^t y - u^t w \\ &\text{sujeito a} && A^t y + z - w = c \\ &&& (z, w) \geq 0, \end{aligned} \tag{2.2}$$

em que  $y \in \mathbb{R}^m$  representa o vetor das variáveis duais livres,  $z \in \mathbb{R}^n$  representa o vetor das variáveis de folga duais e  $w$  é o vetor coluna de dimensão  $n$  de variáveis duais.

Nos problemas primal e dual, o ponto  $(x, s, y, z, w)$  é interior quando  $(x, s, z, w) > 0$ . Além disso, o ponto interior  $(x, y, z, w)$  é factível se satisfaz  $Ax = b$ ,  $x + s = u$  e  $A^t y + z - w = c$ .

O *gap de dualidade* é a diferença entre os valores das funções objetivo dos problemas primal e dual, e é dado por  $\gamma = c^t x - b^t y + u^t w$  que se reduz a  $\gamma = x^t z + s^t w$ , para pontos primais e duais factíveis.

**Teorema 2.1.1.** Um ponto  $(x, s, y, z, w)$  é um ponto ótimo para os problemas primal e dual se, e somente se, satisfaz as condições de otimalidade de primeira ordem (Karush-Kuhn-Tucker) do par de problemas primal e dual (2.1) e (2.2), que são dadas por [40]:

(i) Factibilidade primal

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

(ii) Factibilidade dual

$$A^t y + z - w - c = 0, \quad (z, w) \geq 0, \tag{2.3}$$

(iii) Complementaridade

$$\begin{aligned} XZe &= 0, \\ SWe &= 0, \end{aligned}$$

em que  $X = \text{diag}(x)$ ,  $Z = \text{diag}(z)$ ,  $S = \text{diag}(s)$  e  $W = \text{diag}(w)$  e  $e \in \mathbb{R}^n$ , tal que  $e = (1, 1, \dots, 1)^t$ .

Os métodos de pontos interiores do tipo primal dual utilizam variantes do método de Newton para seu desenvolvimento. Assim, uma breve descrição deste método é apresentada a seguir.

## 2.2 Método de Newton

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

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

O método de Newton é um método iterativo para encontrar uma solução aproximada  $x^* \in \mathbb{R}^n$  de  $F(x) = 0$  [14]. Dada uma aproximação inicial  $x \in \mathbb{R}^n$ , o método busca uma direção  $d$  de forma que  $F(x + d)$  se aproxime do vetor nulo. 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  $x$ , dada por:

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

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

$$J(x)d = -F(x). \quad (2.4)$$

Dessa maneira, o processo iterativo do método de Newton é dado por:

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

em que a direção  $d$  é a solução de (2.4). O processo iterativo é aplicado até que um critério de convergência seja satisfeito.

**Observação 2.2.1.** A direção  $d$  em (2.4) 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. 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 [14].

## 2.3 Método de pontos interiores do tipo primal dual

O método de pontos interiores do tipo primal dual é um método iterativo que inicia com um ponto interior e calcula um novo ponto interior em cada iteração. O método encontra uma solução para o problema aplicando o método de Newton a  $F(x, s, y, z, w) = 0$ , sendo esta equação formada pelas condições de otimalidade do PPL (2.3), desconsiderando  $(x, s, z, w) \geq 0$ . Além disso, o método modifica o tamanho de passo de Newton para que a desigualdade  $(x, s, z, w) > 0$  seja satisfeita, ou seja, para que se obtenha um ponto interior. Dessa maneira, os problemas primal e dual são resolvidos simultaneamente.

Considere:

$$F(x, s, y, z, w) = \begin{bmatrix} Ax - b \\ x + s - u \\ A^t y + z - w - c \\ XZe \\ SWe \end{bmatrix} = - \begin{bmatrix} r_p \\ r_u \\ r_d \\ r_a \\ r_b \end{bmatrix}, \quad (2.5)$$

em que  $r_p, r_u, r_d, r_a, r_b$  são os resíduos primal, dual e os produtos complementares, respectivamente.

Então, determinar a solução ótima do PPL consiste em obter  $(x, s, y, z, w)$  tal que  $F(x, s, y, z, w) = 0$ , pois teremos as condições de otimalidade (2.3) satisfeitas. Assim, o método de Newton é utilizado para determinar a solução.

O processo iterativo do método de Newton utilizando a função  $F$  dada em (2.5), 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

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

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

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{bmatrix} \begin{bmatrix} \tilde{d}x \\ \tilde{d}s \\ \tilde{d}y \\ \tilde{d}z \\ \tilde{d}w \end{bmatrix} = \begin{bmatrix} r_p \\ r_u \\ r_d \\ r_a \\ r_b \end{bmatrix}, \quad (2.6)$$

ou seja,

$$A\tilde{d}x = r_p \quad (2.7)$$

$$\tilde{d}x + \tilde{d}s = r_u \quad (2.8)$$

$$A^t\tilde{d}y + \tilde{d}z - \tilde{d}w = r_d \quad (2.9)$$

$$X\tilde{d}z + Z\tilde{d}x = r_a \quad (2.10)$$

$$S\tilde{d}w + W\tilde{d}s = r_b. \quad (2.11)$$

A direção  $\tilde{d}$ , solução do sistema (2.6), também é conhecida como direção afim escala.

Resolvendo o sistema de equações (2.7) a (2.11) por eliminação de variáveis, encontramos a solução  $(\tilde{d}x, \tilde{d}s, \tilde{d}y, \tilde{d}z, \tilde{d}w)$ . Pelas equações (2.10) e (2.11), temos:

$$\tilde{d}z = X^{-1}(r_a - Z\tilde{d}x), \quad (2.12)$$

$$\tilde{d}w = S^{-1}(r_b - W\tilde{d}s). \quad (2.13)$$

Pela equação (2.8), obtemos:

$$\tilde{d}s = r_u - \tilde{d}x. \quad (2.14)$$

Substituindo  $\tilde{d}s$ ,  $\tilde{d}z$  e  $\tilde{d}w$  em (2.9), temos :

$$A^t\tilde{d}y + X^{-1}(r_a - Z\tilde{d}x) - S^{-1}(r_b - W(r_u - \tilde{d}x)) = r_d$$

$$A^t\tilde{d}y - X^{-1}Z\tilde{d}x - S^{-1}W\tilde{d}x = r_d - X^{-1}r_a + S^{-1}r_b - S^{-1}Wr_u.$$

Assim,

$$A^t\tilde{d}y - (X^{-1}Z + S^{-1}W)\tilde{d}x = r_d - X^{-1}r_a + S^{-1}r_b - S^{-1}Wr_u. \quad (2.15)$$

Definindo  $D = (X^{-1}Z + S^{-1}W)^{-1}$  e reescrevendo a equação (2.15), obtemos:

$$A^t\tilde{d}y - D^{-1}\tilde{d}x = r_d - X^{-1}r_a + S^{-1}r_b - S^{-1}Wr_u. \quad (2.16)$$

Desta forma, por (2.7) e (2.16), temos um sistema aumentado dado por:

$$\begin{bmatrix} -D^{-1} & A^t \\ A & 0 \end{bmatrix} \begin{bmatrix} \tilde{d}x \\ \tilde{d}y \end{bmatrix} = \begin{bmatrix} \tilde{r} \\ r_p \end{bmatrix}, \quad (2.17)$$

em que  $\tilde{r} = r_d - X^{-1}r_a - S^{-1}r_b - S^{-1}Wr_u$ .

Eliminando a variável  $\tilde{d}x$  em (2.17), que é dada por:

$$\tilde{d}x = D(A^t\tilde{d}y - \tilde{r}),$$

obtem-se

$$AD(A^t\tilde{d}y - \tilde{r}) = r_p,$$

ou seja,

$$ADA^t\tilde{d}y = r_p + AD\tilde{r}. \quad (2.18)$$

Como a matriz  $A$  tem posto completo, então  $ADA^t$  é simétrica definida positiva e a fatoração de Cholesky pode ser utilizada para a resolução do sistema (2.18).

Logo,

$$\tilde{d}y = (ADA^t)^{-1}(r_p + AD\tilde{r}). \quad (2.19)$$

Portanto, a direção afim escala dos métodos de pontos interiores primal dual é dada por:

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

$$\tilde{d}x = D(A^t\tilde{d}y - r_d + X^{-1}r_a - S^{-1}r_b + S^{-1}Wr_u),$$

$$\tilde{d}s = r_u - \tilde{d}x,$$

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

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

Para manter o ponto interior em cada iteração do método de pontos interiores, o tamanho do passo de Newton, possivelmente de valor 1, necessita ser reduzido. Desse modo, os métodos de pontos interiores do tipo primal dual são métodos iterativos cujo novo ponto é dado por:

$$x^{k+1} = x^k + \alpha_p \tilde{d}x^k,$$

$$s^{k+1} = s^k + \alpha_p \tilde{d}s^k,$$

$$y^{k+1} = y^k + \alpha_d \tilde{d}y^k,$$

$$z^{k+1} = z^k + \alpha_d \tilde{d}z^k,$$

$$w^{k+1} = w^k + \alpha_d \tilde{d}w^k,$$

em que  $\tilde{d}^k = (\tilde{d}x^k, \tilde{d}s^k, \tilde{d}y^k, \tilde{d}z^k, \tilde{d}w^k)$  é obtida do sistema linear (2.6). Os tamanhos dos passos primal e dual,  $\alpha_p$  e  $\alpha_d$ , respectivamente, são calculados da seguinte forma:

$$\alpha_p = \min \{1, \tau \max \{ \alpha : x^k + \alpha dx^k > 0 \text{ e } s^k + \alpha ds^k > 0 \} \}, \quad (2.20)$$

$$\alpha_d = \min \{1, \tau \max \{ \alpha : z^k + \alpha dz^k > 0 \text{ e } w^k + \alpha dw^k > 0 \} \},$$

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

Considere:

$$\rho_x = \min_{1 \leq l \leq n} \left\{ -\frac{x_l^k}{dx_l^k} : x_l^k + dx_l^k \leq 0 \right\},$$

$$\rho_s = \min_{1 \leq l \leq n} \left\{ -\frac{s_l^k}{ds_l^k} : s_l^k + ds_l^k \leq 0 \right\},$$

$$\rho_z = \min_{1 \leq l \leq n} \left\{ -\frac{z_l^k}{dz_l^k} : z_l^k + dz_l^k \leq 0 \right\},$$

$$\rho_w = \min_{1 \leq l \leq n} \left\{ -\frac{w_l^k}{dw_l^k} : w_l^k + dw_l^k \leq 0 \right\}.$$

Podemos reescrever (2.20), como:

$$\alpha_p = \min \{1, \tau\rho_x, \tau\rho_s\}, \quad (2.21)$$

$$\alpha_d = \min \{1, \tau\rho_z, \tau\rho_w\}.$$

Variações deste método calculam uma combinação de outras direções que satisfazem sistemas que tem a matriz dos coeficientes de (2.6), como veremos mais adiante no caso do método de pontos interiores preditor corretor.

### Critério de parada

O critério de parada dos métodos de pontos interiores é baseado nas condições de otimalidade (2.3).

Sejam  $\varepsilon_p$ ,  $\varepsilon_d$  e  $\varepsilon_o$  tolerâncias de factibilidade primal, factibilidade dual e otimalidade, respectivamente, então o critério de parada é dado por:

$$\frac{\|r_p^k\|}{1 + \|b\|} \leq \varepsilon_p,$$

$$\frac{\|r_u^k\|}{1 + \|u\|} \leq \varepsilon_p,$$

$$\frac{\|r_d^k\|}{1 + \|c\|} \leq \varepsilon_d,$$

$$\frac{(x^k)^t z^k + (s^k)^t w^k}{1 + |c^t x^k|} \leq \varepsilon_o.$$

**Observação 2.3.1.** Na prática, geralmente, a factibilidade primal e dual é obtida antes que a otimalidade seja satisfeita.

### 2.3.1 Método de pontos interiores preditor corretor

No método de pontos interiores preditor corretor, desenvolvido por Mehrotra [31], determina-se uma direção, denominada preditora corretora, que é composta por três componentes: direção afim escala, direção de centragem e direção de correção não linear. A direção afim escala (direção de Newton) é encontrada como descrito anteriormente em (2.6). A direção de centragem se justifica porque os produtos complementares  $x_i z_i$  e  $s_i w_i$ , que formam os resíduos  $r_a$  e  $r_b$ , podem convergir para zero com velocidades diferentes, fazendo com que o método falhe ou progrida lentamente. Assim, um parâmetro  $\mu$  é acrescentado ao resíduo das condições de otimalidade, de forma que na direção obtenha-se  $x_i z_i = s_i w_i = \mu$ . Este parâmetro é atualizado em cada iteração e além disso, deve tender a zero à medida que o método se aproxima de uma solução. O valor de  $\mu$  é calculado após o cálculo da direção afim escala (2.6), avaliando o *gap da dualidade*,

$$\tilde{\gamma} = (x + \tilde{\alpha}_P \tilde{d}x)^t (z + \tilde{\alpha}_D \tilde{d}z) + (s + \tilde{\alpha}_P \tilde{d}s)^t (w + \tilde{\alpha}_D \tilde{d}w),$$

que resultaria ao utilizar a direção afim escala. Caso a redução de  $\tilde{\gamma}$  seja grande em relação a  $\gamma = x^t z + s^t w$  então o valor de  $\mu$  é pequeno, caso contrário, esse valor é grande. Procura-se, assim, obter produtos complementares mais próximo entre si, ou seja, pontos mais centrais.

Uma proposta de Mehrotra [31] para o cálculo de  $\mu$  é a seguinte:

$$\mu = \left( \frac{\tilde{\gamma}}{\gamma} \right)^p \bar{\mu},$$

em que  $\bar{\mu} = \left( \frac{\gamma}{2n} \right)$ .

**Observação 2.3.2.** Geralmente, o valor de  $p$  é 2 ou 3 [31].

Caso um tamanho de passo completo ( $\alpha_p = \alpha_d = 1$ ) seja dado, então uma motivação para a correção não linear é que obtemos a factibilidade primal e dual, isto é,

são obtidos  $r_p = r_u = r_d = 0$ . De fato, de (2.7) - (2.9), temos:

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

$$r_u^{k+1} = u - x^{k+1} - s^{k+1} = u - (x^k + \tilde{d}x^k) - (s^k + \tilde{d}s^k) = r_u^k - r_u^k = 0.$$

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

Porém, os novos pares de produtos complementares são dados por:

$$(x_i + \tilde{d}x_i)(z_i + \tilde{d}z_i) = x_i z_i + x_i \tilde{d}z_i + z_i \tilde{d}x_i + \tilde{d}x_i \tilde{d}z_i,$$

$$(s_i + \tilde{d}s_i)(w_i + \tilde{d}w_i) = s_i w_i + s_i \tilde{d}w_i + w_i \tilde{d}s_i + \tilde{d}s_i \tilde{d}w_i,$$

em que  $i = 1, \dots, n$ .

Como segundo (2.10) - (2.11),  $-x_i z_i = x_i \tilde{d}z_i + z_i \tilde{d}x_i$  e  $-s_i w_i = s_i \tilde{d}w_i + w_i \tilde{d}s_i$ , segue que:

$$(x_i + \tilde{d}x_i)(z_i + \tilde{d}z_i) = \tilde{d}x_i \tilde{d}z_i,$$

$$(s_i + \tilde{d}s_i)(w_i + \tilde{d}w_i) = \tilde{d}s_i \tilde{d}w_i.$$

Desta forma, como estes valores são diferentes de zero, uma correção não linear é realizada na tentativa de compensar estes desvios, pois sabemos que  $r_a^{k+1} = -\tilde{D}x\tilde{D}ze$  e  $r_b^{k+1} = -\tilde{D}s\tilde{D}we$ , em que  $\tilde{D}x = \text{diag}(\tilde{d}x)$ ,  $\tilde{D}z = \text{diag}(\tilde{d}z)$ ,  $\tilde{D}s = \text{diag}(\tilde{d}s)$  e  $\tilde{D}w = \text{diag}(\tilde{d}w)$ , quando  $\alpha_p = \alpha_d = 1$ .

Em vez de resolver dois sistemas lineares, para determinar a direção de centralização e correção não linear separadamente, apenas um sistema linear adicional é resolvido para determinar a combinação dessas duas direções, sendo dado por:

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{bmatrix} \begin{bmatrix} \bar{d}x \\ \bar{d}s \\ \bar{d}y \\ \bar{d}z \\ \bar{d}w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \bar{r}_a \\ \bar{r}_b \end{bmatrix}, \quad (2.22)$$

em que  $\bar{r}_a = \mu e - (\tilde{D}x\tilde{D}z)e$ ,  $\bar{r}_b = \mu e - (\tilde{D}s\tilde{D}w)e$ . Como  $X$ ,  $S$ ,  $Z$  e  $W$  não são alterados ao calcular a direção afim escala, podemos observar que temos a mesma matriz do sistema linear (2.6). Denominamos a direção  $\bar{d} = (\bar{d}x, \bar{d}s, \bar{d}y, \bar{d}z, \bar{d}w)$  como direção corretora.

Através de eliminações de variáveis, encontramos a solução:

$$\bar{d}y = (ADA^t)^{-1}AD(-X^{-1}\bar{r}_a + S^{-1}\bar{r}_b),$$

$$\bar{d}x = D(A^t\bar{d}y + X^{-1}\bar{r}_a - S^{-1}\bar{r}_b),$$

$$\bar{d}s = -\bar{d}x,$$

$$\bar{d}w = S^{-1}(\bar{r}_a - W\bar{d}s),$$

$$\bar{d}z = X^{-1}(\bar{r}_b - Z\bar{d}x).$$

Então, a direção preditora corretora é a combinação das direções  $\tilde{d}$  e  $\bar{d}$  obtidas em (2.6) e (2.22), respectivamente, ou seja,  $d = \tilde{d} + \bar{d}$ .

**Observação 2.3.3.** A direção  $d = \tilde{d} + \bar{d}$  pode ser interpretada como a solução do sistema linear:

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{bmatrix} \begin{bmatrix} dx \\ ds \\ dy \\ dz \\ dw \end{bmatrix} = \begin{bmatrix} r_p \\ r_u \\ r_d \\ r_s \\ r_t \end{bmatrix}, \quad (2.23)$$

em que  $r_s = r_a + \bar{r}_a$  e  $r_t = r_b + \bar{r}_b$ .

O método preditor corretor com variáveis canalizadas é resumido no Algoritmo 1.

**Algoritmo 1** Método Preditor Corretor**Parâmetros**

$\tau = 0,99995$ ,  $\varepsilon_p$ ,  $\varepsilon_d$  e  $\varepsilon_o$ ;

**Inicialização**

$k = 0$ ,  $(x^0, s^0, y^0, z^0, w^0)$ , com  $x^0 > 0$ ,  $s^0 > 0$ ,  $z^0 > 0$ ,  $w^0 > 0$ ;

$\gamma^0 = (x^0)^t z^0 + (s^0)^t w^0$ ;

$r_p^0 = b - Ax^0$ ,  $r_d^0 = c - A^t y^0 - z^0 + w^0$ ,  $r_u^0 = u - x^0 - s^0$ ;

**Processo Iterativo**

1. Enquanto  $\left( \frac{\|r_p^k\|}{1 + \|b\|} > \varepsilon_p \text{ ou } \frac{\|r_u^k\|}{1 + \|u\|} > \varepsilon_p \text{ ou } \frac{\|r_d^k\|}{1 + \|c\|} > \varepsilon_d \text{ ou } \frac{\gamma^k}{1 + |c^t x^k|} > \varepsilon_o \right)$  faça
  2. Encontre a direção afim escala  $\bar{d}$  resolvendo o sistema (2.6);
  3. Encontre a direção de centragem e correção não linear  $\tilde{d}$  resolvendo o sistema (2.22);
  4. Faça  $d = \tilde{d} + \bar{d}$ ;
  5. Encontre os tamanhos dos passos primal e dual por (2.21);
  6. Atualize o ponto
 
$$\begin{aligned} x^{k+1} &= x^k + \alpha_p dx; \\ s^{k+1} &= s^k + \alpha_p ds; \\ y^{k+1} &= y^k + \alpha_d dy; \\ z^{k+1} &= z^k + \alpha_d dz; \\ w^{k+1} &= w^k + \alpha_d dw; \end{aligned}$$
  7. Calcule os resíduos e  $\gamma^{k+1}$ 

$$\begin{aligned} r_p^{k+1} &= b - Ax^{k+1}; \\ r_d^{k+1} &= c - A^t y^{k+1}; \\ r_u^{k+1} &= u - x^{k+1} - s^{k+1}; \\ \gamma^{k+1} &= (x^{k+1})^t z^{k+1} + (s^{k+1})^t w^{k+1}; \end{aligned}$$
  8. Atualize  $k = k + 1$ ;
9. Fim enquanto

**Observação 2.3.4.** A infactibilidade primal e dual é reduzida por  $(1 - \alpha_p)$  e  $(1 - \alpha_d)$ , respectivamente, em cada iteração dos métodos de pontos interiores do tipo primal dual.

De fato, de (2.23), temos:

$$r_p^{k+1} = b - Ax^{k+1} = b - A(x^k + \alpha_p dx^k) = b - Ax^k - \alpha_p A dx^k = (1 - \alpha_p)r_p^k.$$

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

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

(2.24)

### 2.3.2 Múltiplas Correções de Centralidade

As múltiplas correções de centralidade propostas por Gondzio [21] e estudadas com mais profundidade em [12], consistem em determinar novas direções de correções resolvendo sistemas lineares do tipo (2.22), com dois propósitos: alcançar passos maiores nos espaços primal e dual e melhorar a centralidade da iterada atual. Dessa forma, ocorre uma redução da infactibilidade primal e dual, melhorando a convergência. Em [21], Gondzio enfatiza que as múltiplas correções de centralidade não tem como meta a redução do *gap da complementaridade*, uma vez que este é potencialmente reduzido ao obter passos maiores na próxima iteração. Apresentamos a seguir como é realizado o cálculo da direção de correção de centralidade.

Considere a direção preditora  $\tilde{d}$ , sendo a direção preditora corretora  $d$ . Suponha que os maiores tamanhos de passos ( $\alpha_p$  e  $\alpha_d$ ) que podem ser dados nesta direção  $\tilde{d}$  e que mantêm o ponto interior tenham sido calculados. Como o propósito das múltiplas correções de centralidade é aumentar os passos, então é tomado

$$\tilde{\alpha}_p = \min \{ \alpha_p + \delta, 1 \} \quad \text{e} \quad \tilde{\alpha}_d = \min \{ \alpha_d + \delta, 1 \}, \quad (2.25)$$

para  $\delta \in (0, 1)$  e o seguinte ponto é avaliado:

$$\begin{aligned} \tilde{x} &= x + \tilde{\alpha}_p \tilde{d}x; \\ \tilde{s} &= s + \tilde{\alpha}_p \tilde{d}s; \\ \tilde{z} &= z + \tilde{\alpha}_d \tilde{d}z; \\ \tilde{w} &= w + \tilde{\alpha}_d \tilde{d}w. \end{aligned}$$

Como esse novo ponto, dificilmente os produtos complementares  $\tilde{X}\tilde{Z}e$  e  $\tilde{S}\tilde{W}e$  são iguais a  $\bar{\mu}$ . Além disso, possivelmente  $\tilde{x}$ ,  $\tilde{s}$ ,  $\tilde{z}$  e  $\tilde{w}$  violam a não negatividade, gerando componentes negativas nos produtos complementares. Assim, uma direção de correção é calculada para compensar esses desvios, movendo o ponto de forma que, os produtos complementares satisfaçam

$$\begin{aligned} \sigma \bar{\mu} &\leq \tilde{x}_i \tilde{z}_i \leq \sigma^{-1} \bar{\mu}, \\ \sigma \bar{\mu} &\leq \tilde{s}_i \tilde{w}_i \leq \sigma^{-1} \bar{\mu}, \end{aligned}$$

em que  $i = 1, \dots, n$ , ou seja, o ponto esteja na vizinhança da trajetória central.

A direção  $\bar{d}$  é obtida resolvendo o seguinte sistema linear:

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{bmatrix} \begin{bmatrix} \bar{d}x \\ \bar{d}s \\ \bar{d}y \\ \bar{d}z \\ \bar{d}w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \bar{r}_a \\ \bar{r}_b \end{bmatrix}, \quad (2.26)$$

em que  $\bar{r}_a, \bar{r}_b \in \mathbb{R}^n$ , para  $l = 1, \dots, n$  são dados por:

$$(\bar{r}_a)_l = \begin{cases} \sigma\bar{\mu} - \tilde{x}_l\tilde{z}_l, & \text{se } \tilde{x}_l\tilde{z}_l \leq \sigma\bar{\mu} \\ \sigma^{-1}\bar{\mu} - \tilde{x}_l\tilde{z}_l, & \text{se } \tilde{x}_l\tilde{z}_l \geq \sigma^{-1}\bar{\mu} \\ 0, & \text{caso contrário,} \end{cases} \quad (2.27)$$

$$(\bar{r}_b)_l = \begin{cases} \sigma\bar{\mu} - \tilde{s}_l\tilde{w}_l, & \text{se } \tilde{s}_l\tilde{w}_l \leq \sigma\bar{\mu} \\ \sigma^{-1}\bar{\mu} - \tilde{s}_l\tilde{w}_l, & \text{se } \tilde{s}_l\tilde{w}_l \geq \sigma^{-1}\bar{\mu} \\ 0, & \text{caso contrário,} \end{cases}$$

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

O sistema linear (2.26) é similar ao (2.22), somente sendo alterados os valores de  $\bar{r}_a$  e  $\bar{r}_b$ . Assim, são resolvidos de maneira análoga, utilizando eliminações de variáveis, e a fatoração de Cholesky já calculada de  $ADA^t$  para determinar  $\bar{d}y$ .

A direção final é dada por  $d = \tilde{d} + \bar{d}$ . Os tamanhos dos passos são calculados para a direção  $d$ . Caso os valores sejam aumentados, dentro de uma certa tolerância estabelecida, então o ponto é atualizado. Caso contrário, utiliza-se somente a direção preditora corretora calculada no método preditor corretor. Pode-se realizar mais de uma correção, não havendo a necessidade de atualizar o ponto e tomando a direção preditora  $\tilde{d}$  como a direção  $d$ . O processo é repetido até que um número máximo de correções seja atingido [21] ou finalizado se não ocorrer um aumento no tamanho dos passos.

Essa abordagem não altera os resultados de complexidade polinomial obtidos em [40], assim obtém melhores resultados na prática mantendo os resultados teóricos [12].

As múltiplas correções de centralidade são dadas resumidamente no Algoritmo

2.

**Algoritmo 2** Múltiplas Correções de Centralidade**Entrada** $(x, s, y, z, w)$  e  $(dx, ds, dy, dz, dw)$ ;**Parâmetros** $\delta, \beta_{min}, \beta_{max}, \sigma \in K$ ;**Inicialização** $k = 0$ ; $(\tilde{d}x, \tilde{d}s, \tilde{d}y, \tilde{d}z, \tilde{d}w) = (dx, ds, dy, dz, dw)$ ;Encontre  $\alpha_p$  e  $\alpha_d$  dados em (2.21)**Processo Iterativo**

1. Enquanto  $k < K$  faça
  2. Encontre  $\tilde{\alpha}_p$  e  $\tilde{\alpha}_d$  dados em (2.25);
  3. Calcule
 
$$\tilde{x} = x + \tilde{\alpha}_p dx;$$

$$\tilde{s} = s + \tilde{\alpha}_p ds;$$

$$\tilde{y} = y + \tilde{\alpha}_d dy;$$

$$\tilde{z} = z + \tilde{\alpha}_d dz;$$

$$\tilde{w} = w + \tilde{\alpha}_d dw;$$
  4. Encontre a direção de correção  $\bar{d}$  resolvendo o sistema (2.26);
  5. Faça  $\hat{d} = \tilde{d} + \bar{d}$ ;
  6. Encontre os tamanhos dos passos primal e dual,  $\hat{\alpha}_p$  e  $\hat{\alpha}_d$ ;
  7. Se  $\hat{\alpha}_p \geq \alpha_p + \sigma\delta$  e  $\hat{\alpha}_d \geq \alpha_d + \sigma\delta$  então
 
$$(\hat{d}x, \hat{d}s, \hat{d}y, \hat{d}z, \hat{d}w) = (\hat{d}x, \hat{d}s, \hat{d}y, \hat{d}z, \hat{d}w);$$

$$k = k + 1;$$
  8. Caso contrário
 
$$(dx, ds, dy, dz, dw) = (\hat{d}x, \hat{d}s, \hat{d}y, \hat{d}z, \hat{d}w);$$
 termina o processo iterativo
  9. Fim se
10. Fim enquanto

---

O maior esforço computacional, em cada iteração, do método de pontos interiores preditor corretor está na resolução dos dois sistemas lineares com a mesma matriz dos coeficientes, para determinar a direção. Utilizando eliminações de variáveis, temos o sistema de equações normais e uma abordagem muito utilizada para resolvê-lo é a fatoração de Cholesky. Para muitos problemas, esta fatoração é muito cara, gerando um alto custo computacional em cada iteração. Ao utilizar as múltiplas correções de centralidade, o método preditor corretor obtém uma direção de busca que possibilita que o método prossiga na vizinhança da trajetória central, que tamanhos de passo maiores sejam dados e que uma melhoria na convergência seja obtida. Dessa forma, muitos pesquisadores consideram essa variação dos métodos de pontos interiores a mais eficiente. Buscando uma outra abordagem para a melhoria do método preditor corretor, além das múltiplas correções de centralidade, apresentamos no capítulo seguinte a iteração continuada.

## Capítulo 3

# Iteração Continuada

A iteração continuada consiste em determinar uma nova direção para o método de pontos interiores a fim de reduzir o número de iterações necessárias para obter uma solução ótima do problema e, conseqüentemente, diminuir o tempo computacional total de resolução. A nova direção é obtida combinando a direção de busca do método com a direção continuada a ser calculada. Como, em cada iteração dos métodos de pontos interiores, as infactibilidades primal e dual são reduzidas por  $(1 - \alpha_p)$  e  $(1 - \alpha_d)$ , a proposta é determinar uma direção continuada com menor alteração possível em relação à direção de busca  $d$ . Isso permite aumentar o tamanho dos passos dados nesta direção, reduzindo as infactibilidades. A proposta inicial da iteração continuada foi desenvolvida para o método dual afim escala [34] e depois estendida para o método de pontos interiores preditor corretor [4]. Neste último caso, uma direção continuada é calculada com as mesmas características da direção preditora corretora, resolvendo os dois sistemas lineares, porém de forma aproximada. No cálculo desta direção continuada, embora seja utilizada a fatoração de Cholesky já calculada, para alguns problema de grande porte pode-se elevar o custo computacional, uma vez que mais sistemas lineares devem ser resolvidos. Neste trabalho, propomos novas direções continuadas que exigem menor esforço comparado à direção desenvolvida em [4], sendo utilizadas com o método de pontos interiores preditor corretor com as múltiplas correções de centralidade. Como ao utilizar as múltiplas correções de centralidade determinamos um ponto próximo da vizinhança central e não desejamos que isso seja perdido ao utilizar a iteração continuada, então a utilizamos após o cálculo da direção preditora corretora.

Para apresentar as direções continuadas, necessitamos das componentes de bloqueio definidas a seguir, pois somente quando elas existem é que o passo máximo não foi dado, podendo dessa forma ser aumentado.

### 3.1 Componentes de Bloqueio

As componentes de bloqueio são as componentes que indicam quanto o tamanho do passo completo de Newton deve ser reduzido para manter o ponto interior.

Sejam o ponto e a direção preditora corretora na iteração  $k$  do método dados por  $(x^k, s^k, y^k, z^k, w^k)$  e  $(dx^k, ds^k, dy^k, dz^k, dw^k)$ . Considere:

$$h_p = \min_{l_1 \in L_1, l_2 \in L_2} \left\{ -\frac{x_{l_1}^k}{dx_{l_1}^k}; -\frac{s_{l_2}^k}{ds_{l_2}^k} \right\} \quad \text{e} \quad h_d = \min_{l_3 \in L_3, l_4 \in L_4} \left\{ -\frac{z_{l_3}^k}{dz_{l_3}^k}; -\frac{w_{l_4}^k}{dw_{l_4}^k} \right\},$$

em que

$$L_1 = \operatorname{argmin}_{1 \leq l \leq n} \left\{ -\frac{x_l^k}{dx_l^k} | x_l^k + dx_l^k \leq 0 \right\}, \quad L_3 = \operatorname{argmin}_{1 \leq l \leq n} \left\{ -\frac{z_l^k}{dz_l^k} | z_l^k + dz_l^k \leq 0 \right\},$$

$$L_2 = \operatorname{argmin}_{1 \leq l \leq n} \left\{ -\frac{s_l^k}{ds_l^k} | s_l^k + ds_l^k \leq 0 \right\}, \quad L_4 = \operatorname{argmin}_{1 \leq l \leq n} \left\{ -\frac{w_l^k}{dw_l^k} | w_l^k + dw_l^k \leq 0 \right\}.$$

Considere  $q_t$  a cardinalidade de  $L_t$ , ou seja,  $q_t = |L_t|$ , para  $t = 1, 2, 3, 4$ .

Se  $h_p = -x_{l_1}^k/dx_{l_1}^k$ , então  $L_1$  é o conjunto das componentes de bloqueio no espaço primal que estão na direção  $dx^k$ . Se  $h_p = -s_{l_2}^k/ds_{l_2}^k$ , então  $L_2$  é o conjunto das componentes de bloqueio no espaço primal que estão na direção  $ds^k$ .

Os conjuntos das componentes de bloqueio no espaço dual são definidas de forma similar. Se  $h_d = -z_{l_3}^k/dz_{l_3}^k$ , então  $L_3$  é o conjunto das componentes de bloqueio no espaço dual que estão na direção  $dz^k$ . Se  $h_d = -w_{l_4}^k/dw_{l_4}^k$ , então  $L_4$  é o conjunto das componentes de bloqueio dual que estão na direção  $dw^k$ .

Dados os conjuntos das componentes de bloqueio  $L_t$ , definimos as matrizes  $E_t \in \mathbb{R}^{n \times q_t}$  formadas pelos vetores canônicos nas colunas e com valor um na posição  $l$ , para todo  $l \in L_t$ , desde que  $L_t$  seja conjunto das componentes de bloqueio. Observe que  $E_t$  corresponde a  $E_x$ ,  $E_s$ ,  $E_z$  ou  $E_w$  para  $t = 1, \dots, 4$ , respectivamente. Utilizaremos adiante estas matrizes para determinar as direções continuadas.

Com os conjuntos das componentes de bloqueio definidas, podemos reescrever o tamanho de passo dado em (2.21) como

$$\alpha_p = \min \{1, \tau h_p\},$$

$$\alpha_d = \min \{1, \tau h_d\}.$$

Dessa forma, se existirem as componentes de bloqueio, os tamanhos dos passos máximos que podem ser dados são:

$$\alpha_p = \tau h_p < 1 \quad \text{e} \quad \alpha_d = \tau h_d < 1,$$

mantendo interiores todas as componentes do ponto. Porém, para as componentes que não realizam bloqueio poderíamos aumentar o tamanho de passo, tomando

$$\alpha_p + \delta_p \leq 1 \quad \text{e} \quad \alpha_d + \delta_d \leq 1,$$

para  $\delta_p > 0$  e  $\delta_d > 0$ , ainda mantendo o ponto interior. Dessa forma, com a iteração continuada no caso do ponto primal, buscamos encontrar o seguinte ponto

$$\hat{x}_l^{k+1} = \begin{cases} x_{l_1}^k + \theta_p^1 dx_{l_1}^k, & \text{se } l = l_1 \in L_1 \text{ realizou bloqueio em } dx^k; \\ x_l^k + \theta_p^2 dx_l^k, & \text{caso contrário;} \end{cases} \quad (3.1)$$

$$\hat{s}_l^{k+1} = \begin{cases} s_{l_2}^k + \theta_p^1 ds_{l_2}^k, & \text{se } l = l_2 \in L_2 \text{ realizou bloqueio em } ds^k; \\ s_l^k + \theta_p^2 ds_l^k, & \text{caso contrário;} \end{cases}$$

em que  $\theta_p^1 = \alpha_p - \delta_p^1 \leq \alpha_p$  e  $\theta_p^2 = \alpha_p + \delta_p^2 > \alpha_p$ , sendo  $\delta_p^1 \geq 0$  e  $\delta_p^2 > 0$ , para  $l = 1, \dots, n$ .

Podemos reescrever (3.1) como:

$$\hat{x}^{k+1} = x^k + \theta_p^2 (dx - E_x E_x^t dx) + \theta_p^1 E_x E_x^t dx,$$

$$\hat{s}^{k+1} = s^k + \theta_p^2 (ds - E_s E_s^t ds) + \theta_p^1 E_s E_s^t ds,$$

ou

$$\hat{x}^{k+1} = x^k + \theta_p^2 \left[ I - \left( 1 - \frac{\theta_p^1}{\theta_p^2} \right) E_x E_x^t \right] dx, \quad (3.2)$$

$$\hat{s}^{k+1} = s^k + \theta_p^2 \left[ I - \left( 1 - \frac{\theta_p^1}{\theta_p^2} \right) E_s E_s^t \right] ds.$$

O mesmo ocorre para as demais variáveis  $z$  e  $w$ , que podem ser escritas como:

$$\hat{z}^{k+1} = z^k + \theta_d^2 \left[ I - \left( 1 - \frac{\theta_d^1}{\theta_d^2} \right) E_z E_z^t \right] dz, \quad (3.3)$$

$$\hat{w}^{k+1} = w^k + \theta_d^2 \left[ I - \left( 1 - \frac{\theta_d^1}{\theta_d^2} \right) E_w E_w^t \right] dw,$$

em que  $\theta_d^1 = \alpha_d - \delta_d^1 \leq \alpha_d$  e  $\theta_d^2 = \alpha_d + \delta_d^2 > \alpha_d$ , sendo  $\delta_d^1 \geq 0$  e  $\delta_d^2 > 0$ .

Como a variável  $y$  é livre, o tamanho de passo pode ser aumentado sem restrição, sendo dado por:

$$\hat{y}^{k+1} = y^k + \theta_d^2 dy. \quad (3.4)$$

Os valores de  $\delta_p^1$ ,  $\delta_p^2$ ,  $\delta_d^1$  e  $\delta_d^2$  dependem de qual forma de utilização da direção continuada estamos considerando. Antes de apresentar as formas de utilização, apresentaremos as direções continuadas propostas e depois como utilizá-las para obter os tamanhos de passos (3.2) e (3.3).

Neste trabalho, consideramos que existe ao menos uma componente de bloqueio em cada espaço, primal e dual, ou seja,  $h_p \neq \emptyset$  e  $h_d \neq \emptyset$ . Optamos, dessa maneira, por um avanço nos espaços primal e dual, para facilidade de exposição das ideias.

## 3.2 Direções Continuadas

As direções continuadas ( $\hat{d}$ ) devem proporcionar a menor alteração possível em relação à direção preditora corretora  $d$ , de modo que a combinação,  $d + \delta\hat{d}$ ,  $\delta > 0$  possibilite encontrar as variáveis dadas em (3.2) e (3.3). Como a direção  $d$  é solução única do sistema linear (2.23), então calculamos  $\hat{d}$  como solução aproximada deste sistema linear. Além disso, como não pode ocorrer aumento no passo das componentes de bloqueio na direção  $d$ , devemos ter

$$E_x^t \hat{d}x = 0, \quad E_s^t \hat{d}s = 0, \quad (3.5)$$

$$E_z^t \hat{d}z = 0 \quad \text{e} \quad E_w^t \hat{d}w = 0. \quad (3.6)$$

Apresentamos a seguir, as propostas para as direções continuadas.

### 3.2.1 Direção preditora corretora continuada

A direção preditora corretora continuada proposta em [4] procura calcular uma direção  $\hat{d}$  muito próxima de  $d$ . Então, determina  $\hat{d}$  resolvendo aproximadamente os mesmos sistemas lineares (2.6) e (2.22), que permitiram calcular a direção  $d$ . Assim, no cálculo dessa direção, determinamos uma direção afim escala e depois uma direção de centragem e correção não linear, conforme descrito a seguir.

#### Cálculo da direção afim escala continuada

A direção afim escala continuada,  $\tilde{d}$ , deve ser solução aproximada do sistema linear que determina a direção afim escala (2.6), ou seja,

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{bmatrix} \begin{bmatrix} \tilde{d}x \\ \tilde{d}s \\ \tilde{d}y \\ \tilde{d}z \\ \tilde{d}w \end{bmatrix} \simeq \begin{bmatrix} r_p \\ r_u \\ r_d \\ r_a \\ r_b \end{bmatrix}, \quad (3.7)$$

em que  $\tilde{d}x, \tilde{d}s, \tilde{d}z$  e  $\tilde{d}w \in \mathbb{R}^n$ ,  $\tilde{d}y \in \mathbb{R}^m$ . Além disso,  $\tilde{d}$  deve satisfazer

$$E_x^t \tilde{d}x = 0, \quad E_s^t \tilde{d}s = 0, \quad (3.8)$$

$$E_z^t \tilde{d}z = 0 \quad \text{e} \quad E_w^t \tilde{d}w = 0. \quad (3.9)$$

Para encontrar  $\tilde{d}$ , primeiramente, determinaremos  $\tilde{d}x$  utilizando um problema auxiliar, com a equação  $A\tilde{d}x = r_p$ , do sistema linear (3.7), com (3.8) e (3.9).

**Observação 3.2.1.** Quando  $E_t$  não está definido,  $f_t$  para  $t = x, s$ , também não está definido. O mesmo é válido para o vetor  $g$ . Além disso, caso haja somente uma componente de bloqueio no espaço primal,  $f$  é um escalar. O mesmo é válido no caso dual.

Formulamos um problema auxiliar da seguinte forma:

$$\begin{aligned} \min \quad & \frac{1}{2} \|D^{-\frac{1}{2}} \tilde{d}x - D^{-\frac{1}{2}} dx\|^2 \\ \text{s.a} \quad & A\tilde{d}x = r_p \\ & E_i^t \tilde{d}x = f \\ & E_j^t \tilde{d}x = g, \end{aligned} \quad (3.10)$$

em que  $D = (X^{-1}Z + S^{-1}W)^{-1}$ ,

$$\begin{aligned} E_i &= \begin{bmatrix} E_x & E_s \end{bmatrix}, & E_j &= \begin{bmatrix} E_z & E_w \end{bmatrix}, \\ f &= \begin{bmatrix} f_x \\ f_s \end{bmatrix} & \text{e} & g = \begin{bmatrix} g_z \\ g_w \end{bmatrix}, \end{aligned} \quad (3.11)$$

para  $f_x \in \mathbb{R}^{q_1}$ ,  $f_s \in \mathbb{R}^{q_2}$ ,  $g_z \in \mathbb{R}^{q_3}$  e  $g_w \in \mathbb{R}^{q_4}$ . Na função objetivo, introduzimos a matriz  $D^{-\frac{1}{2}}$  para obtermos a matriz  $ADA^t$  na resolução do problema. Desse modo, a fatoração de Cholesky já calculada é novamente utilizada para resolver os sistemas lineares envolvidos. Os valores de  $f$  e  $g$  são determinados a seguir.

a) Determinando o valor de  $f = [f_x^t \ f_s^t]^t$ .

- Devemos ter  $E_x^t \tilde{d}x = 0$  então  $f_x = 0$ .
- Para obter  $E_s^t \tilde{d}s = 0$ , devemos encontrar os valores das componentes  $E_s^t \tilde{d}x$  correspondentes. De (3.7), temos a relação:

$$\tilde{d}x = r_u - \tilde{d}s.$$

Logo,

$$E_s^t \tilde{d}x = E_s^t r_u - E_s^t \tilde{d}s.$$

Como  $E_s^t \tilde{d}s = 0$ , devemos ter:

$$f_s = E_s^t \tilde{d}x = E_s^t r_u.$$

b) Determinando o valor de  $g = [g_z^t \ g_w^t]^t$ .

- Para que  $E_z^t \tilde{d}z = 0$ , encontramos o valor da componente  $E_z^t \tilde{d}x$  correspondente pela equação  $Z\tilde{d}x + X\tilde{d}z = r_a$  de (3.7). Usando (2.5), temos:

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

Então,

$$\begin{aligned} E_z^t \tilde{d}x &= -E_z^t x - E_z^t Z^{-1} X \tilde{d}z = \\ &= -E_z^t x - E_z^t Z^{-1} X E_z E_z^t \tilde{d}z. \end{aligned}$$

Como  $E_z^t \tilde{d}z = 0$ , devemos ter:

$$g_z = E_z^t \tilde{d}x = -E_z^t x.$$

- Para que  $E_w^t \tilde{d}w = 0$ , encontramos o valor da componente  $E_w^t \tilde{d}x$  correspondente através das equações  $W\tilde{d}s + S\tilde{d}w = r_b$  e  $\tilde{d}x = r_u - \tilde{d}s$  de (3.7). Usando (2.5), temos:

$$\begin{aligned} \tilde{d}s &= W^{-1}(r_b - S\tilde{d}w) = \\ &= W^{-1}(-SWe - S\tilde{d}w) = \\ &= -Se - W^{-1}S\tilde{d}w. \end{aligned}$$

Logo,

$$\begin{aligned} E_w^t \tilde{d}s &= -E_w^t s - E_w^t W^{-1}S\tilde{d}w = \\ &= -E_w^t s - E_w^t W^{-1}SE_w E_w^t \tilde{d}w. \end{aligned}$$

Como  $E_w^t \tilde{d}w = 0$ , devemos ter:

$$E_w^t \tilde{d}s = -E_w^t s.$$

Além disso, como temos  $\tilde{d}x = r_u - \tilde{d}s$ , multiplicando essa equação por  $E_w$  e substituindo o valor de  $E_w^t \tilde{d}s$  encontrado, obtemos:

$$g_w = E_w^t \tilde{d}x = E_w^t (r_u + s).$$

Logo, o problema auxiliar para determinar  $\tilde{d}x$  é dado por (3.10), em que

$$\begin{aligned} f &= \begin{bmatrix} 0 \\ E_s^t r_u \end{bmatrix}; \\ g &= \begin{bmatrix} -E_z^t x \\ E_w^t (r_u + s) \end{bmatrix}. \end{aligned} \tag{3.12}$$

**Observação 3.2.2.** As componentes de bloqueio devem ser diferentes entre si, pois, por exemplo, caso  $\beta$  realizasse bloqueio na direção  $dx$  e  $dz$ , teríamos no problema auxiliar  $dx_\beta = 0$  e  $dx_\beta = -x_\beta$ , o que não pode ocorrer. O mesmo pode ser verificado com as componentes de bloqueio na direção  $ds$  e  $dw$ .

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

$$l(\tilde{d}x, v, p, q) = \frac{1}{2}(D^{-\frac{1}{2}}\tilde{d}x - D^{-\frac{1}{2}}dx)^t (D^{-\frac{1}{2}}\tilde{d}x - D^{-\frac{1}{2}}dx) + v^t (A\tilde{d}x - r_p) + p^t (E_i^t \tilde{d}x - f) + q^t (E_j^t \tilde{d}x - g), \tag{3.13}$$

em que  $v = (v_1, v_2, \dots, v_m)^t \in \mathbb{R}^m$  e  $p \in \mathbb{R}^{q_1+q_2}$ ,  $q \in \mathbb{R}^{q_3+q_4}$ .

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

$$\begin{cases} D^{-1}\tilde{d}x - D^{-1}dx + A^t v + E_i p + E_j q = 0 & (3.14) \end{cases}$$

$$\begin{cases} A\tilde{d}x - r_p = 0 & (3.15) \end{cases}$$

$$\begin{cases} E_i^t \tilde{d}x - f = 0 & (3.16) \end{cases}$$

$$\begin{cases} E_j^t \tilde{d}x - g = 0. & (3.17) \end{cases}$$

Inicialmente, vamos considerar que existe somente uma componente de bloqueio  $i$  no espaço primal e outra  $j$  no espaço dual. Dessa forma,  $E_i = e_i$ ,  $E_j = e_j$  e  $p, q$  são escalares.

Por (3.14),

$$\begin{aligned} D^{-1}\tilde{d}x &= D^{-1}dx - A^t v - p e_i - q e_j. \\ \tilde{d}x &= dx - DA^t v - p D e_i - q D e_j. \end{aligned}$$

Logo,

$$\tilde{d}x = dx - DA^t v - p d_{ii} e_i - q d_{jj} e_j. \quad (3.18)$$

Por (3.15) e (3.18),

$$r_p = A\tilde{d}x = A dx - ADA^t v - p d_{ii} A_i - q d_{jj} A_j.$$

Como  $r_p = A dx$ , temos:

$$ADA^t v = -d_{ii} A_i p - d_{jj} A_j q.$$

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

$$v = -d_{ii} (ADA^t)^{-1} A_i p - d_{jj} (ADA^t)^{-1} A_j q. \quad (3.19)$$

Por (3.16) e (3.18),

$$f = \tilde{d}x_i = dx_i - d_{ii} A_i^t v - p d_{ii}. \quad (3.20)$$

Substituindo o valor de  $v$  em (3.20), temos:

$$f = dx_i - d_{ii} A_i^t (-p d_{ii} (ADA^t)^{-1} A_i - q d_{jj} (ADA^t)^{-1} A_j) - p d_{ii}.$$

$$f = dx_i + p d_{ii}^2 A_i^t (ADA^t)^{-1} A_i + q d_{ii} d_{jj} A_i^t (ADA^t)^{-1} A_j - p d_{ii}.$$

Logo,

$$p d_{ii} - p d_{ii}^2 A_i^t (ADA^t)^{-1} A_i = dx_i + q d_{ii} d_{jj} A_i^t (ADA^t)^{-1} A_j - f.$$

Então,

$$p = \frac{dx_i - f + qd_{ii}d_{jj}A_i^t(ADA^t)^{-1}A_j}{d_{ii}(1 - d_{ii}A_i^t(ADA^t)^{-1}A_i)}. \quad (3.21)$$

Por (3.17) e (3.18),

$$g = \tilde{d}x_j = dx_j - d_{jj}A_j^t v - qd_{jj}. \quad (3.22)$$

Substituindo o valor de  $v$ , dado por (3.19), em (3.22) temos:

$$g = dx_j - d_{jj}A_j^t(-pd_{ii}(ADA^t)^{-1}A_i - qd_{jj}(ADA^t)^{-1}A_j) - qd_{jj}.$$

Logo,

$$g = dx_j + pd_{ii}d_{jj}A_j^t(ADA^t)^{-1}A_i + qd_{jj}^2A_j^t(ADA^t)^{-1}A_j - qd_{jj}. \quad (3.23)$$

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 (3.23) e substituindo o valor de  $p$ , dado por (3.21), obtemos:

$$\begin{aligned} g &= dx_j + \left[ \frac{dx_i - f + qd_{ii}d_{jj}\theta_{ij}}{d_{ii}(1 - d_{ii}\theta_i)} \right] d_{ii}d_{jj}\theta_{ij} + qd_{jj}^2\theta_j - qd_{jj}. \\ g &= dx_j + \left[ \frac{(dx_i - f)d_{jj}\theta_{ij}}{1 - d_{ii}\theta_i} \right] + \left[ \frac{qd_{ii}d_{jj}^2\theta_{ij}^2}{1 - d_{ii}\theta_i} \right] + qd_{jj}^2\theta_j - qd_{jj}. \end{aligned}$$

Assim,

$$qd_{jj} - q \left[ \frac{d_{ii}d_{jj}^2\theta_{ij}^2}{1 - d_{ii}\theta_i} \right] - qd_{jj}^2\theta_j = dx_j - g + \left[ \frac{(dx_i - f)d_{jj}\theta_{ij}}{1 - d_{ii}\theta_i} \right]. \quad (3.24)$$

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 (3.24), temos:

$$qd_{jj}\xi = dx_j - g + \left[ \frac{(dx_i - f)d_{jj}\theta_{ij}}{1 - d_{ii}\theta_i} \right].$$

Logo,

$$q = \frac{dx_j - g + \left[ \frac{(dx_i - f)d_{jj}\theta_{ij}}{1 - d_{ii}\theta_i} \right]}{d_{jj}\xi}.$$

Então, teremos 2 sistemas lineares envolvendo a matriz  $ADA^t$  no cálculo de  $\tilde{d}x$ . São eles:

$$\begin{aligned}v_1 &= (ADA^t)^{-1}A_i \\v_2 &= (ADA^t)^{-1}A_j.\end{aligned}$$

Considere agora o caso geral em que podemos ter mais de uma componente de bloqueio primal e/ou dual.

Por (3.14),

$$D^{-1}\tilde{d}x = D^{-1}dx - A^t v - E_i p - E_j q.$$

Logo,

$$\tilde{d}x = dx - DA^t v - DE_i p - DE_j q. \quad (3.25)$$

Por (3.16) e (3.25),

$$f = E_i^t \tilde{d}x = E_i^t dx - E_i^t DA^t v - E_i^t DE_i p, \quad (3.26)$$

pois  $E_i^t DE_j = 0$ . Então,

$$E_i^t DE_i p = E_i^t dx - f - E_i^t DA^t v.$$

Temos que  $\tilde{D}_i = E_i^t DE_i$  é uma matriz diagonal de ordem  $(q_1 + q_2) \times (q_1 + q_2)$  sendo  $(\tilde{d}_i)_l = d_l$ , em que  $l$  é a componente que realiza bloqueio no espaço primal ( $l \in L_1$  e/ou  $l \in L_2$ ). Como  $\tilde{D}_i$  é uma matriz diagonal não nula, então existe  $\tilde{D}_i^{-1}$ . Assim:

$$p = \tilde{D}_i^{-1}(E_i^t dx - f - E_i^t DA^t v). \quad (3.27)$$

Por (3.17) e (3.25),

$$g = E_j^t \tilde{d}x = E_j^t dx - E_j^t DA^t v - E_j^t DE_j q. \quad (3.28)$$

Temos que  $\tilde{D}_j = E_j^t DE_j$ , é uma matriz diagonal de ordem  $(q_3 + q_4) \times (q_3 + q_4)$  sendo  $(\tilde{d}_j)_l = d_l$ , em que  $l$  é a componente que realiza bloqueio no espaço dual ( $l \in L_3$  e/ou  $l \in L_4$ ). Assim,  $\tilde{D}_j$  é uma matriz diagonal não nula, logo existe  $\tilde{D}_j^{-1}$ , então:

$$q = \tilde{D}_j^{-1}(E_j^t dx - g - E_j^t DA^t v). \quad (3.29)$$

Por (3.15) e (3.25),

$$r_p = A\tilde{d}x = Adx - ADA^t v - ADE_i p - ADE_j q.$$

Como  $r_p = Adx$ , temos:

$$ADA^t v = -ADE_i p - ADE_j q. \quad (3.30)$$

Substituindo os valores de  $p$  (3.27) e  $q$  (3.29) em (3.30), obtemos:

$$\begin{aligned} ADA^t v &= -ADE_i \tilde{D}_i^{-1} (E_i^t dx - f - E_i^t DA^t v) - ADE_j \tilde{D}_j^{-1} (E_j dx - g - E_j DA^t v) = \\ &= [ADE_i \tilde{D}_i^{-1} E_i^t DA^t + ADE_j \tilde{D}_j^{-1} E_j DA^t] v - ADE_i \tilde{D}_i^{-1} (E_i^t dx - f) - \\ &\quad - ADE_j \tilde{D}_j^{-1} (E_j dx - g). \end{aligned} \quad (3.31)$$

Logo,

$$\begin{aligned} [ADA^t - A(DE_i \tilde{D}_i^{-1} E_i^t D + DE_j \tilde{D}_j^{-1} E_j D) A^t] v &= -ADE_i \tilde{D}_i^{-1} (E_i^t dx - f) - \\ &\quad - ADE_j \tilde{D}_j^{-1} (E_j dx - g). \end{aligned}$$

Considere:

$$\hat{D} = DE_i \tilde{D}_i^{-1} E_i^t D + DE_j \tilde{D}_j^{-1} E_j D. \quad (3.32)$$

Então:

$$[ADA^t - A\hat{D}A^t] v = -ADE_i \tilde{D}_i^{-1} (E_i^t dx - f) - ADE_j \tilde{D}_j^{-1} (E_j dx - g). \quad (3.33)$$

Temos que:

$$\tilde{D}_i^{-1} = \text{diag}(1/(\tilde{d}_i)_u), \quad l = 1, \dots, (q_1 + q_2) \quad \text{e} \quad \tilde{D}_j^{-1} = \text{diag}(1/(\tilde{d}_j)_u), \quad l = 1, \dots, (q_3 + q_4).$$

Logo,  $\hat{D}_i = DE_i \tilde{D}_i^{-1} E_i^t D$  é uma matriz diagonal de ordem  $n \times n$  sendo que:

$$(\hat{d}_i)_u = \begin{cases} d_u, & \text{se } l \text{ realiza bloqueio primal;} \\ 0, & \text{caso contrário,} \end{cases} \quad (3.34)$$

em que  $l = 1, \dots, n$ . Da mesma forma,  $\hat{D}_j = DE_j \tilde{D}_j^{-1} E_j^t D$  é uma matriz diagonal de ordem  $n \times n$  sendo que:

$$(\hat{d}_j)_u = \begin{cases} d_u, & \text{se } l \text{ realiza bloqueio dual;} \\ 0, & \text{caso contrário,} \end{cases} \quad (3.35)$$

em que  $l = 1, \dots, n$ .

Podemos expressar  $\hat{D}_i$  e  $\hat{D}_j$  como:

$$\hat{D}_i = E_i \tilde{D}_i E_i^t,$$

$$\hat{D}_j = E_j \tilde{D}_j E_j^t.$$

Assim, podemos reescrever (3.32) da seguinte forma:

$$\hat{D} = E_i \tilde{D}_i E_i^t + E_j \tilde{D}_j E_j^t = E \tilde{D} E^t, \quad (3.36)$$

em que

$$E = \begin{bmatrix} E_i & E_j \end{bmatrix} \text{ e } \tilde{D} = \begin{bmatrix} \tilde{D}_i & 0 \\ 0 & \tilde{D}_j \end{bmatrix}.$$

Substituindo (3.36) em (3.33), temos a seguinte matriz de coeficientes:

$$ADA^t - A\hat{D}A^t = ADA^t - AE\tilde{D}E^tA^t.$$

**Observação 3.2.3.** A matriz  $\bar{D} = D - \hat{D}$  é uma matriz diagonal formada pelos elementos da matriz  $D$  exceto que  $(\bar{d})_{ll} = 0$  se  $l$  é componente de bloqueio primal ou dual. Na prática, é improvável existirem inúmeras componentes de bloqueio, de forma que  $A\bar{D}A^t = ADA^t - A\hat{D}A^t$  é não singular.

Pela Fórmula de Sherman-Morrison-Woodburry, se  $U$  e  $V$  são matrizes de ordem  $n \times k$  tais que  $(I + V^tC^{-1}U)^{-1}$  existe, então:

$$(C + UV^t)^{-1} = C^{-1} - C^{-1}U(I + V^tC^{-1}U)^{-1}V^tC^{-1}. \quad (3.37)$$

Assim, retornando à equação (3.33), como a matriz  $ADA^t$  é simétrica definida positiva e admite a fatoração de Cholesky, tomando

$$U = -AE\tilde{D}^{\frac{1}{2}},$$

$$V = AE\tilde{D}^{\frac{1}{2}},$$

e considerando  $B = ADA^t$  e  $I$  a matriz identidade de ordem  $(q_1 + q_2 + q_3 + q_4)$ , pela fórmula (3.37), temos que:

$$(ADA^t - A\hat{D}A^t)^{-1} = B^{-1} + B^{-1}AE\tilde{D}^{\frac{1}{2}} \left( I - \tilde{D}^{\frac{1}{2}}E^tA^tB^{-1}AE\tilde{D}^{\frac{1}{2}} \right)^{-1} \tilde{D}^{\frac{1}{2}}EA^tB^{-1}.$$

**Observação 3.2.4.** Como  $A\bar{D}A^t = ADA^t - A\hat{D}A^t$  é não singular, então  $(I - \tilde{D}^{\frac{1}{2}}E^tA^tB^{-1}AE\tilde{D}^{\frac{1}{2}})$  é não singular. Para calcular a matriz inversa desta matriz simétrica e indefinida pode ser utilizada a fatoração de Bunch-Parlett [8] do tipo  $L\dot{D}L^t$ , em que  $\dot{D}$  é uma matriz bloco diagonal, com blocos 1x1 e 2x2.

Dessa forma, a solução para (3.33) é dada por:

$$v = - \left[ B^{-1} + B^{-1}AE\tilde{D}^{\frac{1}{2}} \left( I - \tilde{D}^{\frac{1}{2}}E^tA^tB^{-1}AE\tilde{D}^{\frac{1}{2}} \right)^{-1} \tilde{D}^{\frac{1}{2}}EA^tB^{-1} \right] \bar{w}, \quad (3.38)$$

em que  $\bar{w} = [ADE_i\tilde{D}_i(E_i^tdx - f) - ADE_j\tilde{D}_j(E_j^tdx - g)]$ .

Assim, temos o seguinte sistema linear a ser resolvido:

$$\bar{v}_1 = B^{-1}\bar{w},$$

e para cada coluna  $\tilde{A}_l$  da matriz  $\tilde{A} = AED\tilde{D}^{\frac{1}{2}}$ ,  $l = 1, \dots, (q_1 + q_2 + q_3 + q_4)$ , no cálculo de  $(I - \tilde{D}^{\frac{1}{2}}E^tA^tB^{-1}AED\tilde{D}^{\frac{1}{2}})^{-1}$ , resolvemos:

$$\bar{v}_l = B^{-1}\tilde{A}_l,$$

ou seja, teremos no total  $(q_1 + q_2 + q_3 + q_4 + 1)$  sistemas lineares a serem resolvidos com a matriz  $B = ADA^t$ , no cálculo de  $v$ .

Portanto, a partir de (3.13), encontramos  $\tilde{d}x$ , um ponto crítico para  $l(\tilde{d}x, v, p, q)$ . Analisaremos a matriz Hessiana da função Lagrangiana neste ponto, para verificar que  $\tilde{d}x$  é o mínimo para (3.10).

Considere:

$$\begin{aligned} f(\tilde{d}x) &= \frac{1}{2} \|D^{-\frac{1}{2}}\tilde{d}x - D^{-\frac{1}{2}}dx\|^2, \\ h_{t_1}(\tilde{d}x) &= \tilde{d}x_{t_1} - (f)_{t_1}, \\ h_{t_2}(\tilde{d}x) &= \tilde{d}x_{t_2} - (g)_{t_2}, \\ g_k(\tilde{d}x) &= A_k\tilde{d}x - (r_p)_k, \end{aligned}$$

em que  $A_k$  é a linha  $k$  da matriz  $A$ ,  $1 \leq k \leq m$ ,  $1 \leq t_1 \leq q_1 + q_2$  e  $1 \leq t_2 \leq q_3 + q_4$ . Então, a matriz Hessiana de  $l(\tilde{d}x, v, p, q)$  é dada por:

$$L(\tilde{d}x) = F(\tilde{d}x) + \sum_{k=1}^m v_k G_k(\tilde{d}x) + \sum_{t_1=1}^{q_1+q_2} p_{t_1} H_{t_1}(\tilde{d}x) + \sum_{t_2=1}^{q_3+q_4} q_{t_2} H_{t_2}(\tilde{d}x),$$

em que  $F(\tilde{d}x)$ ,  $G_k(\tilde{d}x)$ ,  $H_{t_1}(\tilde{d}x)$  e  $H_{t_2}(\tilde{d}x)$  são matrizes Hessianas de  $f(\tilde{d}x)$ ,  $g_k(\tilde{d}x)$ ,  $h_{t_1}(\tilde{d}x)$  e  $h_{t_2}(\tilde{d}x)$ , respectivamente.

Temos que:

$$\nabla f(\tilde{d}x) = D \begin{pmatrix} \tilde{d}x_1 - dx_1 \\ \tilde{d}x_2 - dx_2 \\ \vdots \\ \tilde{d}x_n - dx_n \end{pmatrix}, \quad \nabla g_k(\tilde{d}x) = (A_k)^t, \quad \nabla h_{t_1}(\tilde{d}x) = e_{t_1}, \quad \nabla h_{t_2}(\tilde{d}x) = e_{t_2}.$$

Logo,

$$F(\tilde{d}x) = D, \quad G_k(\tilde{d}x) = H_{t_1}(\tilde{d}x) = H_{t_2}(\tilde{d}x) = 0_{n \times n},$$

para todo  $1 \leq k \leq m$ ,  $1 \leq t_1 \leq q_1 + q_2$  e  $1 \leq t_2 \leq q_3 + q_4$ . Assim,  $L(\tilde{d}x) = F(\tilde{d}x)$ . Como  $F(\tilde{d}x)$  é definida positiva para todo valor de  $\tilde{d}x$ , temos que  $\tilde{d}x$  é um ponto de mínimo global. Portanto, a direção  $\tilde{d}x$  encontrada é a solução mínima do problema.

Retornando ao sistema (3.7), encontramos as direções:

$$\begin{aligned} \tilde{d}s &= r_u - \tilde{d}x, \\ \tilde{d}z &= X^{-1}(r_a - Z\tilde{d}x), \\ \tilde{d}w &= S^{-1}(r_b - W\tilde{d}s). \end{aligned}$$

Além disso, encontramos a equação  $A^t \tilde{d}y = r_d - \tilde{d}z + \tilde{d}w$  que, multiplicando por  $D^{\frac{1}{2}}$  e resolvendo por quadrados mínimos, resulta em:

$$\tilde{d}y = (ADA^t)^{-1}AD(r_d - \tilde{d}z + \tilde{d}w).$$

### Direção corretora continuada

Calculando a perturbação  $\mu$ , realizando a correção não linear e resolvendo o sistema linear (2.23) aproximadamente, com  $E_x^t \hat{d}x = 0$ ,  $E_s^t \hat{d}s = 0$ ,  $E_z^t \hat{d}z = 0$  e  $E_w^t \hat{d}w = 0$ , determinamos a direção preditora corretora continuada  $\hat{d}$ . Então, a direção  $\hat{d}$  satisfaz aproximadamente o seguinte sistema linear:

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{bmatrix} \begin{bmatrix} \hat{d}x \\ \hat{d}s \\ \hat{d}y \\ \hat{d}z \\ \hat{d}w \end{bmatrix} \simeq \begin{bmatrix} r_p \\ r_u \\ r_d \\ r_s \\ r_t \end{bmatrix}, \quad (3.39)$$

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

$$\begin{aligned} E_x^t \hat{d}x &= 0, & E_s^t \hat{d}s &= 0 \\ E_z^t \hat{d}z &= 0 & \text{e} & E_w^t \hat{d}w = 0. \end{aligned}$$

De modo análogo ao sistema (3.7), primeiramente, determinamos  $\hat{d}x$  utilizando um problema auxiliar do tipo (3.10), sendo os valores  $f$  e  $g$  determinados a seguir.

a) Determinando o valor de  $f$ .

- Para obter  $E_x^t \hat{d}x = 0$ , tomamos  $f_x = 0$ ;
- Para que  $E_s^t \hat{d}s = 0$ , devemos encontrar o valor da componente  $E_s^t \hat{d}x$  correspondente. De (3.39), temos a equação:

$$\hat{d}x = r_u - \hat{d}s.$$

Logo,

$$E_s^t \hat{d}x = E_s^t r_u - E_s^t \hat{d}s.$$

Como  $E_s^t \hat{d}s = 0$ , devemos ter:

$$f_s = E_s^t \hat{d}x = E_s^t r_u.$$

b) Determinando o valor de  $g$ .

- Para que  $E_z^t \hat{d}z = 0$ , encontramos o valor da componente  $E_z \hat{d}x$  correspondente pela equação  $Z\hat{d}x + X\hat{d}z = r_s$ . De (3.39) e (2.23), temos:

$$\begin{aligned} \hat{d}x &= Z^{-1}(r_s - X\hat{d}z) = \\ &= Z^{-1}(r_a + \mu e - \tilde{D}_x \tilde{D}_z e - X\hat{d}z) = \\ &= Z^{-1}(-XZe + \mu e - \tilde{D}_x \tilde{D}_z e - X\hat{d}z) = \\ &= Z^{-1}(-XZe - X\hat{d}z) + Z^{-1}(\mu e - \tilde{D}_x \tilde{D}_z e) = \\ &= -Xe - Z^{-1}X\hat{d}z + Z^{-1}(\mu e - \tilde{D}_x \tilde{D}_z e). \end{aligned}$$

Logo,

$$E_z^t \hat{d}x = -E_z^t X e - E_z^t Z^{-1} X E_z E_z^t \hat{d}z + E_z^t Z^{-1} \mu e - E_z^t Z^{-1} \tilde{D}_x E_z E_z^t \tilde{D}_z e.$$

Como  $E_z^t \hat{d}z = 0$ , devemos ter:

$$g = E_z^t \hat{d}x = -E_z^t x + E_z^t Z^{-1} \mu e - E_z^t Z^{-1} \tilde{D}_x E_z E_z^t \tilde{d}z.$$

Já que temos  $E_z^t \tilde{d}z = 0$ , segue que:

$$g_z = -E_z^t x + \mu E_z^t Z^{-1} e.$$

- Para que  $E_w^t \hat{d}w = 0$ , encontramos o valor da componente  $E_w^t \hat{d}x$  correspondente através das equações  $W\hat{d}s + S\hat{d}w = r_t$  e  $\hat{d}x = r_u - \hat{d}s$  de (3.39). De (2.23), temos:

$$\begin{aligned} \hat{d}s &= W^{-1}(r_t - S\hat{d}w) = \\ &= W^{-1}(r_b + \mu e - \tilde{D}_s \tilde{D}_w e - S\hat{d}w) = \\ &= W^{-1}(-SWe + \mu e - \tilde{D}_s \tilde{D}_w e - S\hat{d}w) = \\ &= W^{-1}(-SWe - S\hat{d}w) + W^{-1}(\mu e - \tilde{D}_s \tilde{D}_w e) = \\ &= -Se - W^{-1}S\hat{d}w + W^{-1}(\mu e - \tilde{D}_s \tilde{D}_w e). \end{aligned}$$

Então,

$$E_w^t \hat{d}s = -E_w^t S e - E_w^t W^{-1} S E_w E_w^t \hat{d}w + E_w^t W^{-1} \mu e - E_w^t W^{-1} \tilde{D}_s E_w E_w^t \tilde{d}w.$$

Como  $E_w^t \hat{d}w = 0$  e  $E_w^t \tilde{d}w = 0$ , devemos ter:

$$E_w^t \hat{d}s = -E_w^t S e + \mu E_w^t W^{-1} e.$$

Como  $\hat{d}x = r_u - \hat{d}s$ , então  $E_w^t \hat{d}x = E_w^t(r_u - \hat{d}s)$ . Substituindo o valor de  $E_w^t \hat{d}s$  encontrado, obtemos:

$$g_w = E_w^t \hat{d}x = E_w^t r_u + E_w^t S e - \mu E_w^t W^{-1} e.$$

Logo, o problema auxiliar para determinar  $\hat{d}x$  é dado por:

$$\begin{aligned} \min \quad & \frac{1}{2} \|D^{-\frac{1}{2}}\hat{d}x - D^{-\frac{1}{2}}dx\|^2 \\ \text{s.a} \quad & A\hat{d}x = r_p \\ & E_i\hat{d}x = f \\ & E_j^t\hat{d}x = g, \end{aligned} \tag{3.40}$$

em que

$$\begin{aligned} f &= \begin{bmatrix} 0 \\ E_s^t r_u \end{bmatrix}; \\ g &= \begin{bmatrix} -E_z^t x + \mu E_z^t Z^{-1} e \\ E_w^t r_u + E_w^t S e - \mu E_w^t W^{-1} e \end{bmatrix}. \end{aligned} \tag{3.41}$$

Observe que o problema (3.40) é similar a (3.10), sendo alterados apenas os valores  $f$  e  $g$ . Logo, calculamos  $\hat{d}x$  da mesma forma, obtendo a seguinte solução:

$$\hat{d}x = dx - DA^t v - DE_i p - DE_j q,$$

$$p = (E_i^t DE_i)^{-1} (E_i^t dx - f - E_i^t DA^t v),$$

$$q = (E_j^t DE_j)^{-1} (E_j^t dx - g - E_j^t DA^t v),$$

$$v = - \left[ B^{-1} + B^{-1} A E \tilde{D}^{\frac{1}{2}} \left( I - \tilde{D}^{\frac{1}{2}} E^t A^t B^{-1} A E \tilde{D}^{\frac{1}{2}} \right)^{-1} \tilde{D}^{\frac{1}{2}} E A^t B^{-1} \right] \bar{w},$$

$$\bar{w} = [A D E_i \tilde{D}_i (E_i^t dx - f) - A D E_j \tilde{D}_j (E_j^t dx - g)].$$

Neste caso, é necessário somente resolver novamente o sistema linear envolvendo a matriz  $B^{-1} = (A D A^t)^{-1}$ , para calcular  $\bar{v}_1 = B^{-1} \bar{w}$ , já que  $f$  e  $g$  são alterados. Os sistemas lineares  $\bar{v}_l = B^{-1} \tilde{A}_l$ , que são resolvidos no cálculo de  $(I - \tilde{D}^{\frac{1}{2}} E^t A^t B^{-1} A E \tilde{D}^{\frac{1}{2}})^{-1}$ , na direção preditora continuada, não são alterados. Assim, não é necessário resolvê-los novamente.

Retornando ao sistema (3.39), encontramos as direções:

$$\begin{aligned} \hat{d}s &= r_u - \hat{d}x, \\ \hat{d}z &= X^{-1} (r_s - Z \hat{d}x), \\ \hat{d}w &= S^{-1} (r_t - W \hat{d}s). \end{aligned}$$

Além disso, encontramos a equação  $A^t \hat{d}y = r_d - \hat{d}z + \hat{d}w$  que, multiplicando por  $D^{\frac{1}{2}}$  e resolvendo por quadrados mínimos, fornece a solução:

$$\hat{d}y = (A D A^t)^{-1} A D (r_d - \hat{d}z + \hat{d}w).$$

Portanto, a direção preditora corretora continuada é solução do seguinte sistema linear:

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{bmatrix} \begin{bmatrix} \hat{d}x \\ \hat{d}s \\ \hat{d}y \\ \hat{d}z \\ \hat{d}w \end{bmatrix} = \begin{bmatrix} r_p \\ r_u \\ r_d \\ r_s \\ r_t \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ r_3 \\ 0 \\ 0 \end{bmatrix}, \quad (3.42)$$

em que  $r_3 = r_d - A^t \hat{d}y - \hat{d}z + \hat{d}w$ , ou seja, o resíduo da projeção de quadrados mínimos.

### 3.2.2 Direção Moderada

A direção continuada  $\hat{d}$ , definida nesta subseção, é determinada de forma que a alteração para a direção preditora corretora  $d$  calculada seja a menor possível. Assim, o sistema linear (2.23) que calcula a direção  $d$  é resolvido novamente, de forma aproximada, para determinar  $\hat{d}$ . Além disso, devemos ter  $E_x^t \hat{d}x = 0$ ,  $E_s^t \hat{d}s = 0$ ,  $E_z^t \hat{d}z = 0$  e  $E_w^t \hat{d}w = 0$ . A proposta dessa direção é similar à direção preditora corretora continuada, porém acreditamos que não é preciso calcular uma nova direção do tipo preditora corretora, pois já temos  $d$  que é uma direção deste tipo. Assim, somente um sistema linear adicional é resolvido para determinar  $\hat{d}$ . Esta direção continuada é denominada moderada, sendo solução aproximada do sistema linear (3.39).

Assim, calculamos a direção moderada da mesma forma, utilizando o problema auxiliar (3.40) em que  $D = (X^{-1}Z + S^{-1}W)^{-1}$ ;  $\hat{d}x, dx \in \mathbb{R}^n$ ;  $E_i \in \mathbb{R}^{n \times (q_1+q_2)}$  e  $E_j \in \mathbb{R}^{n \times (q_3+q_4)}$ ;

$$f = \begin{bmatrix} 0 \\ E_s^t r_u \end{bmatrix};$$

$$g = \begin{bmatrix} E_z^t Z^{-1} r_s \\ E_w^t r_u - E_w^t W^{-1} r_t \end{bmatrix}.$$

A solução é dada por:

$$\hat{d}x = dx - DA^t v - DE_i p - DE_j q,$$

$$p = (E_i^t DE_i)^{-1} (E_i^t dx - f - E_i^t DA^t v),$$

$$q = (E_j^t DE_j)^{-1} (E_j^t dx - g - E_j^t DA^t v),$$

$$v = - \left[ B^{-1} + B^{-1} AE \tilde{D}^{\frac{1}{2}} \left( I - \tilde{D}^{\frac{1}{2}} E^t A^t B^{-1} AE \tilde{D}^{\frac{1}{2}} \right)^{-1} \tilde{D}^{\frac{1}{2}} E A^t B^{-1} \right] \bar{w},$$

$$\bar{w} = [ADE_i \tilde{D}_i (E_i^t dx - f) - ADE_j \tilde{D}_j (E_j^t dx - g)].$$

Como visto na Subseção 3.2.1, no caso de haver somente uma componente de bloqueio primal e dual, dois sistemas lineares adicionais envolvendo a matriz  $ADA^t$  são resolvidos no cálculo de  $v$ . No entanto, no caso em que há mais de uma componente de bloqueio em cada espaço, temos  $(q_1 + q_2 + q_3 + q_4 + 1)$  sistemas lineares adicionais a serem resolvidos.

As demais direções são dadas por:

$$\hat{d}s = r_u - \hat{d}x,$$

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

$$\hat{d}w = S^{-1}(r_b - W\hat{d}s),$$

$$\hat{d}y = (ADA^t)^{-1}AD(r_d - \hat{d}z + \hat{d}w).$$

A direção moderada pode ser interpretada como solução do sistema linear (3.42), em que  $r_3$  é o resíduo da projeção de quadrados mínimos para determinar  $\hat{d}y^k$ .

### 3.2.3 Direção Simples

Sejam a direção preditora corretora ( $d$ ) e as matrizes  $E_x$ ,  $E_s$ ,  $E_z$  e  $E_w$  já determinadas. A direção simples  $\hat{d}$  é a mesma  $d$ , porém fixando em zero as componentes de bloqueio da direção  $d$ , ou seja, considerando  $E_x^t \hat{d}x = 0$ ,  $E_s^t \hat{d}s = 0$ ,  $E_z^t \hat{d}z = 0$  e  $E_w^t \hat{d}w = 0$ .

Assim, a direção  $\hat{d} = (\hat{d}x, \hat{d}s, \hat{d}y, \hat{d}z, \hat{d}w)$  é determinada da seguinte forma:

$$\hat{d}x = dx - E_x E_x^t dx,$$

$$\hat{d}s = ds - E_s E_s^t ds,$$

$$\hat{d}z = dz - E_z E_z^t dz,$$

$$\hat{d}w = dw - E_w E_w^t dw,$$

$$\hat{d}y = dy.$$

Para determinar a direção simples não é necessário resolver sistemas lineares adicionais com a matriz  $ADA^t$ , ao contrário da direção moderada. Além disso, esta direção

pode ser interpretada como a solução do seguinte sistema linear:

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{bmatrix} \begin{bmatrix} \hat{d}x \\ \hat{d}s \\ \hat{d}y \\ \hat{d}z \\ \hat{d}w \end{bmatrix} = \begin{bmatrix} r_p \\ r_u \\ r_d \\ r_s \\ r_t \end{bmatrix} - \begin{bmatrix} AE_x E_x^t dx \\ E_x E_x^t dx + E_s E_s^t ds \\ E_z E_z^t dz - E_w E_w^t dw \\ ZE_x E_x^t dx + XE_z E_z^t dz \\ WE_s E_s^t ds + SE_w E_w^t dw \end{bmatrix}.$$

Portanto, propomos direções continuadas  $\hat{d}$  como soluções aproximadas do sistema linear que possibilitou o cálculo da direção preditora corretora, ou seja:

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{bmatrix} \begin{bmatrix} \hat{d}x \\ \hat{d}s \\ \hat{d}y \\ \hat{d}z \\ \hat{d}w \end{bmatrix} = \begin{bmatrix} r_p \\ r_u \\ r_d \\ r_s \\ r_t \end{bmatrix} + \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{bmatrix},$$

com  $E_x^t \hat{d}x = 0$ ,  $E_s^t \hat{d}s = 0$ ,  $E_z^t \hat{d}z = 0$  e  $E_w^t \hat{d}w = 0$ .

Logo, as soluções  $\hat{d}$  são dadas da seguinte forma:

$$(\hat{d}x, \hat{d}s) = (dx, ds) - (E_x E_x^t dx, E_s E_s^t ds) + (\epsilon_x, \epsilon_s), \quad (3.43)$$

$$(\hat{d}y, \hat{d}z, \hat{d}w) = (dy, dz, dw) - (0, E_z E_z^t dz, E_w E_w^t dw) + (\epsilon_y, \epsilon_z, \epsilon_w),$$

em que  $\epsilon_x, \epsilon_s, \epsilon_z, \epsilon_w \in \mathbb{R}^n$ ,  $E_x^t \epsilon_x = 0$ ,  $E_s^t \epsilon_s = 0$ ,  $E_z^t \epsilon_z = 0$  e  $E_w^t \epsilon_w = 0$ .

No caso da direção moderada, temos  $\|r_1\| = \|r_2\| = \|r_4\| = \|r_5\| = 0$  e  $\|r_3\| \neq 0$ . Além disso,  $\|\epsilon_x\| \neq 0$ ,  $\|\epsilon_s\| \neq 0$ ,  $\|\epsilon_z\| \neq 0$ ,  $\|\epsilon_w\| \neq 0$ . Por outro lado, na direção simples, temos  $\|r_1\| \neq 0$ ,  $\|r_2\| \neq 0$ ,  $\|r_3\| \neq 0$ ,  $\|r_4\| \neq 0$ ,  $\|r_5\| \neq 0$  e  $\epsilon_x = \epsilon_s = \epsilon_z = \epsilon_w = 0$ .

### 3.3 Formas de utilização da direção continuada

Com as direções continuadas definidas, apresentamos duas formas de utilizá-las e avaliamos o novo ponto ao utilizar a nova direção resultante da combinação da direção preditora corretora e a direção continuada para obter os novos pontos primal e dual dados em (3.2) e (3.3). Se uma melhoria no método é obtida, a direção continuada é utilizada e, em seguida, as múltiplas correções de centralidade são realizadas, as quais fornecem mais um possível aumento dos tamanhos dos passos.

#### 3.3.1 Forma avançada

Na forma avançada, avaliamos o novo ponto resultante ao utilizar uma nova direção  $d$ , dada pela combinação da direção preditora corretora e a direção continuada

$(d = d + \hat{d})$  da seguinte forma:

$$(\hat{x}^{k+1}, \hat{s}^{k+1}) = (x^k, s^k) + \hat{\alpha}_p(dx^k + \hat{d}x^k, ds^k + \hat{d}s^k), \quad (3.44)$$

$$(\hat{y}^{k+1}, \hat{z}^{k+1}, \hat{w}^{k+1}) = (y^k, z^k, w^k) + \hat{\alpha}_d(dy^k + \hat{d}y^k, dz^k + \hat{d}z^k, dw^k + \hat{d}w^k).$$

Para a direção  $\hat{d}$  dada em (3.43), obtemos o ponto primal:

$$\begin{aligned} \hat{x}^{k+1} &= x^k + \hat{\alpha}_p(2dx - E_x E_x^t dx + \epsilon_x) = \\ &= x^k + 2\hat{\alpha}_p \left( I - \frac{1}{2} E_x E_x^t \right) dx + \hat{\alpha}_p \epsilon_x = \\ &= x^k + 2\hat{\alpha}_p \left[ I - \left( 1 - \frac{\hat{\alpha}_p}{2\hat{\alpha}_p} E_x E_x^t \right) dx \right] + \hat{\alpha}_p \epsilon_x. \end{aligned} \quad (3.45)$$

De modo análogo, obtemos os demais pontos  $s$ ,  $z$ ,  $w$ :

$$\begin{aligned} \hat{s}^{k+1} &= s^k + 2\hat{\alpha}_p \left[ I - \left( 1 - \frac{\hat{\alpha}_p}{2\hat{\alpha}_p} \right) E_s E_s^t \right] ds + \hat{\alpha}_p \epsilon_s. \\ \hat{z}^{k+1} &= z^k + 2\hat{\alpha}_d \left[ I - \left( 1 - \frac{\hat{\alpha}_d}{2\hat{\alpha}_d} \right) E_z E_z^t \right] dz + \hat{\alpha}_d \epsilon_z. \\ \hat{w}^{k+1} &= w^k + 2\hat{\alpha}_d \left[ I - \left( 1 - \frac{\hat{\alpha}_d}{2\hat{\alpha}_d} \right) E_w E_w^t \right] dw + \hat{\alpha}_d \epsilon_w. \end{aligned} \quad (3.46)$$

Comparando (3.45) e (3.46) com (3.2) e (3.3), temos:

$$\theta_p^1 = \hat{\alpha}_p,$$

$$\theta_p^2 = 2\hat{\alpha}_p$$

e

$$\theta_d^1 = \hat{\alpha}_d,$$

$$\theta_d^2 = 2\hat{\alpha}_d.$$

Desse modo, analisamos os valores de  $\hat{\alpha}_p$ , verificando se  $\theta_p^1 \leq \alpha_p$  e  $\theta_p^2 > \alpha_p$ . O mesmo é feito para  $\hat{\alpha}_d$ .

Análise do tamanho do passo  $\hat{\alpha}_p$

Por (3.45) e (3.46), podemos escrever o ponto primal na seguinte forma:

$$\hat{x}_l^{k+1} = \begin{cases} x_{l_1}^k + \hat{\alpha}_p dx_{l_1}^k, & \text{se } l = l_1 \in L_1 \text{ realizou bloqueio em } dx, \\ x_l^k + 2\hat{\alpha}_p dx_l^k + \hat{\alpha}_p(\epsilon_x)_l, & \text{caso contrário,} \end{cases} \quad (3.47)$$

$$\hat{s}_l^{k+1} = \begin{cases} s_{l_2}^k + \hat{\alpha}_p ds_{l_2}^k, & \text{se } l = l_2 \in L_2 \text{ realizou bloqueio em } ds, \\ s_l^k + 2\hat{\alpha}_p ds_l^k + \hat{\alpha}_p(\epsilon_s)_l & \text{caso contrário,} \end{cases}$$

Para manter os pontos primais interiores, o tamanho de passo  $\hat{\alpha}_p$  é determinado da seguinte forma:

$$\hat{\alpha}_p = \tau \max \left\{ \hat{\alpha} : x^k + \hat{\alpha}(dx^k + \hat{d}x^k) > 0 \text{ e } s^k + \hat{\alpha}(ds^k + \hat{d}s^k) > 0 \right\}. \quad (3.48)$$

Assim, para o ponto primal dado em (3.47), podemos reescrever (3.48) como

$$\hat{\alpha}_p = \tau \min \{ \rho_x, \rho_s \},$$

em que

$$\rho_x = \min \left\{ -\frac{x_{l_1}^k}{dx_{l_1}^k}, \min_{\substack{1 \leq l \leq n, \\ l \neq l_1}} \left\{ -\frac{x_l^k}{2dx_l^k + \epsilon_x} : x_l^k + 2dx_l^k + (\epsilon_x)_l \leq 0 \right\} \right\}, \quad (3.49)$$

$$\rho_s = \min \left\{ -\frac{s_{l_2}^k}{ds_{l_2}^k}, \min_{\substack{1 \leq l \leq n, \\ l \neq l_2}} \left\{ -\frac{s_l^k}{2ds_l^k + \epsilon_s} : s_l^k + 2ds_l^k + (\epsilon_s)_l \leq 0 \right\} \right\}.$$

**Observação 3.3.1.** O valor de  $\hat{\alpha}_p$  não pode ser um, pois a iteração continuada é realizada se existe ao menos uma componente de bloqueio nos espaço primal e dual, ou seja, temos  $\alpha_p < 1$ , visto que  $\alpha_p$  é calculado da seguinte forma:

$$\alpha_p = \tau h_p = \tau \min \left\{ -\frac{x_{l_1}}{dx_{l_1}}, -\frac{s_{l_2}}{ds_{l_2}} \right\}. \quad (3.50)$$

Logo,

$$\alpha_p = -\tau \frac{x_{l_1}}{dx_{l_1}} \leq -\tau \frac{s_{l_2}}{ds_{l_2}},$$

ou

$$\alpha_p = -\tau \frac{s_{l_2}}{ds_{l_2}} \leq -\tau \frac{x_{l_1}}{dx_{l_1}}.$$

Vamos à análise dos valores para  $\hat{\alpha}_p$  :

- Se  $\hat{\alpha}_p = -\tau \frac{x_{l_1}^k}{dx_{l_1}^k}$  e/ou  $\hat{\alpha}_p = -\tau \frac{s_{l_2}^k}{ds_{l_2}^k}$  então, por (3.50), segue que:

$$\hat{\alpha}_p = \alpha_p.$$

- Se  $\hat{\alpha}_p = -\tau \frac{x_l^k}{2dx_l^k + (\epsilon_x)_l}$  em que  $x_l^k + 2dx_l^k + (\epsilon_x)_l < 0$ , sendo  $l \neq l_1$ , então:

$$\hat{\alpha}_p = -\tau \frac{x_l^k}{2dx_l^k + (\epsilon_x)_l} < \tau \min \left\{ -\frac{x_{l_1}}{dx_{l_1}}, -\frac{s_{l_2}}{ds_{l_2}} \right\} = \alpha_p.$$

Assim,

$$\hat{\alpha}_p < \alpha_p.$$

Por outro lado, considerando  $(\epsilon_x)_l \geq 0$  e  $|dx_l| > (\epsilon_x)_l$ , temos:

$$\hat{\alpha}_p = -\tau \frac{x_l^k}{2dx_l^k + (\epsilon_x)_l} \geq -\tau \frac{x_l^k}{2dx_l^k}.$$

Pela definição de componente de bloqueio, temos que:

$$-\frac{x_{l_1}}{dx_{l_1}} < -\frac{x_l}{dx_l},$$

em que  $x_l + dx_l < 0$ . Logo:

$$2\hat{\alpha}_p \geq -\tau \frac{x_l^k}{dx_l^k} > -\tau \frac{x_{l_1}^k}{dx_{l_1}^k} \geq \alpha_p.$$

- Se  $\hat{\alpha}_p = -\tau \frac{s_l^k}{2ds_l^k + (\epsilon_s)_l}$ , em que  $s_l^k + 2ds_l^k + (\epsilon_s)_l < 0$ , sendo  $l \neq l_2$ , então:

$$\hat{\alpha}_p = -\tau \frac{s_l^k}{2ds_l^k + (\epsilon_s)_l} < \tau \min \left\{ -\frac{x_{l_1}}{dx_{l_1}}, -\frac{s_{l_2}}{ds_{l_2}} \right\} = \alpha_p.$$

Além disso, considerando  $(\epsilon_s)_l \geq 0$  e  $|ds_l| > (\epsilon_s)_l$ , temos:

$$\hat{\alpha}_p = -\tau \frac{s_l^k}{2ds_l^k + (\epsilon_s)_l} \geq -\tau \frac{s_l^k}{2ds_l^k}.$$

Pela definição de componente de bloqueio, segue que:

$$2\hat{\alpha}_p \geq -\tau \frac{s_l^k}{ds_l^k} > -\tau \frac{s_{l_2}^k}{ds_{l_2}^k} \geq \alpha_p.$$

Então, em todos os casos teremos,  $\hat{\alpha}_p \leq \alpha_p < 2\hat{\alpha}_p$ . De modo análogo, para o tamanho do passo dual, obtém-se as conclusões equivalentes,  $\hat{\alpha}_d \leq \alpha_d < 2\hat{\alpha}_d$ .

Portanto, nas componentes de bloqueio mantemos ou reduzimos o valor tamanho de passo de  $\alpha_p$  para  $\hat{\alpha}_p$  e nas demais componentes aumentamos o tamanho de passo dado para  $2\hat{\alpha}_p > \alpha_p$ . O mesmo é realizado nas componentes de bloqueio dual, em que o tamanho do passo é  $\hat{\alpha}_d \leq \alpha_d$ , e nas demais componentes aumentamos o tamanho do passo dado para  $2\hat{\alpha}_d > \alpha_d$ . Caso  $2\hat{\alpha}_p > 1$  e  $2\hat{\alpha}_d > 1$ , para manter igual a 1 o maior tamanho do passo dado em uma direção igual a 1, adotamos  $\hat{\alpha}_p = 1/2$  e  $\hat{\alpha}_d = 1/2$ .

### Múltiplas iterações continuadas

Quando a primeira iteração continuada é realizada, por (3.44), concluímos que a nova direção é dada por:

$$(dx^k, ds^k) = (dx^k + \hat{d}x^k, ds^k + \hat{d}s^k), \quad (3.51)$$

$$(dy^k, dz^k, dw^k) = (dy^k + \hat{d}y^k, dz^k + \hat{d}z^k, dw^k + \hat{d}w^k).$$

Ao realizar a iteração continuada na iteração do método, potencialmente,  $2\hat{\alpha}_p < 1$  e  $2\hat{\alpha}_d < 1$ . Dessa forma, o cálculo de novas direções continuadas pode ser realizado, calculando componentes de bloqueio na nova direção  $d$  (3.51), aumentando  $2\hat{\alpha}_p$  e  $2\hat{\alpha}_d$ .

Na primeira iteração continuada, considere  $L_1^1$  e  $L_2^1$  os conjuntos das componentes de bloqueio em  $x$  e  $s$ , respectivamente, e  $E_x^1$  e  $E_s^1$  as matrizes formadas pelos vetores canônicos correspondentes a  $l_1$  e  $l_2$ , para todo  $l_1 \in L_1^1$  e  $l_2 \in L_2^1$ . Então, obtemos a primeira direção continuada  $\hat{d}x^1, \hat{d}s^1$ :

$$\hat{d}x^1 = dx - E_x^1(E_x^1)^t dx + \epsilon_x^1 = [I - E_x^1(E_x^1)^t]dx + \epsilon_x^1,$$

$$\hat{d}s^1 = ds - E_s^1(E_s^1)^t ds + \epsilon_s^1 = [I - E_s^1(E_s^1)^t]ds + \epsilon_s^1.$$

Adotamos a notação  $dx^*$  e  $ds^*$  como a direção final que é dada após cada iteração continuada.

A nova direção após a primeira iteração continuada é dada por:

$$dx^* = dx + \hat{d}x^1 = [2I - E_x^1(E_x^1)^t]dx + \epsilon_x^1, \quad (3.52)$$

$$ds^* = ds + \hat{d}s^1 = [2I - E_s^1(E_s^1)^t]ds + \epsilon_s^1.$$

Dessa forma, caso somente uma iteração continuada seja realizada, obtemos o ponto  $\hat{x}^{k+1} = x^k + \hat{\alpha}_p dx^*$  e  $\hat{s}^{k+1} = s^k + \hat{\alpha}_p ds^*$  em que:

$$\hat{x}_l^{k+1} = \begin{cases} x_{l_1}^k + \hat{\alpha}_p dx_{l_1}^k, & \text{se } l = l_1 \in L_1^1 \text{ realizou bloqueio em } dx \\ x_l^k + 2\hat{\alpha}_p dx_l^k + \hat{\alpha}_p(\epsilon_x^1)_l, & \text{caso contrário,} \end{cases}$$

$$\hat{s}_l^{k+1} = \begin{cases} s_{l_2}^k + \hat{\alpha}_p ds_{l_2}^k, & \text{se } l = l_2 \in L_2^1 \text{ realizou bloqueio em } ds \\ s_l^k + 2\hat{\alpha}_p ds_l^k + \hat{\alpha}_p(\epsilon_s^1)_l, & \text{caso contrário.} \end{cases}$$

Caso  $2\hat{\alpha}_p < 1$ , realizamos a segunda iteração continuada, determinamos os novos conjuntos das componentes de bloqueio  $L_1^2$  e  $L_2^2$  referentes à direção (3.52). Fixamos em zeros as componentes de bloqueio da segunda direção continuada  $\hat{dx}^2$ , possibilitando que um aumento do tamanho de passo seja dado na nova direção  $dx^* + \hat{dx}^2$ . A segunda direção continuada é dada por:

$$\hat{dx}^2 = dx - E_x^2(E_x^2)^t dx + \epsilon_x^2 = [I - E_x^2(E_x^2)^t] dx + \epsilon_x^2.$$

$$\hat{ds}^2 = ds - E_s^2(E_s^2)^t ds + \epsilon_s^2 = [I - E_s^2(E_s^2)^t] ds + \epsilon_s^2.$$

Logo, a nova direção após a segunda iteração continuada é dada por:

$$dx^* = dx^* + \hat{dx}^2 = [3I - E_x^1(E_x^1)^t - E_x^2(E_x^2)^t] dx + \epsilon_x^1 + \epsilon_x^2$$

$$ds^* = ds^* + \hat{ds}^2 = [3I - E_s^1(E_s^1)^t - E_s^2(E_s^2)^t] ds + \epsilon_s^1 + \epsilon_s^2$$

Assim, temos o ponto  $\hat{x}^{k+1} = x^k + \hat{\alpha}_p dx^*$  e  $\hat{s}^{k+1} = s^k + \hat{\alpha}_p ds^*$ , sendo que:

$$\hat{x}_l^{k+1} = \begin{cases} x_{l_1}^k + \hat{\alpha}_p dx_{l_1}^k, & \text{se } l = l_1 \in L_1^1 \text{ realizou bloqueio em } dx \\ x_{l_1}^k + \hat{\alpha}_p dx_{l_1}^k + \hat{\alpha}_p(\epsilon_x^1)_l, & \text{se } l = l_1 \in L_1^2 \text{ realizou bloqueio em } \hat{dx}^1 \\ x_l^k + 3\hat{\alpha}_p dx_l^k + \hat{\alpha}_p[(\epsilon_x^1)_l + (\epsilon_x^2)_l], & \text{caso contrário,} \end{cases}$$

$$\hat{s}_l^{k+1} = \begin{cases} s_{l_2}^k + \hat{\alpha}_p ds_{l_2}^k, & \text{se } l = l_2 \in L_2^1 \text{ realizou bloqueio em } ds \\ s_{l_2}^k + \hat{\alpha}_p ds_{l_2}^k + \hat{\alpha}_p(\epsilon_s^1)_{l_2} & \text{se } l = l_2 \in L_2^2, \text{ realizou bloqueio em } \hat{ds}^1 \\ s_l^k + 3\hat{\alpha}_p ds_l^k + \hat{\alpha}_p[(\epsilon_s^1)_l + (\epsilon_s^2)_l], & \text{caso contrário.} \end{cases}$$

Assim, sucessivamente, podemos calcular as novas direções continuadas enquanto  $(k_c + 1)\hat{\alpha}_p < 1$ , em que  $k_c$  é o número de iterações continuadas realizadas em uma iteração do método.

Considere  $E_x^l$ ,  $E_s^l$ ,  $E_z^l$  e  $E_w^l$  as matrizes formadas pelos vetores canônicos correspondentes as variáveis  $x_{l_1}$ ,  $s_{l_2}$ ,  $z_{l_3}$  e  $w_{l_4}$  que realizam bloqueio, respectivamente, em cada iteração continuada  $l$  realizada, para  $l = 1, \dots, k_c$ .

As novas direções após o cálculo das iterações continuadas podem ser escritas como:

$$\begin{aligned}
 dx &= dx + \sum_{l=1}^{k_c} \hat{d}x^l = dx + \sum_{l=1}^{k_c} [I - E_x^l (E_x^l)^t] dx + \epsilon_x^l, \\
 ds &= ds + \sum_{l=1}^{k_c} \hat{d}s^l = ds + \sum_{l=1}^{k_c} [I - E_s^l (E_s^l)^t] ds + \epsilon_s^l, \\
 dz &= dz + \sum_{l=1}^{k_c} \hat{d}z^l = dz + \sum_{l=1}^{k_c} [I - E_z^l (E_z^l)^t] dz + \epsilon_z^l, \\
 dw &= dw + \sum_{l=1}^{k_c} \hat{d}w^l = dw + \sum_{l=1}^{k_c} [I - E_w^l (E_w^l)^t] dw + \epsilon_w^l, \\
 dy &= dy + \sum_{l=1}^{k_c} \hat{d}y^l = (k_c + 1) dy.
 \end{aligned} \tag{3.53}$$

E o novo ponto primal  $x$  passa a ser:

$$\begin{aligned}
 \hat{x} &= x + \hat{\alpha}_p dx = \\
 &= x + \hat{\alpha}_p [dx + [I - E_x^1 (E_x^1)^t] dx + [I - E_x^2 (E_x^2)^t] dx + \dots + [I - E_x^{k_c} (E_x^{k_c})^t] dx] + \bar{\epsilon}_x = \\
 &= x + \hat{\alpha}_p [(k_c + 1)I - E_x^1 (E_x^1)^t + E_x^2 (E_x^2)^t + \dots + E_x^{k_c} (E_x^{k_c})^t] dx + \bar{\epsilon}_x,
 \end{aligned}$$

em que  $\bar{\epsilon}_x = \hat{\alpha}_p \hat{\epsilon}_x = \hat{\alpha}_p \sum_{l=1}^{k_c} \epsilon_x^l$ .

De forma geral, o novo ponto é dado por:

$$\hat{x} = x + \hat{\alpha}_p[(k_c + 1)I - \sum_{l=1}^{k_c} E_x^l (E_x^l)^t] dx + \bar{\epsilon}_x,$$

$$\hat{s} = s + \hat{\alpha}_p[(k_c + 1)I - \sum_{l=1}^{k_c} E_s^l (E_s^l)^t] ds + \bar{\epsilon}_s,$$

$$\hat{z} = z + \hat{\alpha}_d[(k_c + 1)I - \sum_{l=1}^{k_c} E_z^l (E_z^l)^t] dz + \bar{\epsilon}_z,$$

$$\hat{w} = w + \hat{\alpha}_d[(k_c + 1)I - \sum_{l=1}^{k_c} E_w^l (E_w^l)^t] dw + \bar{\epsilon}_w,$$

$$\hat{y} = y + (k_c + 1)\hat{\alpha}_d dy,$$

em que  $\bar{\epsilon}_s = \sum_{l=1}^{k_c} \hat{\alpha}_p \epsilon_s^l$ ,  $\bar{\epsilon}_z = \sum_{l=1}^{k_c} \hat{\alpha}_d \epsilon_z^l$  e  $\bar{\epsilon}_w = \sum_{l=1}^{k_c} \hat{\alpha}_d \epsilon_w^l$ .

### Análise dos Novos Resíduos

Nesta seção, analisamos os novos resíduos, verificando se é necessária alguma restrição para aplicação da iteração continuada ao método para garantir menores resíduos em cada iteração.

Em cada direção continuada  $l$  calculada, para  $l = 1, \dots, k_c$ , temos

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{bmatrix} \begin{bmatrix} \hat{d}x^l \\ \hat{d}s^l \\ \hat{d}y^l \\ \hat{d}z^l \\ \hat{d}w^l \end{bmatrix} = \begin{bmatrix} r_p \\ r_u \\ r_d \\ r_s \\ r_t \end{bmatrix} + \begin{bmatrix} r_1^l \\ r_2^l \\ r_3^l \\ r_4^l \\ r_5^l \end{bmatrix},$$

Adotamos a notação “ $\hat{\phantom{x}}$ ” para os novos pontos e resíduos obtidos com a utilização da iteração continuada. Assim,  $\hat{r}_p^l$  representa o resíduo primal utilizando a direção dada em (3.53) com o cálculo de  $\hat{d}^l$  e  $r_p^{k+1}$  é o resíduo primal sem utilizar a direção continuada. Temos, então:

$$\begin{aligned}
\hat{r}_p^l &= b - A\hat{x}^{k+1} = \\
&= b - A \left[ x^k + \hat{\alpha}_p \left( dx^k + \sum_{q=1}^l \hat{d}x^q \right) \right] = \\
&= r_p^k - \hat{\alpha}_p A dx^k - \hat{\alpha}_p \sum_{q=1}^l A \hat{d}x^q = \\
&= (1 - \hat{\alpha}_p) r_p^k - \hat{\alpha}_p \sum_{q=1}^l (r_p^k + r_1^q) = \\
&= [1 - (l+1)\hat{\alpha}_p] r_p^k - \hat{\alpha}_p \sum_{q=1}^l r_1^q.
\end{aligned}$$

De forma similar, para os demais resíduos, temos:

$$\hat{r}_u^l = [1 - (l+1)\hat{\alpha}_p] r_u^k - \hat{\alpha}_p \sum_{q=1}^l r_2^q,$$

$$\hat{r}_d^l = [1 - (l+1)\hat{\alpha}_p] r_d^k - \hat{\alpha}_p \sum_{q=1}^l r_3^q.$$

Observe que, quando  $l = 1$ , temos:

$$\hat{r}_p^1 = (1 - 2\hat{\alpha}_p) r_p^k - \hat{\alpha}_p r_1^1.$$

Logo, se  $r_1^1 = 0$ , como ocorre pela direção moderada (3.42), teríamos

$$\|\hat{r}_p^1\| \leq (1 - 2\hat{\alpha}_p) \|r_p^k\| < (1 - \alpha_p) \|r_p^k\| = \|r_p^{k+1}\|.$$

Assim, o novo resíduo  $\|\hat{r}_p^1\|$  obtido na primeira iteração continuada seria menor do que resíduo  $\|r_p^{k+1}\|$ , obtido se esta abordagem não estivesse sendo utilizada.

Na segunda iteração continuada, ou seja, para  $l = 2$ , obteríamos:

$$\hat{r}_p^2 = (1 - 3\hat{\alpha}_p) r_p^k - \hat{\alpha}_p (r_1^1 + r_1^2).$$

Logo, visto que  $r_1^1 = 0$ , caso  $r_1^2 = 0$ , teríamos

$$\|\hat{r}_p^2\| \leq (1 - 3\hat{\alpha}_p) \|r_p^k\| < (1 - 2\hat{\alpha}_p) \|r_p^k\| = \|\hat{r}_p^1\|.$$

Assim, o resíduo após a segunda iteração continuada seria menor que o anterior. O mesmo pode ser verificado para os demais resíduos. No entanto, podemos ter:

$$\left\| \sum_{q=1}^l r_1^q \right\| \neq 0,$$

$$\left\| \sum_{q=1}^l r_2^q \right\| \neq 0,$$

$$\left\| \sum_{q=1}^l r_3^q \right\| \neq 0.$$

Desse modo, não temos garantida a redução dos resíduos. Então, considerando  $\hat{r}_p^0 = (1 - \alpha_p)r_p^k$ ,  $\hat{r}_u^0 = (1 - \alpha_u)r_u^k$  e  $\hat{r}_d^0 = (1 - \alpha_d)r_d^k$ . Na iteração continuada  $l$ , a nova direção é utilizada se:

$$\|\hat{r}_p^l\| < \|\hat{r}_p^{l-1}\|,$$

$$\|\hat{r}_u^l\| < \|\hat{r}_u^{l-1}\|,$$

$$\|\hat{r}_d^l\| < \|\hat{r}_d^{l-1}\|,$$

para  $l = 1, 2, \dots, k_c$ .

A forma avançada de utilizar a iteração continuada é dada resumidamente no Algoritmo 3.

---

**Algoritmo 3** Iteração continuada pela forma avançada

---

**Parâmetros** $dx, ds, dy, dz, dw;$  $\alpha_p, \alpha_d;$  $r_p, r_u, r_d;$ **Inicialização** $l = 1,$  $(\bar{dx}, \bar{ds}, \bar{dy}, \bar{dz}, \bar{dw}) = (dx, ds, dy, dz, dw);$  $\hat{r}_p^0 = (1 - \alpha_p)r_p, \hat{r}_u^0 = (1 - \alpha_p)r_u, \hat{r}_d^0 = (1 - \alpha_d)r_d;$  $\bar{\alpha}_p = \alpha_p, \bar{\alpha}_d = \alpha_d;$ **Processo Iterativo**

1. Enquanto  $\bar{\alpha}_p < 1$  e  $\bar{\alpha}_d < 1$  faça
    2. Encontre as componentes de bloqueio na direção  $d$ ;
    3. Calcule a direção continuada  $\hat{d}$ ;
    4. Faça
 
$$(\bar{dx}, \bar{ds}) = (\bar{dx}, \bar{ds}) + (\hat{dx}, \hat{ds});$$

$$(\bar{dy}, \bar{dz}, \bar{dw}) = (\bar{dy}, \bar{dz}, \bar{dw}) + (\hat{dy}, \hat{dz}, \hat{dw});$$
    5. Encontre os tamanhos dos passos primal e dual  $(\hat{\alpha}_p, \hat{\alpha}_d)$  na direção  $\bar{d}$ ;
    6. Calcule os resíduos
 
$$\hat{r}_p^l, \hat{r}_d^l, \hat{r}_u^l;$$
    7. Se  $\|\hat{r}_p^l\| < \|\hat{r}_p^{l-1}\|, \|\hat{r}_u^l\| < \|\hat{r}_u^{l-1}\|, \|\hat{r}_d^l\| < \|\hat{r}_d^{l-1}\|$  então
 

Atualize a direção

$$(dx, ds) = (\bar{dx}, \bar{ds});$$

$$(dy, dz, dw) = (\bar{dy}, \bar{dz}, \bar{dw});$$

Calcule

$$\bar{\alpha}_p = (l + 1)\hat{\alpha}_p$$

$$\bar{\alpha}_d = (l + 1)\hat{\alpha}_d$$
    8. Caso contrário
 

A iteração continuada é finalizada;
    9. Fim se
    9. Atualize  $l = l + 1$ ;
  10. Fim enquanto
-

### 3.3.2 Forma atrasada

Na forma atrasada de utilizar a direção continuada, a nova direção  $d$  é dada por  $d = d + \delta \hat{d}$ , em que  $\delta \in (0, 1)$  e depende de  $\alpha_p$  e  $\alpha_d$ , tamanhos dos passos máximos que podem ser dados na direção preditora corretora.

Considere:

$$\hat{\alpha}_p = \min \left\{ 1, \tau \max \left\{ \alpha : x^k + \alpha \hat{d}x^k > 0 \text{ e } s^k + \alpha \hat{d}s^k > 0 \right\} \right\}, \quad (3.54)$$

$$\hat{\alpha}_d = \min \left\{ 1, \tau \max \left\{ \alpha : z^k + \alpha \hat{d}z^k > 0 \text{ e } w^k + \alpha \hat{d}w^k > 0 \right\} \right\},$$

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

O novo ponto ao usar a forma atrasada da direção continuada é

$$(\hat{x}^{k+1}, \hat{s}^{k+1}) = (x^k, s^k) + \alpha_p \left[ (dx^k, ds^k) + \frac{\delta_p}{\alpha_p} (\hat{d}x^k, \hat{d}s^k) \right], \quad (3.55)$$

$$(\hat{y}^{k+1}, \hat{z}^{k+1}, \hat{w}^{k+1}) = (y^k, z^k, w^k) + \alpha_d \left[ (dx^k, ds^k) + \frac{\delta_d}{\alpha_d} (\hat{d}y^k, \hat{d}z^k, \hat{d}w^k) \right],$$

em que  $\delta_p = \hat{\alpha}_p - \alpha_p$  e  $\delta_d = \hat{\alpha}_d - \alpha_d$ .

Dessa forma, a nova direção  $d$  é dada por:

$$(dx^k, ds^k) = (dx^k, ds^k) + \frac{\delta_p}{\alpha_p} (\hat{d}x^k, \hat{d}s^k), \quad (3.56)$$

$$(dy^k, dz^k, dw^k) = (dy^k, dz^k, dw^k) + \frac{\delta_d}{\alpha_d} (\hat{d}y^k, \hat{d}z^k, \hat{d}w^k).$$

Então, para a direção  $\hat{d}x = dx - E_x E_x^t + \epsilon_x$  obtemos o ponto primal:

$$\begin{aligned} \hat{x}^{k+1} &= x^k + \alpha_p \left[ dx^k + \frac{\delta_p}{\alpha_p} (dx^k - E_x E_x^t dx^k + \epsilon_x) \right] = \\ &= x^k + (\alpha_p + \delta_p) dx^k - \delta_p (E_x E_x^t dx^k + \epsilon_x) = \\ &= x^k + (\alpha_p + \delta_p) \left( I - \frac{\delta_p}{\alpha_p + \delta_p} E_x E_x^t \right) dx^k + \delta_p \epsilon_x = \\ &= x^k + (\alpha_p + \delta_p) \left[ I - \left( 1 - \frac{\alpha_p}{\alpha_p + \delta_p} \right) E_x E_x^t \right] dx^k + \delta_p \epsilon_x. \end{aligned} \quad (3.57)$$

De modo similar, para os demais pontos  $s$ ,  $z$  e  $w$ , temos:

$$\begin{aligned}\hat{s}^{k+1} &= s^k + (\alpha_p + \delta_p) \left[ I - \left( 1 - \frac{\alpha_p}{\alpha_p + \delta_p} \right) E_s E_s^t \right] ds^k + \delta_p \epsilon_s, \\ \hat{z}^{k+1} &= z^k + (\alpha_d + \delta_d) \left[ I - \left( 1 - \frac{\alpha_d}{\alpha_d + \delta_d} \right) E_z E_z^t \right] dz^k + \delta_d \epsilon_z, \\ \hat{w}^{k+1} &= w^k + (\alpha_d + \delta_d) \left[ I - \left( 1 - \frac{\alpha_d}{\alpha_d + \delta_d} \right) E_w E_w^t \right] dw^k + \delta_d \epsilon_w.\end{aligned}\tag{3.58}$$

Comparando (3.57) e (3.58) com (3.2) e (3.3), temos

$$\begin{aligned}\theta_p^1 &= \alpha_p, \\ \theta_p^2 &= \alpha_p + \delta_p\end{aligned}$$

e

$$\begin{aligned}\theta_d^1 &= \alpha_d, \\ \theta_d^2 &= \alpha_d + \delta_d,\end{aligned}$$

em que  $\delta_p = \hat{\alpha}_p - \alpha_p$  e  $\delta_d = \hat{\alpha}_d - \alpha_d$ . Então, devemos verificar se  $\delta_p > 0$  e  $\delta_d > 0$ , ou seja,  $\hat{\alpha}_p > \alpha_p$  e  $\hat{\alpha}_d > \alpha_d$ .

**Análise do tamanho de passo  $\hat{\alpha}_p$**

Pela forma atrasada (3.55) obtemos os pontos primais dados em (3.57) e (3.58), que podem ser escritos na seguinte forma:

$$\begin{aligned}\hat{x}_l^{k+1} &= \begin{cases} x_{l_1}^k + \alpha_p dx_{l_1}^k, & \text{se } l = l_1 \in L_1 \text{ realizou bloqueio em } \hat{dx}; \\ x_l^k + \hat{\alpha}_p [dx_l^k + (\epsilon_x)_l], & \text{caso contrário.} \end{cases} \\ \hat{s}_l^{k+1} &= \begin{cases} s_{l_2}^k + \alpha_p ds_{l_2}^k, & \text{se } t = l_2 \in L_2 \text{ realizou bloqueio em } \hat{ds}; \\ s_l^k + \hat{\alpha}_p [ds_l^k + (\epsilon_s)_l], & \text{caso contrário.} \end{cases}\end{aligned}$$

Para garantir que este ponto  $(\hat{x}^{k+1}, \hat{s}^{k+1})$  seja interior, o tamanho do passo  $\hat{\alpha}_p$  é dado por:

$$\hat{\alpha}_p = \min \{1, \tau \rho_x, \tau \rho_s\},$$

em que

$$\begin{aligned}\rho_x &= \min_{\substack{1 \leq l \leq n, \\ l \neq l_1}} \left\{ -\frac{x_l^k}{dx_l^k + (\epsilon_x)_l} : x_l^k + dx_l^k + (\epsilon_x)_l \leq 0 \right\}, \\ \rho_s &= \min_{\substack{1 \leq l \leq n, \\ l \neq l_2}} \left\{ -\frac{s_l^k}{ds_l^k + (\epsilon_s)_l} : s_l^k + ds_l^k + (\epsilon_s)_l \leq 0 \right\}.\end{aligned}\tag{3.59}$$

Casos a considerar:

- $\hat{\alpha}_p = 1$ .

Nesse caso,  $\hat{\alpha}_p > \alpha_p$  e  $\delta_p = \hat{\alpha}_p - \alpha_p = 1 - \alpha_p$ .

- $\hat{\alpha}_p = \tau \rho_x = -\tau \frac{x_l^k}{dx_l^k + (\epsilon_x)_l}$ , em que  $x_l^k + dx_l^k + (\epsilon_x)_l < 0$ .

Considerando  $(\epsilon_x)_l \geq 0$  e  $|dx_l^k| > (\epsilon_x)_l$ , segue que:

$$-\frac{x_l^k}{dx_l^k + (\epsilon_x)_l} \geq -\frac{x_l^k}{dx_l^k}.$$

Pela definição de componente de bloqueio, temos

$$-\frac{x_{l_1}}{dx_{l_1}} < -\frac{x_l}{dx_l},$$

em que  $x_l + dx_l < 0$ .

Dessa forma:

$$\hat{\alpha}_p = \tau \rho_x = -\tau \frac{x_l^k}{dx_l^k + (\epsilon_x)_l} \geq -\frac{x_l^k}{dx_l^k} > -\tau \frac{x_{l_1}^k}{dx_{l_1}^k} \geq \alpha_p.$$

Logo,  $\delta_p = \hat{\alpha}_p - \alpha_p > 0$ .

- $\hat{\alpha}_p = \tau \rho_s = -\tau \frac{s_l^k}{ds_l^k + (\epsilon_s)_l}$ , em que  $s_l^k + ds_l^k + (\epsilon_s)_l < 0$ .

Considerando que  $(\epsilon_s)_l \geq 0$  e  $|ds_l^k| > (\epsilon_s)_l$ , então:

$$-\frac{s_l^k}{ds_l^k + (\epsilon_s)_l} \geq -\frac{s_l^k}{ds_l^k} > -\frac{s_{l_2}^k}{ds_{l_2}^k}.$$

Assim,

$$\hat{\alpha}_p = \tau \rho_s = -\tau \frac{s_l^k}{ds_l^k + (\epsilon_s)_l} > -\tau \frac{s_{l_2}^k}{ds_{l_2}^k} \geq \alpha_p.$$

Logo,  $\delta_p = \hat{\alpha}_p - \alpha_p > 0$ .

A mesma análise é feita para o tamanho do passo dual final, obtendo-se, similarmente,  $\hat{\alpha}_d > \alpha_d$  e  $\delta_d > 0$ .

Portanto, nas componentes de bloqueio primais mantemos o tamanho do passo em  $\alpha_p$  e nas demais componentes aumentamos o tamanho do passo dado para  $\alpha_p + \delta_p < 1$ . O mesmo é realizado nas componentes de bloqueio duais, em que mantemos o tamanho do passo em  $\alpha_d$ , enquanto nas demais componentes aumentamos o tamanho do passo dado para  $\alpha_d + \delta_d < 1$ .

### Múltiplas iterações continuadas

Ao realizar uma iteração continuada no método preditor corretor, os tamanhos dos passos  $\alpha_p + \delta_p$  e  $\alpha_d + \delta_d$  podem ser menores que o passo de Newton. Nesse caso, novas direções continuadas podem ser realizadas, calculando componentes de bloqueio na nova direção já obtida

$$(dx^k, ds^k) = (dx^k, ds^k) + \frac{\delta_p}{\alpha_p}(\hat{dx}^k, \hat{ds}^k),$$

$$(dy^k, dz^k, dw^k) = (dy^k, dz^k, dw^k) + \frac{\delta_d}{\alpha_d}(\hat{dy}^k, \hat{dz}^k, \hat{dw}^k),$$

umentando  $\delta_p$  e  $\delta_d$  desde que  $\alpha_p + \delta_p \leq 1$  e  $\alpha_d + \delta_d \leq 1$ .

Na primeira iteração continuada, considere  $L_1^1$  e  $L_2^1$  os conjuntos com as componentes de bloqueio em  $x$  e  $s$ , respectivamente, e  $E_x^1$  e  $E_s^1$  as matrizes formadas pelos vetores canônicos correspondentes a  $l_1$  e  $l_2$  para todo  $l_1 \in L_1^1$  e  $l_2 \in L_2^1$ . Então, obtemos a primeira direção continuada:

$$(\hat{dx}^1, \hat{ds}^1) = (dx, ds) - (E_x^1(E_x^1)^t dx, E_s^1(E_s^1)^t ds) + (\epsilon_x^1, \epsilon_s^1).$$

Se  $\hat{\alpha}_p$  é o maior tamanho de passo que pode ser dado em  $(x, s)$  na direção  $(\hat{dx}^1, \hat{ds}^1)$  então  $\delta_p^1 = \hat{\alpha}_p - \alpha_p$ .

Logo, a nova direção ao final da primeira iteração continuada, que denotaremos por  $(dx^*, ds^*)$ , é dada por:

$$(dx^*, ds^*) = (dx, ds) + \frac{\delta_p^1}{\alpha_p}(\hat{dx}^1, \hat{ds}^1). \quad (3.60)$$

Se somente uma iteração continuada é realizado, o ponto  $\hat{x}^{k+1} = x^k + \alpha_p dx^*$  e  $\hat{s}^{k+1} = s^k + \alpha_p ds^*$  é dado na seguinte forma:

$$\hat{x}_l^{k+1} = \begin{cases} x_{l_1}^k + \alpha_p dx_{l_1}^k, & \text{se } l = l_1 \in L_1^1 \text{ realizou bloqueio em } dx, \\ x_l^k + (\alpha_p + \delta_p^1)[dx_l^k + (\epsilon_x)_l^1], & \text{caso contrário.} \end{cases}$$

$$\hat{s}_l^{k+1} = \begin{cases} s_{l_2}^k + \alpha_p ds_{l_2}^k, & \text{se } l = l_2 \in L_2^1, \text{ realizou bloqueio em } ds \\ s_l^k + (\alpha_p + \delta_p^1)[ds_l^k + (\epsilon_s)_l^1], & \text{caso contrário.} \end{cases}$$

Caso  $(\alpha_p + \delta_p^1) < 1$ , a segunda iteração continuada é realizada, calculando-se as componentes de bloqueio na direção (3.60) e obtendo  $L_1^2$  e  $L_2^2$ . A segunda direção continuada é dada por:

$$(\hat{dx}^2, \hat{ds}^2) = (\hat{dx}^1, \hat{ds}^1) - (E_x^2(E_x^2)^t dx, E_s^2(E_s^2)^t ds) + (\epsilon_x^2, \epsilon_s^2).$$

Se  $\hat{\alpha}_p$  é o maior tamanho de passo que pode ser dado em  $(x, s)$  na direção  $(\hat{dx}^2, \hat{ds}^2)$ , então  $\delta_p^2 = \hat{\alpha}_p - \alpha_p - \delta_p^1$ .

Logo, a nova direção ao final da segunda iteração continuada tem a seguinte forma:

$$(dx^*, ds^*) = (dx, ds) + \frac{\delta_p^1}{\alpha_p}(\hat{dx}^1, \hat{ds}^1) + \frac{\delta_p^2}{\alpha_p}(\hat{dx}^2, \hat{ds}^2).$$

Assim, temos o ponto  $\hat{x}^{k+1} = x^k + \alpha_p dx^*$  e  $\hat{s}^{k+1} = s^k + \alpha_p ds^*$  em que:

$$\hat{x}_l^{k+1} = \begin{cases} x_{l_1}^k + \alpha_p dx_{l_1}^k, & \text{se } l = l_1 \in L_1^1 \text{ realizou bloqueio em } dx \\ x_{l_1}^k + (\alpha_p + \delta_p^1)dx_{l_1}^k + \delta_p^1(\epsilon_x^1)_{l_1}, & \text{se } l = l_1 \in L_1^2 \text{ realizou bloqueio em } \hat{dx}^1 \\ x_l^k + (\alpha_p + \delta_p^1 + \delta_p^2)dx_l^k + \delta_p^1(\epsilon_x^1)_l + \delta_p^2(\epsilon_x^2)_l, & \text{caso contrário,} \end{cases}$$

$$\hat{s}_l^{k+1} = \begin{cases} s_{l_2}^k + \alpha_p ds_{l_2}^k, & \text{se } l = l_2 \in L_2^1 \text{ realizou bloqueio em } ds \\ s_{l_2}^k + (\alpha_p + \delta_p^1)ds_{l_2}^k + \delta_p^1(\epsilon_s^1)_{l_2}, & \text{se } l = l_2 \in L_2^2 \text{ realizou bloqueio em } \hat{ds}^1 \\ s_l^k + (\alpha_p + \delta_p^1 + \delta_p^2)ds_l^k + \delta_p^1(\epsilon_s^1)_l + \delta_p^2(\epsilon_s^2)_l, & \text{caso contrário.} \end{cases}$$

Logo, estamos aumentando ainda mais o tamanho do passo nas componentes que não realizaram bloqueio.

Assim, sucessivamente, podemos calcular novas direções continuadas enquanto  $\alpha_p + \delta_p^1 + \delta_p^2 + \delta_p^3 + \dots < 1$ .

Considere  $E_x^l, E_s^l, E_z^l$  e  $E_w^l$  as matrizes em que cada linha é formada pelos vetores canônicos correspondentes às variáveis  $x_{l_1}, s_{l_2}, z_{l_3}$  e  $w_{l_4}$  que realizam bloqueio, respectivamente, em cada iteração continuada  $l$  realizada, para  $l = 1, \dots, k_c$ , sendo  $k_c$  o número de iterações continuadas feitas em uma iteração do método. Realizando múltiplas

iterações continuadas, a nova direção pode ser escrita como:

$$\begin{aligned}
 dx &= dx + \frac{1}{\alpha_p} \sum_{l=1}^{k_c} \delta_p^l \hat{d}x^l, \\
 ds &= ds + \frac{1}{\alpha_p} \sum_{l=1}^{k_c} \delta_p^l \hat{d}s^l, \\
 dz &= dz + \frac{1}{\alpha_d} \sum_{l=1}^{k_c} \delta_d^l \hat{d}z^l, \\
 dw &= dw + \frac{1}{\alpha_d} \sum_{l=1}^{k_c} \delta_d^l \hat{d}w^l, \\
 dy &= dy + \sum_{l=1}^{k_c} \delta_d^l \hat{d}y^l,
 \end{aligned} \tag{3.61}$$

em que

$$\hat{d}x^l = \hat{d}x^{l-1} - E_x^l (E_x^l)^t + \epsilon_x^l$$

$$\hat{d}s^l = \hat{d}s^{l-1} - E_s^l (E_s^l)^t + \epsilon_s^l$$

$$\hat{d}z^l = \hat{d}z^{l-1} - E_z^l (E_z^l)^t + \epsilon_z^l$$

$$\hat{d}w^l = \hat{d}w^{l-1} - E_w^l (E_w^l)^t + \epsilon_w^l$$

$$\hat{d}y^l = dy.$$

O novo ponto primal  $x$  passa a ser:

$$\begin{aligned}
 \hat{x} &= x + \alpha_p dx = \\
 &= x + \alpha_p \left[ dx + \frac{\delta_p^1}{\alpha_p} \hat{d}x^1 + \frac{\delta_p^2}{\alpha_p} \hat{d}x^2 + \dots + \frac{\delta_p^{k_c}}{\alpha_p} \hat{d}x^{k_c} \right] = \\
 &= x + \alpha_p dx + \delta_p^1 [I - E_x^1 (E_x^1)^t] dx + \delta_p^2 [I - E_x^1 (E_x^1)^t - E_x^2 (E_x^2)^t] dx + \\
 &+ \dots + \delta_p^{k_c} [I - E_x^1 (E_x^1)^t - E_x^2 (E_x^2)^t - \dots - E_x^{k_c} (E_x^{k_c})^t] dx + \bar{\epsilon}_x,
 \end{aligned}$$

em que  $\bar{\epsilon}_x = \sum_{l=1}^{k_c} \delta_p^l \epsilon_x^l$ .

De forma geral, utilizando a direção continuada pela forma atrasada, o ponto é dado por:

$$\begin{aligned}
\hat{x} &= x + \alpha_p dx + \sum_{l=1}^{k_c} \delta_p^l [I - \sum_{q=1}^l E_x^q (E_x^q)^t] dx + \bar{\epsilon}_x, \\
\hat{s} &= s + \alpha_p ds + \sum_{l=1}^{k_c} \delta_p^l [I - \sum_{q=1}^l E_s^q (E_s^q)^t] ds + \bar{\epsilon}_s, \\
\hat{z} &= z + \alpha_d dz + \sum_{l=1}^{k_c} \delta_d^l [I - \sum_{q=1}^l E_z^q (E_z^q)^t] dz + \bar{\epsilon}_z, \\
\hat{w} &= w + \alpha_d dw + \sum_{l=1}^{k_c} \delta_d^l [I - \sum_{q=1}^l E_w^q (E_w^q)^t] dw + \bar{\epsilon}_w, \\
\hat{y} &= y + (\alpha_d + \sum_{l=1}^{k_c} \delta_d^l) dy,
\end{aligned} \tag{3.62}$$

em que  $\bar{\epsilon}_s = \sum_{l=1}^{k_c} \delta_p^l \epsilon_s^l$ ,  $\bar{\epsilon}_z = \sum_{l=1}^{k_c} \delta_d^l \epsilon_z^l$  e  $\bar{\epsilon}_w = \sum_{l=1}^{k_c} \delta_d^l \epsilon_w^l$ .

### Análise dos Novos Resíduos

Analisando os novos resíduos, verificamos possíveis restrições para a aplicação da iteração continuada ao método preditor corretor.

Em cada direção continuada  $l = 1, \dots, k_c$  calculada, temos

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & A^t & I & -I \\ Z & 0 & 0 & X & 0 \\ 0 & W & 0 & 0 & S \end{bmatrix} \begin{bmatrix} \hat{dx}^l \\ \hat{ds}^l \\ \hat{dy}^l \\ \hat{dz}^l \\ \hat{dw}^l \end{bmatrix} = \begin{bmatrix} r_p \\ r_u \\ r_d \\ r_s \\ r_t \end{bmatrix} + \begin{bmatrix} r_1^l \\ r_2^l \\ r_3^l \\ r_4^l \\ r_5^l \end{bmatrix}.$$

Dessa forma:

$$\begin{aligned}
\hat{r}_p^l &= b - A\hat{x}^{k+1} = \\
&= b - A \left[ x^k + \alpha_p \left( dx^k + \sum_{q=1}^l \frac{\delta_p^q}{\alpha_p} \hat{dx}^q \right) \right] = \\
&= r_p^k - \alpha_p A dx^k - \sum_{q=1}^l \delta_p^q A \hat{dx}^q = \\
&= (1 - \alpha_p) r_p^k - \sum_{q=1}^l \delta_p^q (r_p^k + r_1^q) = \\
&= \left[ 1 - \left( \alpha_p + \sum_{q=1}^l \delta_p^q \right) \right] r_p^k - \sum_{q=1}^l \delta_p^q r_1^q.
\end{aligned}$$

Definindo  $\hat{r}_p^0 = r_p^{k+1} = (1 - \alpha_p)r_p^k$ , temos

$$\hat{r}_p^l = \hat{r}_p^{l-1} - \delta_p^l (r_p^k + r_1^l),$$

para  $l = 1, \dots, k_c$ .

De forma similar, obtemos os demais resíduos. Para  $\hat{r}_u^0 = r_u^{k+1} = (1 - \alpha_p)r_u^k$  e  $\hat{r}_d^0 = r_d^{k+1} = (1 - \alpha_d)r_d^k$ , temos:

$$\hat{r}_u^l = \hat{r}_u^{l-1} - \delta_p^l (r_u^k + r_2^l),$$

$$\hat{r}_d^l = \hat{r}_d^{l-1} - \delta_d^l (r_d^k + r_3^l).$$

para  $l = 1, \dots, k_c$ .

Note que, quando  $l = 1$ , para o resíduo  $\hat{r}_p^1$ , temos:

$$\begin{aligned}
\hat{r}_p^1 &= \hat{r}_p^0 - \delta_p^1 (r_p^k + r_1^1) = \\
&= (1 - \alpha_p) r_p^k - \delta_p^1 r_p^k - \delta_p^1 r_1^1 = \\
&= [1 - (\alpha_p + \delta_p^1)] r_p^k - \delta_p^1 r_1^1.
\end{aligned}$$

Assim, caso  $\|r_1^1\| = 0$ , teríamos:

$$\begin{aligned}
\|\hat{r}_p^1\| &\leq [1 - (\alpha_p + \delta_p^1)] \|r_p^k\| + \delta_p^1 \|r_1^1\| \leq \\
&\leq (1 - \alpha_p) \|r_p^k\|.
\end{aligned}$$

No caso  $l = 2$ , temos

$$\begin{aligned}\hat{r}_p^2 &= \hat{r}_p^0 - \delta_p^1(r_p^k + r_1^1) - \delta_p^2(r_p^k + r_1^2) = \\ &= (1 - \alpha_p)r_p^k - (\delta_p^1 + \delta_p^2)r_p^k - \delta_p^1r_1^1 - \delta_p^2r_1^2 = \\ &= [1 - (\alpha_p + \delta_p^1 + \delta_p^2)]r_p^k - \delta_p^1r_1^1 - \delta_p^2r_1^2.\end{aligned}$$

Logo, se  $\|r_1^1\| = \|r_1^2\| = 0$ , obtemos:

$$\begin{aligned}\|\hat{r}_p^2\| &\leq [1 - (\alpha_p + \delta_p^1 + \delta_p^2)]\|r_p^k\| + \delta_p^1\|r_1^1\| + \delta_p^2\|r_1^2\| \leq \\ &\leq (1 - \alpha_p + \delta_p^1)\|r_p^k\| = \|\hat{r}_p^1\|,\end{aligned}$$

ou seja, estaríamos reduzindo ainda mais o resíduo primal  $\hat{r}_p^l$  na segunda iteração continuada, comparado com a primeira continuada realizada.

O mesmo pode ser verificado para os demais resíduos  $\hat{r}_u^l$  e  $\hat{r}_d^l$ , em relação a  $r_2^l$  e  $r_3^l$ . No entanto, não é garantido que  $\|r_1^l\| = \|r_2^l\| = \|r_3^l\| = 0$  para todo  $l = 1, 2, \dots, k_c$ . Logo, na iteração continuada  $l$ , a nova direção é utilizada se:

$$\|\hat{r}_p^l\| = \|\hat{r}_p^{l-1} - \delta_p^l(r_p^k + r_1^l)\| < \|\hat{r}_p^{l-1}\|,$$

$$\|\hat{r}_u^l\| = \|\hat{r}_u^{l-1} - \delta_p^l(r_u^k + r_2^l)\| < \|\hat{r}_u^{l-1}\|,$$

$$\|\hat{r}_d^l\| = \|\hat{r}_d^{l-1} - \delta_d^l(r_d^k + r_3^l)\| < \|\hat{r}_d^{l-1}\|.$$

A forma de utilizar a direção continuada atrasada é dada resumidamente no Algoritmo 4.

---

**Algoritmo 4** Iteração continuada pela forma atrasada

---

**Parâmetros** $dx, ds, dy, dz, dw;$  $\alpha_p, \alpha_d;$  $r_p, r_u, r_d;$ **Inicialização** $l = 1,$  $(\bar{dx}, \bar{ds}, \bar{dy}, \bar{dz}, \bar{dw}) = (dx, ds, dy, dz, dw);$  $\hat{r}_p^0 = (1 - \alpha_p)r_p, \hat{r}_u^0 = (1 - \alpha_p)r_u, \hat{r}_d^0 = (1 - \alpha_d)r_d;$  $\bar{\delta}_p = 0, \bar{\delta}_d = 0;$ **Processo Iterativo**

1. Enquanto  $\alpha_p + \bar{\delta}_p < 1$  e  $\alpha_d + \bar{\delta}_d < 1$  faça
    2. Encontre as componentes de bloqueio na direção  $d$ ;
    3. Calcule a direção continuada  $\hat{d}$ ;
    4. Encontre os tamanhos dos passos primal e dual na direção  $\hat{d}$ ;
    5. Calcule
 
$$\delta_p = \hat{\alpha}_p - \alpha_p - \bar{\delta}_p, \delta_d = \hat{\alpha}_d - \alpha_d - \bar{\delta}_d;$$

$$\bar{\delta}_p = \bar{\delta}_p + \delta_p, \bar{\delta}_d = \bar{\delta}_d + \delta_d;$$
    6. Faça
 
$$(\bar{dx}, \bar{ds}) = (\bar{dx}, \bar{ds}) + \frac{\delta_p}{\alpha_p}(\hat{dx}, \hat{ds});$$

$$(\bar{dy}, \bar{dz}, \bar{dw}) = (\bar{dy}, \bar{dz}, \bar{dw}) + \frac{\delta_d}{\alpha_d}(\hat{dy}, \hat{dz}, \hat{dw});$$
    7. Calcule os resíduos
 
$$\hat{r}_p^l, \hat{r}_d^l, \hat{r}_u^l;$$
    8. Se  $\|\hat{r}_p^l\| < \|\hat{r}_p^{l-1}\|, \|\hat{r}_u^l\| < \|\hat{r}_u^{l-1}\|, \|\hat{r}_d^l\| < \|\hat{r}_d^{l-1}\|$  então
 

Atualize a direção

$$(dx, ds) = (\bar{dx}, \bar{ds});$$

$$(dy, dz, dw) = (\bar{dy}, \bar{dz}, \bar{dw});$$
    9. Caso contrário
 

A iteração continuada é finalizada.
    10. Fim se
    11. Atualize  $l = l + 1$ ;
  12. Fim enquanto
-

### 3.4 Critério para utilização da direção continuada

Considere:

$$\hat{\lambda} = \frac{\hat{x}^t \hat{z} + \hat{s}^t \hat{w}}{1 + |c^t \hat{x}|} \quad \text{e} \quad \lambda = \frac{x^t z + s^t w}{1 + |c^t x|}.$$

Possivelmente, ao utilizar a direção continuada tanto na forma avançada quanto na forma atrasada, a redução dos resíduos primal e dual e os produtos complementares não ocorre simultaneamente. Dessa forma, sugerimos o critério a seguir para utilização da iteração continuada, buscando uma melhoria nos resíduos e nos produtos complementares de forma simultânea:

$$\left\| \begin{array}{c} \hat{r}_p^l \\ \hat{r}_u^l \\ \hat{r}_d^l \\ \hat{\lambda} \end{array} \right\|_2 < \omega \left\| \begin{array}{c} \hat{r}_p^{l-1} \\ \hat{r}_u^{l-1} \\ \hat{r}_d^{l-1} \\ \lambda \end{array} \right\|_2, \quad (3.63)$$

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

**Observação 3.4.1.** Vale ressaltar que as duas formas de utilização da direção continuada resultam em aumento dos tamanhos dos passos nas componentes das direções que realizaram bloqueio. No entanto, nas componentes diferentes do bloqueio, enquanto a forma avançada reduz os tamanhos dos passos  $\alpha_p$  e  $\alpha_d$ , a forma atrasada mantém estes valores.

O método de pontos interiores com as múltiplas correções de centralidade utilizando a iteração continuada é descrito no Algoritmo 5. Observe que nos passos 6 e 7 representam as alterações no método preditor corretor em relação a abordagem proposta.

---

**Algoritmo 5** Método preditor corretor com múltiplas correções de centralidade utilizando a iteração continuada

---

**Parâmetros**

$\tau = 0,99995$ ,  $\varepsilon_p$ ,  $\varepsilon_d$  e  $\varepsilon_o$ ;

**Inicialização**

$k = 0$ ,  $(x^0, s^0, y^0, z^0, w^0)$ , com  $x^0 > 0$ ,  $s^0 > 0$ ,  $z^0 > 0$ ,  $w^0 > 0$ ;

$\gamma^0 = (x^0)^t z^0 + (s^0)^t w^0$ ;

$r_p^0 = b - Ax^0$ ,  $r_d^0 = c - A^t y^0 - z^0 + w^0$ ,  $r_u^0 = u - x^0 - s^0$ ;

**Processo Iterativo**

1. Enquanto  $\left( \frac{\|r_p^k\|}{1 + \|b\|} > \varepsilon_p \text{ ou } \frac{\|r_u^k\|}{1 + \|u\|} > \varepsilon_p \text{ ou } \frac{\|r_d^k\|}{1 + \|c\|} > \varepsilon_d \text{ ou } \frac{\gamma^k}{1 + |c^t x^k|} > \varepsilon_o \right)$  faça

2. Encontre a direção afim escala  $\bar{d}$  resolvendo o sistema (2.6);

3. Encontre a direção de centragem e correção não linear  $\tilde{d}$  resolvendo o sistema (2.22);

4. Faça  $d = \tilde{d} + \bar{d}$ ;

5. Encontre os tamanhos dos passos primal e dual por (2.21);

6. Faça a iteração continuada (Algoritmo 3 ou 4);

7. Faça as correções de centralidade (Algoritmo 2);

8. Atualize o ponto

$$x^{k+1} = x^k + \alpha_p dx;$$

$$s^{k+1} = s^k + \alpha_p dz;$$

$$y^{k+1} = y^k + \alpha_d dy;$$

$$z^{k+1} = z^k + \alpha_d dz;$$

$$w^{k+1} = w^k + \alpha_d dw;$$

9. Calcule os resíduos e  $\gamma^{k+1}$

$$r_p^{k+1} = b - Ax^{k+1};$$

$$r_d^{k+1} = c - A^t y^{k+1};$$

$$r_u^{k+1} = u - x^{k+1} - s^{k+1};$$

$$\gamma^{k+1} = (x^{k+1})^t z^{k+1} + (s^{k+1})^t w^{k+1};$$

10. Atualize  $k = k + 1$ ;

9. Fim enquanto

---

## Capítulo 4

# Pontos iniciais para o método de pontos interiores

Para utilizar os métodos de pontos interiores somente é necessário fornecer um ponto inicial que satisfaça as condições de não negatividade. Por outro lado, a convergência do método é sensível a este ponto. Assim, o desempenho dos métodos de pontos interiores pode melhorar se um bom ponto inicial for utilizado. Neste capítulo, apresentamos o algoritmo de ajustamento ótimo para  $p$  coordenadas combinado com uma heurística proposta por Mehrotra [31] para determinar melhores pontos iniciais para os métodos de pontos interiores.

### 4.1 Algoritmo de ajustamento ótimo para $p$ coordenadas

O algoritmo de ajustamento ótimo para  $p$  coordenadas foi proposto por Silva [38] e é uma generalização do algoritmo de ajustamento pelo par ótimo [19]. Estes são algoritmos baseados nas ideias do algoritmo de von Neumann, proposto para encontrar uma solução factível para o seguinte conjunto de restrições:

$$\begin{aligned} Bx &= 0 \\ e^t x &= 1 \\ x &\geq 0, \end{aligned} \tag{4.1}$$

em que  $B \in \mathbb{R}^{m \times n}$ ,  $x \in \mathbb{R}^n$ ,  $e \in \mathbb{R}^n$  é o vetor com todas as coordenadas iguais a um e toda coluna  $B_j$  de  $B$ ,  $j = 1, \dots, n$ , tem  $\|B_j\| = 1$ .

O problema (4.1) pode ser interpretado, geometricamente, como encontrar um ponto em uma hipersfera  $m$  dimensional com raio unitário e centro na origem. Além disso, resolver esse problema pode ser visto como atribuir ponderações  $x_j$  não negativas às colunas  $B_j$  de modo que, depois de reescalado, seu centro de gravidade seja a origem.

As vantagens do algoritmo de ajustamento para  $p$  coordenadas para resolver o problema (4.1) são possuir um avanço inicial rápido e ser simples, uma vez que este algoritmo envolve apenas multiplicação de matriz por vetor e solução de um sistema linear com uma matriz definida positiva de dimensão pequena. Além disso, sua convergência é mais rápida que do algoritmo de ajustamento pelo par ótimo e do algoritmo de von Neumann, embora ainda seja lenta.

Todo problema de programação linear pode ser reescrito na forma (4.1) [18, 19]. Essa transformação é apresentada a seguir. Assim, podemos utilizar o algoritmo de ajustamento para  $p$  coordenadas para resolver problemas de programação linear.

Transformação do problema de programação linear para a forma (4.1)

Considere o par de problemas primal e dual, dados em (2.1) e (2.2). Obter uma solução ótima para este par de problemas é equivalente a obter uma solução factível para o seguinte problema:

$$\begin{aligned} Ax &= b \\ x + s &= u \\ A^t y + z - w &= c \\ c^t x - b^t y - u^t w &= 0 \\ x, s, z, w &\geq 0. \end{aligned} \quad (4.2)$$

Como  $y$  é uma variável livre, podemos escrever  $y = y^+ - y^-$ , para  $y^+ \geq 0$  e  $y^- \geq 0$ . Dessa forma, reescrevemos (4.2) como segue:

$$\begin{aligned} Ax &= b \\ x + s &= u \\ A^t y^+ - A^t y^- + z - w &= c \\ c^t x - b^t y^+ + b^t y^- - u^t w &= 0 \\ x, s, y^+, y^-, z, w &\geq 0. \end{aligned} \quad (4.3)$$

Escrevendo (4.3) na forma matricial, temos:

$$\begin{aligned} \hat{A}\hat{x} &= \hat{b} \\ \hat{x} &\geq 0, \end{aligned}$$

em que

$$\hat{A} = \begin{bmatrix} A & 0 & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 & 0 \\ 0 & 0 & A^t & -A^t & I & -I \\ c^t & 0 & -b^t & b^t & 0 & -u^t \end{bmatrix}, \quad \hat{x} = \begin{bmatrix} x \\ s \\ y^+ \\ y^- \\ z \\ w \end{bmatrix} \quad \text{e} \quad \hat{b} = \begin{bmatrix} b \\ u \\ c \\ 0 \end{bmatrix}.$$

**Teorema 4.1.1.** Considere a região factível definida pelas restrições

$$\begin{aligned}\hat{A}\hat{x} &= \hat{b} \\ \hat{x} &\geq 0,\end{aligned}\tag{4.4}$$

$$\begin{aligned}\hat{A}\hat{x} - \hat{b}\hat{v} &= 0 \\ e^t\hat{x} + \hat{v} &= 1 \\ \hat{x}, \hat{v} &\geq 0,\end{aligned}\tag{4.5}$$

em que  $\hat{v} \in \mathbb{R}$ .

(i) Se  $\hat{x}, \hat{v}$ , em que  $\hat{v} \neq 0$ , é uma solução factível para (4.5), então  $\tilde{x} = \hat{x}/\hat{v}$  é uma solução factível para (4.4).

(ii) Se  $\tilde{x}$  é uma solução factível para (4.4) então, para  $\hat{v} = 1/e^t\tilde{x} + 1$ ,  $\hat{x} = \hat{v}\tilde{x}$  é uma solução factível para (4.5).

*Demonstração.*

(i) Seja  $\hat{x}, \hat{v}$ , em que  $\hat{v} \neq 0$ , uma solução factível para (4.5). Então,  $\hat{A}\hat{x} - \hat{b}\hat{v} = 0$ . Como  $\hat{v} \neq 0$ , segue que  $\hat{A}(\hat{x}/\hat{v}) = \hat{b}$ . Além disso, como  $\hat{x} \geq 0$ ,  $\hat{v} > 0$ , então  $\tilde{x} = \hat{x}/\hat{v}$  é uma solução factível para (4.4).

(ii) Seja  $\tilde{x}$  uma solução factível para (4.4). Então, de  $A\tilde{x} = b$ , segue que  $A\hat{v}\tilde{x} - \hat{v}b = 0$ , para  $\hat{v} \geq 0$ . Além disso,  $e^t(\hat{v}\tilde{x}) + \hat{v} = (e^t\tilde{x} + 1)\hat{v}$ . Dessa forma, para  $\hat{v} = 1/(e^t\tilde{x} + 1)$ , temos  $e^t(\hat{v}\tilde{x}) + \hat{v} = 1$ . Logo,  $\hat{x} = \hat{v}\tilde{x}$  é uma solução factível para (4.5).  $\square$

Pelo Teorema (4.1.1), uma solução factível do problema (4.4) é solução do problema (4.5) e vice-versa. Então, consideremos o problema (4.5), que pode ser reescrito como:

$$\begin{aligned}\tilde{A}\tilde{x} &= 0 \\ e^t\tilde{x} &= 1, \\ \tilde{x} &\geq 0\end{aligned}\tag{4.6}$$

em que  $\tilde{A} = \begin{bmatrix} \hat{A} & -\hat{b} \end{bmatrix}$  e  $\tilde{x} = \begin{bmatrix} \hat{x}^t & \hat{v} \end{bmatrix}^t$ .

No problema (4.6), para cada coluna  $\tilde{A}_j$  da matriz  $\tilde{A}$ ,  $j = 1, \dots, \bar{n}$ , em que  $\bar{n} = 4n + 2m + 1$ , consideremos:

$$B_j = \frac{\tilde{A}_j}{\|\tilde{A}_j\|}.$$

Então:

$$\begin{aligned}B\bar{x} &= 0 \\ e^t\bar{x} &= 1 \\ \bar{x} &\geq 0.\end{aligned}\tag{4.7}$$

em que

$$\bar{x} = \begin{bmatrix} x \\ s \\ y^+ \\ y^- \\ z \\ w \\ v \end{bmatrix}, \quad (4.8)$$

$x, s, z, w \in \mathbb{R}^n$ ,  $y^-, y^+ \in \mathbb{R}^m$  e  $v \in \mathbb{R}$ .

Pelo Teorema (4.1.2) a seguir, uma solução factível do problema (4.6) é solução do problema (4.7) e vice-versa. Assim o problema na forma (4.1) é obtido.

**Teorema 4.1.2.** Considere a região factível definida pelas seguintes restrições:

$$\begin{aligned} \tilde{A}\tilde{x} &= \tilde{b} \\ e^t\tilde{x} &= 1, \\ \tilde{x} &\geq 0 \end{aligned} \quad (4.9)$$

$$\begin{aligned} B\bar{x} &= 0 \\ e^t\bar{x} &= 1, \\ \bar{x} &\geq 0 \end{aligned} \quad (4.10)$$

(i) Se  $\bar{x}$  é uma solução factível para (4.10), então  $\tilde{x}$  é uma solução factível para (4.9), em que

$$\tilde{x}_j = \left( \frac{\bar{x}_j}{\|\tilde{A}_j\|} \right) \bigg/ \left( \sum_{k=1}^{\bar{n}} \frac{\bar{x}_k}{\|\tilde{A}_k\|} \right).$$

(ii) Se  $\tilde{x}$  é uma solução factível para (4.9), então  $\bar{x}$  é uma solução factível para (4.10), em que

$$\bar{x}_j = (\tilde{x}_j \|\tilde{A}_j\|) \bigg/ \left( \sum_{k=1}^{\bar{n}} \tilde{x}_k \|\tilde{A}_k\| \right).$$

*Demonstração.*

(i) Sejam  $\bar{x}$  uma solução factível para (4.10) e

$$\tilde{x}_j = \left( \frac{\bar{x}_j}{\|\tilde{A}_j\|} \right) \bigg/ \left( \sum_{k=1}^{\bar{n}} \frac{\bar{x}_k}{\|\tilde{A}_k\|} \right).$$

Então:

$$\begin{aligned}
B\bar{x} = 0 &\Leftrightarrow \sum_{j=1}^{\bar{n}} B_j \bar{x}_j = 0 \\
&\Leftrightarrow \sum_{j=1}^{\bar{n}} \frac{\tilde{A}_j}{\|\tilde{A}_j\|} \bar{x}_j = 0 \\
&\Leftrightarrow \frac{1}{\left(\sum_{k=1}^{\bar{n}} \frac{\bar{x}_k}{\|\tilde{A}_k\|}\right)} \sum_{j=1}^{\bar{n}} \frac{\tilde{A}_j}{\|\tilde{A}_j\|} \bar{x}_j = 0 \\
&\Leftrightarrow \sum_{j=1}^{\bar{n}} \frac{\tilde{A}_j}{\|\tilde{A}_j\|} \bar{x}_j \Big/ \left(\sum_{k=1}^{\bar{n}} \frac{\bar{x}_k}{\|\tilde{A}_k\|}\right) = 0 \\
&\Leftrightarrow \tilde{A}\tilde{x} = 0.
\end{aligned}$$

Além disso,

$$\begin{aligned}
e^t \tilde{x} &= \sum_{j=1}^{\bar{n}} \tilde{x}_j = \sum_{j=1}^{\bar{n}} \left(\frac{\bar{x}_j}{\|\tilde{A}_j\|}\right) \Big/ \left(\sum_{k=1}^{\bar{n}} \frac{\bar{x}_k}{\|\tilde{A}_k\|}\right) = \\
&= \frac{1}{\left(\sum_{k=1}^{\bar{n}} \frac{\bar{x}_k}{\|\tilde{A}_k\|}\right)} \left(\sum_{j=1}^{\bar{n}} \frac{\bar{x}_j}{\|\tilde{A}_j\|}\right) = \\
&= 1.
\end{aligned}$$

Como  $\bar{x}_j \geq 0$ , para todo  $j = 1, \dots, \bar{n}$ , segue que  $\tilde{x}_j \geq 0$ . Então,  $\tilde{x}$  é solução factível de (4.9).

(ii) Sejam  $\tilde{x}$  uma solução factível de (4.9) e

$$\bar{x}_j = (\bar{x}_j \|\tilde{A}_j\|) \Big/ \left(\sum_{k=1}^{\bar{n}} \bar{x}_k \|\tilde{A}_k\|\right).$$

Então,

$$\begin{aligned}
\tilde{A}\tilde{x} = 0 &\Leftrightarrow \sum_{j=1}^{\bar{n}} \tilde{A}_j \tilde{x}_j = 0 \\
&\Leftrightarrow \frac{1}{\left(\sum_{k=1}^{\bar{n}} \frac{\tilde{x}_k}{\|\tilde{A}_k\|}\right)} \sum_{j=1}^{\bar{n}} \tilde{A}_j \tilde{x}_j = 0 \\
&\Leftrightarrow \frac{1}{\left(\sum_{k=1}^{\bar{n}} \frac{\tilde{x}_k}{\|\tilde{A}_k\|}\right)} \sum_{j=1}^{\bar{n}} \frac{\tilde{A}_j}{\|\tilde{A}_j\|} \tilde{x}_j \|\tilde{A}_j\| = 0 \\
&\Leftrightarrow \sum_{j=1}^{\bar{n}} \frac{\tilde{A}_j}{\|\tilde{A}_j\|} \tilde{x}_j \|\tilde{A}_j\| \Big/ \left(\sum_{k=1}^{\bar{n}} \frac{\tilde{x}_k}{\|\tilde{A}_k\|}\right) = 0 \\
&\Leftrightarrow \sum_{j=1}^{\bar{n}} B_j \tilde{x}_j \|\tilde{A}_j\| \Big/ \left(\sum_{k=1}^{\bar{n}} \frac{\tilde{x}_k}{\|\tilde{A}_k\|}\right) = 0 \\
&\Leftrightarrow B\bar{x} = 0.
\end{aligned}$$

E

$$\begin{aligned}
e^t \bar{x} &= \sum_{j=1}^{\bar{n}} \bar{x}_j = \sum_{j=1}^{\bar{n}} (\tilde{x}_j \|\tilde{A}_j\|) \Big/ \left(\sum_{k=1}^{\bar{n}} \tilde{x}_k \|\tilde{A}_k\|\right) = \\
&= \frac{1}{\left(\sum_{k=1}^{\bar{n}} \tilde{x}_k \|\tilde{A}_k\|\right)} \left(\sum_{j=1}^{\bar{n}} \tilde{x}_j \|\tilde{A}_j\|\right) = \\
&= 1.
\end{aligned}$$

Logo, como  $\tilde{x}_j \geq 0$ , para todo  $j = 1, \dots, \bar{n}$ , segue que  $\bar{x}_j \geq 0$ . Portanto,  $\bar{x}$  é solução factível de (4.10).  $\square$

O algoritmo de ajustamento ótimo para  $p$  coordenadas

O algoritmo de ajustamento ótimo para  $p$  coordenadas é uma generalização do algoritmo de ajustamento pelo par ótimo para resolver o problema dado em (4.1). O algoritmo de ajustamento pelo par ótimo é um método iterativo baseado na ideia de que, a partir de um  $x^0$  inicial, o resíduo  $b^{k-1} = Bx^{k-1}$ ,  $k = 1, 2, \dots$ , pode ser movido de tal forma a aproximar-se da origem 0. Isto é feito aumentando um peso  $x_i$  de alguma coluna  $B_i$  e reduzindo um peso  $x_j$  de certa coluna  $B_j$ , sendo as demais variáveis ajustadas de

acordo com  $x_i$  e  $x_j$  determinados. Para o resíduo  $b^k$ , é esperado que esteja mais próximo da origem comparado com o resíduo anterior  $b^{k-1}$ . Então, o algoritmo de ajustamento pelo par ótimo identifica as colunas  $B_{s_+}$  e  $B_{s_-}$  de  $B$  que formam o maior e o menor ângulo com o vetor  $b^{k-1}$ , respectivamente, depois determina  $x_{s_+}^k$  e  $x_{s_-}^k$  que minimizam a distância de  $b^k$  a origem satisfazendo a convexidade e as restrições de não negatividade do problema (4.1).

O algoritmo de ajustamento para  $p$  coordenadas, generaliza este algoritmo identificando as  $s_1$  e  $s_2$  colunas da matriz  $B$  que fazem o maior e o menor ângulo com o vetor  $b^{k-1} = Bx^{k-1}$ , respectivamente, de forma que, para o número de colunas  $p$  a ser priorizada, se tenha  $s_1 + s_2 = p$ . Depois, um subproblema de otimização é resolvido buscando minimizar a distância de  $b^k = Bx^k$  à origem, satisfazendo a convexidade e as restrições de não negatividade. A resolução do subproblema é realizada utilizando método de pontos interiores. Apresentamos em 6, de forma resumida, este algoritmo.

---

**Algoritmo 6** Algoritmo de ajustamento ótimo para  $p$  coordenadas

---

**Parâmetros**

$B$ ,  $x^0 \geq 0$ , com  $e^t x^0 = 1$

**Inicialização**

$b^0 = Bx^0$

**Processo Iterativo**

1. Para  $k = 1, 2, 3, \dots$

2. Encontre:

$s_1$  colunas de  $B$  que formam o maior ângulo com  $b^{k-1}$  :  $\{B_{\eta_1^+}, B_{\eta_2^+}, \dots, B_{\eta_{s_1}^+}\}$

$s_2$  colunas de  $B$  que formam o menor ângulo com  $b^{k-1}$  :  $\{B_{\eta_1^-}, B_{\eta_2^-}, \dots, B_{\eta_{s_2}^-}\}$  tais que

$x_i^{k-1} > 0, i = \eta_1^-, \dots, \eta_{s_2}^-$ .

$v^{k-1} = \min_{i=\eta_1^+, \dots, \eta_{s_1}^+} B_{i^+}^t b^{k-1}$

3. Se  $v^{k-1} > 0$  então PARE. O problema (4.1) é infactível

4. Resolva o subproblema:

$\min \|\bar{b}\|^2$

s.a  $\lambda_0 \left( 1 - \sum_{i=1}^{s_1} x_{\eta_i^+}^{k-1} - \sum_{j=1}^{s_2} x_{\eta_j^-}^{k-1} \right) + \sum_{i=1}^{s_1} \lambda_{\eta_i^+}^{k-1} + \sum_{i=1}^{s_1} \lambda_{\eta_j^-}^{k-1} = 1,$

$\lambda_{\eta_i^+}^{k-1} \geq 0$ , para  $i = \eta_1, \dots, \eta_{s_1}$ ,

$\lambda_{\eta_j^-}^{k-1} \geq 0$ , para  $j = \eta_1, \dots, \eta_{s_2}$ ,

em que

$\bar{b} = \lambda_0 \left( b^{k-1} - \sum_{i=1}^{s_1} x_{\eta_i^+}^{k-1} B_{\eta_i^+} - \sum_{j=1}^{s_2} x_{\eta_j^-}^{k-1} B_{\eta_j^-} \right) + \sum_{i=1}^{s_1} \lambda_{\eta_i^+}^{k-1} B_{\eta_i^+} + \sum_{i=1}^{s_1} \lambda_{\eta_j^-}^{k-1} B_{\eta_j^-}.$

5. Atualize:

$b^k = \lambda_0 \left( b^{k-1} - \sum_{i=1}^{s_1} x_{\eta_i^+}^{k-1} B_{\eta_i^+} - \sum_{j=1}^{s_2} x_{\eta_j^-}^{k-1} B_{\eta_j^-} \right) + \sum_{i=1}^{s_1} \lambda_{\eta_i^+}^{k-1} B_{\eta_i^+} + \sum_{i=1}^{s_1} \lambda_{\eta_j^-}^{k-1} B_{\eta_j^-}$

$x_j^k = \begin{cases} \lambda_0 x_j^{k-1}, & j \notin \{\eta_1^+, \dots, \eta_{s_1}^+, \eta_1^-, \dots, \eta_{s_2}^-\} \\ \lambda_{\eta_i^+}, & j = \eta_i^+, i = 1, \dots, s_1, \\ \lambda_{\eta_j^-}, & j = \eta_j^-, j = 1, \dots, s_2, \end{cases}$

$k = k + 1.$

6. Fim

---

No Passo 3 do Algoritmo 6, se  $v_{k-1} > 0$ , então todas as colunas  $b_j$  da matriz  $B$  estão do mesmo lado do hiperplano que passa pela origem e é perpendicular ao vetor  $b^{k-1}$ . Desse modo, não é possível encontrar nenhuma combinação convexa das colunas da matriz  $B$  que resulte na origem. Assim, o problema é infactível.

O subproblema do Passo 4 de (6) pode ser resolvido utilizando métodos de pontos interiores. Considere:

$$\begin{aligned} Q &= \left[ \bar{q} \quad B_{\eta_1^+} \quad B_{\eta_2^+} \quad \dots \quad B_{\eta_{s_1}^+} \quad B_{\eta_1^-} \quad B_{\eta_2^-} \quad \dots \quad B_{\eta_{s_2}^-} \right] \\ \bar{q} &= b^{k-1} - \sum_{i=1}^{s_1} x_{\eta_i^+}^{k-1} B_{\eta_i^+} - \sum_{j=1}^{s_2} x_{\eta_j^-}^{k-1} B_{\eta_j^-}, \\ \lambda &= \left( \lambda_0, \lambda_{\eta_1^+}, \dots, \lambda_{\eta_{s_1}^+}, \lambda_{\eta_1^-}, \dots, \lambda_{\eta_{s_2}^-} \right), \\ a &= \left( 1 - \sum_{i=1}^{s_1} x_{\eta_i^+}^{k-1} - \sum_{j=1}^{s_2} x_{\eta_j^-}^{k-1}, 1, \dots, 1 \right). \end{aligned}$$

Então, podemos reescrever o subproblema como:

$$\begin{aligned} \min \quad & \frac{1}{2} \|Q\lambda\|^2 \\ \text{s.a} \quad & a^t \lambda = 1 \\ & -\lambda \leq 0. \end{aligned} \tag{4.11}$$

As condições de KKT para subproblema (4.11) são dadas por:

$$\begin{aligned} Q^t Q \lambda + a \theta - v &= 0, \\ v^t \lambda &= 0, \\ a^t \lambda - 1 &= 0, \\ -\lambda &\leq 0, \end{aligned} \tag{4.12}$$

em que  $Q^t Q \in \mathbb{R}^{(p+1) \times (p+1)}$   $\theta$  é livre,  $v \geq 0$  são os multiplicadores de Lagrange.

Aplicando o métodos de pontos interiores a (4.12), obtemos o seguinte sistema linear para obter a direção em cada iteração do método:

$$\begin{bmatrix} Q^t Q & a & -I \\ V & 0 & U \\ a^t & 0 & 0 \end{bmatrix} \begin{bmatrix} d\lambda \\ d\theta \\ dv \end{bmatrix} = \begin{bmatrix} r_a \\ r_b \\ r_c \end{bmatrix} \tag{4.13}$$

em que  $V = \text{diag}(\theta)$ ,  $U = \text{diag}(v)$ ,  $r_a = v - a\theta - Q^t Q \lambda$ ,  $r_b = -v\lambda$  e  $r_c = 1 - a^t \lambda$ .

Pela segunda equação de (4.13), temos:

$$V d\theta + U dv = r_b.$$

Como  $U$ , é uma matriz diagonal positiva, segue que:

$$dv = U^{-1}(r_b - V\lambda).$$

Pela primeira equação de (4.13), temos:

$$Q^t Q d\lambda + ad\theta - dv = r_a.$$

Substituindo o valor de  $dv$ , obtemos:

$$Q^t Q d\lambda + ad\theta - U^{-1}(r_b - V\lambda) = r_a. \quad (4.14)$$

Segue que:

$$(Q^t Q + U^{-1}V)d\lambda = r_a + U^{-1}r_b - ad\theta.$$

Desse modo,

$$d\lambda = (Q^t Q + U^{-1}V)^{-1}r_a + U^{-1}r_b - ad\theta.$$

Já pela última equação de (4.13), temos

$$a^t d\lambda = r_c.$$

Logo,

$$r_c = a^t d\lambda = a^t (Q^t Q + U^{-1}V)^{-1} (r_a + U^{-1}r_b - ad\theta).$$

Então,

$$a^t (Q^t Q + U^{-1}V)^{-1} ad\theta = a^t (Q^t Q + U^{-1}V)^{-1} (r_a + U^{-1}r_b - r_c).$$

Portanto, para calcular as direções é necessário resolver os seguintes sistemas lineares:

$$\begin{aligned} (Q^t Q + U^{-1}V)\tilde{v} &= a \\ (Q^t Q + U^{-1}V)\tilde{u} &= r_a + U^{-1}r_b. \end{aligned}$$

Visto que a matriz  $Q^t Q + U^{-1}V$  é definida positiva de ordem  $(p+1) \times (p+1)$ , resolver estes sistemas lineares não apresenta dificuldades para valores pequenos de  $p$ . Caso  $p$  seja um número grande, determinar o fator de Cholesky para esta matriz requer um alto custo computacional e resolver estes sistemas lineares dessa forma não é recomendado. Em [37], é proposta uma escolha do parâmetro  $p$ , sendo dada por:

$$p = \varrho \sqrt{mn},$$

em que  $\varrho$  é uma constante de ajuste,  $m$  é o número de linhas e  $n$  é o número de colunas da matriz do problema de programação linear a ser resolvido. Em [37], o valor usado para  $\varrho$  foi  $2.10^{-2}$ .

Como a convergência deste método é lenta, em vez de utilizarmos este algoritmo até encontrar um solução ótima para os problemas de programação linear, podemos utilizá-lo até que os resíduos satisfaçam uma certa tolerância. Dessa forma, a aplicação do algoritmo de ajustamento para  $p$  coordenadas pode ser utilizada para determinar melhores pontos iniciais para os métodos de pontos interiores. No entanto, este algoritmo também necessita de um ponto inicial, podendo a heurística de Mehrotra [30] ser utilizada para determinar este ponto inicial. Mehrotra propõe uma heurística para o determinar o ponto inicial de modo que este seja centralizado, conforme apresentado, a seguir.

## 4.2 Heurística de Mehrotra

A heurística de Mehrotra, utilizada para determinar um ponto inicial para o método de pontos interiores no código do PCx, é dada resumidamente no Algoritmo 7.

---

### Algoritmo 7 Heurística de Mehrotra para o cálculo do ponto inicial

---

1. Calcule os pontos:

$$\tilde{x} = \frac{A^t(AA^t)^{-1}(2b - Au) + u}{2}$$

$$\tilde{s} = u - \tilde{x}$$

$$\tilde{y} = (AA^t)^{-1}Ac$$

$$\tilde{z} = \frac{c - A^t\tilde{y}}{2}$$

$$\tilde{w} = -\tilde{z}$$

2. Encontre os valores  $\delta_p$  e  $\delta_d$  tais que  $(x, s) + (\delta_p, \delta_p)$  e  $(z, w) + (\delta_d, \delta_d)$  sejam não-negativos:

$$\delta_p = \max \{-1.5 \min \tilde{x}_i, -1.5 \min \tilde{s}_i, 0\}$$

$$\delta_d = \max \{-1.5 \min \tilde{z}_i, -1.5 \min \tilde{w}_i, 0\}$$

3. Determine  $\tilde{\delta}_p$  e  $\tilde{\delta}_d$  tais que os pontos  $x^0$ ,  $s^0$ ,  $z^0$  e  $w^0$  sejam centralizados:

$$\tilde{\delta}_p = \delta_p + \frac{1}{2} \frac{(\tilde{x} + \delta_p e)^t (\tilde{z} + \delta_d e) + (\tilde{s} + \delta_p e)^t (\tilde{w} + \delta_d e)}{\sum_{i=1}^n (\tilde{z}_i + \tilde{w}_i + 2\delta_d)},$$

$$\tilde{\delta}_d = \delta_d + \frac{1}{2} \frac{(\tilde{x} + \delta_p e)^t (\tilde{z} + \delta_d e) + (\tilde{s} + \delta_p e)^t (\tilde{w} + \delta_d e)}{\sum_{i=1}^n (\tilde{x}_i + \tilde{s}_i + 2\delta_d)},$$

4. Calcule os pontos iniciais:

$$x^0 = \tilde{x} + \tilde{\delta}_p$$

$$s^0 = \tilde{s} + \tilde{\delta}_p$$

$$z^0 = \tilde{z} + \tilde{\delta}_d$$

$$w^0 = \tilde{w} + \tilde{\delta}_d$$

$$y^0 = \tilde{y}$$


---

Como o Algoritmo 7 calcula pontos centralizados no Passo 3 e a centralidade é importante para os métodos de pontos interiores, então algumas iterações do algoritmo de ajustamento ótimo para  $p$  coordenadas são realizadas antes da etapa de centralização, ou seja, entre o Passo 2 e o Passo 3. Mais detalhes sobre a heurística de Mehrotra podem ser encontrados em [31] e sobre o algoritmo de ajustamento ótimo para  $p$  coordenadas em [16, 37].

# Capítulo 5

## Experimentos computacionais

Neste capítulo, apresentamos os experimentos computacionais realizados ao incorporar a iteração continuada no método de pontos interiores preditor corretor com múltiplas correções de centralidade e utilizando o algoritmo de ajustamento ótimo para  $p$  coordenadas para determinar bons pontos iniciais para o mesmo método.

### 5.1 Código PCx

O código PCx [13], desenvolvido no Optimization Technology Center at Argonne National Laboratory and Northwestern University, é uma implementação do método de pontos interiores preditor corretor com múltiplas correções [21]. A maioria das suas rotinas está implementada na linguagem C. As rotinas responsáveis pela solução dos sistemas lineares, que utiliza uma fatoração esparsa de Cholesky desenvolvida por Ng e Peyton [33], estão implementadas em FORTRAN77. O PCx utiliza os problemas de programação linear no formato MPS, podendo os problemas conter restrições de igualdade e desigualdade, variáveis canalizadas ou livres. Uma rotina de preprocessamento do PCx converte os problemas em um formato padrão, e após a solução ser determinada, esta é transformada em termos da formulação original.

### 5.2 Implementação

A direção continuada em suas diferentes formas de uso, tanto com a direção moderada quanto com a direção simples, foi implementada em linguagem C e incorporada ao código do PCx [13] com a opção das múltiplas correções ligada.

Os experimentos computacionais foram realizados em um processador Intel Core i7, 8 GB RAM, 1TB HD e Sistema Operacional Linux.

O parâmetro adotado para que a iteração continuada seja utilizada de acordo com o Critério (3.63) é  $\omega = 0,96$ . O valor *default* do PCx para otimalidade é  $\epsilon_o =$

$10^{-10}$ ,  $\epsilon_p = 10^{-8}$  e  $\epsilon_d = 10^{-8}$ . Para o algoritmo de ajustamento ótimo para  $p$  coordenadas é utilizado  $p = 0,025\sqrt{m+n}$ , o número máximo de iterações é igual a 2 e a tolerância é  $10^{-4}$ .

### 5.3 Problemas teste

Os experimentos computacionais foram realizados com problemas de livre acesso na internet. Foram selecionados 38 problemas de médio a grande porte. Na maioria deles, o PCx leva mais de cinco segundos para encontrar uma solução ótima. Os problemas teste pertencem às seguintes coleções:

- Netlib [9]

Nesta coleção, selecionamos o problema dfl001, um problema de atribuição de frota de uma companhia aérea, disponível em:

<http://www.netlib.org/lp/>

- Mészáros

Os problemas selecionados desta coleção estão disponíveis em:

[http://www.sztaki.hu/~meszaros/public\\_ftp/lptestset/](http://www.sztaki.hu/~meszaros/public_ftp/lptestset/)

- PDS: Patient distribution (evacuation) system, [10].

Disponível em:

<http://plato.asu.edu/ftp/lptestset/pds/>

- Fome

Disponível em:

<http://plato.asu.edu/ftp/lptestset/fome/>

- Watson originário de programação linear estocástica.

Disponível em:

<http://plato.asu.edu/ftp/lptestset/misc/>

Na primeira coluna da Tabela 1, está o nome dos problemas selecionados, na segunda coluna, a dimensão dos problemas (número de linhas e colunas) após o pré-processamento do PCx e na terceira coluna, está a densidade do fator de Cholesky, fornecida pelo PCx, que realiza um reordenamento procurando que este fator seja esparso. A última coluna mostra a coleção a que o problema pertence.

Tabela 1 – Problemas Teste

Problema	Dimensão		Densidade	Coleção
	Linhas	Colunas		
baxter	23871	30122	0,02064	Mészáros
dano3mip	3202	15851	0,21235	Mészáros
dbic1	34205	174457	0,00316	Mészáros
dbir1	14032	38763	0,02556	Mészáros
dbir2	14107	38838	0,02825	Mészáros
ex3sta1	17442	17515	0,16834	Mészáros
lp22	2958	16392	0,21515	Mészáros
lp11	34125	89383	0,00162	Mészáros
nsct1	15278	26430	0,04636	Mészáros
nsct2	15360	26512	0,04779	Mészáros
rlfprim	57422	61268	0,18400	Mészáros
route	20894	43019	0,01422	Mészáros
seymour1	4827	6082	0,38727	Mészáros
slptsk	2861	3347	0,99846	Mészáros
south31	17983	35686	0,03502	Mészáros
Ts-palko	22002	47235	0,06484	Mészáros
tp-6	142752	1014301	0,00280	Mészáros
degme	185501	659415	0,00325	Mészáros
df1001	5984	12143	0,09132	Netlib
pds06	9156	28472	0,01395	PDS
pds10	15648	48780	0,01372	PDS
pds20	38722	106180	0,01357	PDS
pds30	47968	156042	0,01271	PDS
pds40	64276	214385	0,01363	PDS
pds50	80339	272513	0,01293	PDS
pds60	96514	332862	0,01247	PDS
pds70	111896	386238	0,01257	PDS
pds80	126120	430800	0,01185	PDS
pds90	139752	471538	0,01136	PDS
pds100	152300	498530	0,01036	PDS
fome11	11968	24286	0,04566	Fome
fome12	23936	48572	0,02283	Fome
fome13	47872	97144	0,01142	Fome
fome20	32287	106180	0,01357	Fome
fome21	64574	212360	0,00679	Fome
Watson-1	119782	233426	0,00024	Watson
Watson-2	209614	411177	0,00014	Watson

## 5.4 Resultados computacionais

Para analisar o desempenho do PCx ao incorporar a iteração continuada, foram realizados testes com as formas avançada e atrasada da direção continuada, tanto com a direção moderada quanto com a direção simples.

Considere:

AVDM: PCx com a forma avançada da direção moderada.

ATDM: PCx com a forma atrasada da direção moderada.

AVDS: PCx com a forma avançada da direção simples.

ATDS: PCx com a forma atrasada da direção simples.

#### 5.4.1 Comparação AVDM, ATDM , AVDS e ATDS

Inicialmente, foi realizada uma comparação com os métodos AVDM, ATDM , AVDS e ATDS para decidir qual deles pode ser o mais competitivo com PCx. Na Tabela 2, as colunas de AVDM, ATDM, AVDS, ATDS contêm o número total de iterações (k) realizadas pelos métodos, e o tempo total de resolução dos problemas em segundos (t), além do número de iterações continuadas (ic) realizadas.

**Observação 5.4.1.** Nos estudos iniciais da iteração continuada, em que consideramos somente uma componente de bloqueio no espaço primal e outra no espaço dual, a utilização da direção moderada não apresentava resultados competitivos com a direção simples. Ao existir a possibilidade de ter mais de uma componente de bloqueio em cada espaço e analisar a direção moderada, verificamos que mais de dois sistemas lineares com a matriz  $ADA^t$  seria resolvidos, tornando ainda mais caro computacionalmente o cálculo desta direção continuada comparado com a direção simples. Então, não foi considerado em nossa implementação o caso que pode ocorrer mais de uma componente de bloqueio no espaço primal e dual, na direção moderada, pois ao realizar os experimentos computacionais foi verificado que somente em 10 dentre os 38 problemas teste poderia ocorrer este fato. Assim, estes problemas foram excluídos da comparação da Tabela 2.

Tabela 2 – Comparação AVDM, ATDM , AVDS e ATDS.

Problema	AVDM			ATDM			AVDS			ATDS		
	k	ic	t(s)	k	ic	t(s)	k	ic	t(s)	k	ic	t(s)
baxter	<b>26</b>	2	24,09	28	2	25,88	<b>26</b>	14	<b>23,30</b>	<b>26</b>	12	24,34
dano3mip	*	19	*	<b>31</b>	7	7,89	*	9	*	<b>31</b>	13	<b>7,70</b>
dbic1	<b>43</b>	19	13,49	44	18	13,59	46	20	12,93	46	22	<b>12,77</b>
dbir1	24	0	28,76	24	0	28,94	<b>21</b>	3	<b>25,29</b>	24	7	28,34
dbir2	<b>25</b>	5	35,25	26	1	36,28	<b>25</b>	12	<b>34,75</b>	26	8	35,91
ex3sta1	8	8	133,12	<b>7</b>	1	<b>117,36</b>	8	0	131,43	8	2	131,30
lp22	59	4	10,06	35	4	6,32	48	8	8,05	<b>29</b>	7	<b>5,20</b>
nsct1	22	2	82,42	22	2	83,12	22	4	81,43	<b>21</b>	7	<b>78,13</b>
nsct2	30	5	112,09	31	1	117,70	28	12	106,04	<b>26</b>	10	<b>99,22</b>
rlfdual	14	0	6,12	14	0	6,28	14	0	6,06	14	0	6,05
seymour1	<b>15</b>	5	21,88	<b>15</b>	3	<b>21,72</b>	16	8	22,72	16	8	22,73
slptsk	20	4	34,94	<b>18</b>	1	<b>31,60</b>	19	14	32,42	20	12	34,20
south31	<b>15</b>	2	33,30	<b>15</b>	0	33,64	16	7	35,10	<b>15</b>	10	<b>33,00</b>
rlfprim	10	1	5704,09	10	1	5706,79	<b>9</b>	0	<b>5185,99</b>	<b>9</b>	0	5195,92
fome21	48	0	277,29	47	1	275,97	47	5	271,33	<b>46</b>	4	<b>268,21</b>
Ts-palko	19	7	134,23	<b>18</b>	5	126,53	<b>18</b>	11	<b>125,06</b>	19	13	130,10
tp-6	27	24	588,55	26	8	554,14	<b>21</b>	11	<b>448,00</b>	23	17	486,99
degme	27	13	1434,49	<b>24</b>	9	1271,96	<b>24</b>	4	<b>1266,44</b>	25	23	1308,58
karted	15	5	461,79	15	7	460,58	15	5	457,10	15	13	<b>455,30</b>
df001	<b>38</b>	1	13,31	43	1	15,63	42	5	14,31	<b>38</b>	4	<b>12,97</b>
pds-20	44	2	128,17	43	0	129,01	42	3	125,31	<b>40</b>	5	<b>120,09</b>
pds-40	50	6	961,77	50	1	964,09	50	4	954,11	<b>49</b>	11	<b>933,71</b>
pds-70	57	2	4898,49	53	1	4578,97	54	3	4637,79	<b>52</b>	5	<b>4475,60</b>
pds-80	<b>51</b>	1	<b>6300,83</b>	<b>51</b>	0	6307,07	55	7	6743,84	52	14	6388,71
pds-90	<b>51</b>	2	7754,29	52	3	7901,77	<b>51</b>	2	<b>7723,56</b>	53	9	8001,88
pds-100	<b>53</b>	3	<b>8925,26</b>	54	2	9083,84	54	10	9051,68	54	7	9055,51
Watson-1	83	23	11,68	82	13	11,44	<b>79</b>	17	<b>9,16</b>	83	19	9,34
Watson-2	<b>66</b>	28	18,57	<b>66</b>	19	17,72	67	7	14,68	<b>66</b>	8	<b>14,08</b>

\* O método não resolveu o problema

Na Tabela 2, temos os melhores resultados em destaque entre os métodos. Observe que somente ATDM e ATDS resolvem todos os problemas testes. ATDS é o que apresenta o maior número de problemas em destaque, resolvendo 13 problemas teste em menor número de iterações e tempo total. Este bom desempenho de ATDS pode ser confirmado nas Figuras 1 e 2, com o perfil de desempenho em relação ao número de iterações e tempo total, respectivamente, dos resultados da Tabela 2. Pelas Figuras 1 e 2, concluímos que ATDS e AVDS são robustos para esses problemas testes. No entanto, o mais eficiente é ATDS, resolvendo aproximadamente cerca de 58% dos problemas testes em menor número de iterações e 52% em tempo total reduzido. Dessa forma, ATDS é o mais competitivo entre os métodos propostos. A seguir, apresentamos a comparação entre PCx e ATDS.

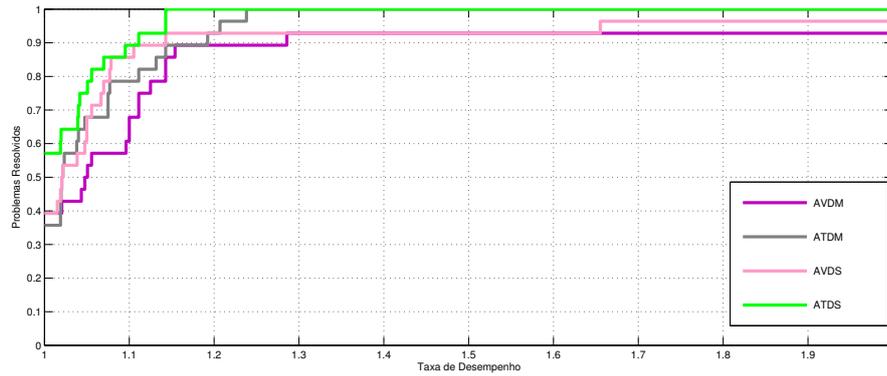


Figura 1 – Número de iterações.

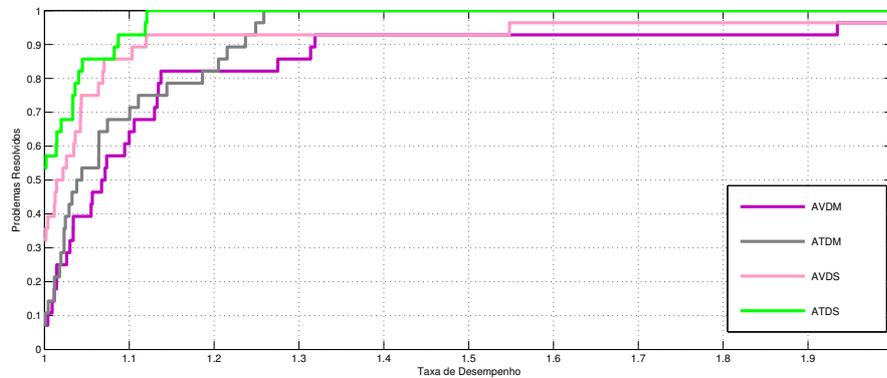


Figura 2 – Tempo total.

**Observação 5.4.2.** A iteração continuada utilizando a direção preditora corretora continuada é proposta em [4] para PCx sem múltiplas correções de centralidade. Ao utilizar as múltiplas correções de centralidade no PCx, esta direção não proporciona resultados satisfatórios, pois dificilmente é realizada. Desta forma, esta direção continuada não se aplica para o PCx com múltiplas correções de centralidade.

### 5.4.2 Comparação PCx e ATDS

Na Tabela 3, temos a comparação dos resultados obtidos por PCx e ATDS em todos os problemas teste.

Tabela 3 – Comparação entre PCx e ATDS.

Problema	PCx		ATDS		
	k	t(s)	k	t(s)	ic
baxter	28	24,92	<b>26</b>	<b>24,34</b>	12
dano3mip	*	*	<b>31</b>	<b>7,70</b>	13
dbic1	<b>45</b>	<b>11,87</b>	46	12,77	22
dbir1	24	<b>27,66</b>	24	28,34	7
dbir2	26	<b>35,06</b>	26	35,91	8
ex3sta1	8	131,72	8	<b>131,30</b>	2
lp22	56	9,24	<b>29</b>	<b>5,20</b>	7
nsct1	21	78,40	21	<b>78,13</b>	7
nsct2	28	105,88	<b>26</b>	<b>99,22</b>	10
rldual	14	<b>5,94</b>	14	6,05	0
route	20	3,31	<b>18</b>	<b>3,02</b>	4
seymour1	<b>15</b>	<b>21,45</b>	16	22,73	8
slptsk	<b>19</b>	<b>32,60</b>	20	34,20	12
south31	15	33,10	15	<b>33,00</b>	10
rldprim	9	<b>5185,61</b>	9	5195,92	0
fome11	<b>45</b>	<b>30,62</b>	47	31,96	4
fome12	43	58,94	<b>40</b>	<b>54,77</b>	4
fome13	42	115,16	<b>39</b>	<b>107,46</b>	4
fome20	43	127,03	<b>40</b>	<b>120,94</b>	5
fome21	48	275,98	<b>46</b>	<b>268,21</b>	4
tspalko	<b>18</b>	<b>124,53</b>	19	130,10	13
tp-6	<b>21</b>	<b>412,05</b>	23	486,99	17
degme	<b>24</b>	<b>1267,40</b>	25	1308,58	23
karted	16	484,64	<b>15</b>	<b>455,30</b>	13
df001	45	15,26	<b>38</b>	<b>12,97</b>	4
pds-06	29	<b>2,44</b>	29	2,48	10
pds-10	33	12,70	<b>31</b>	<b>12,12</b>	7
pds-20	43	126,81	<b>40</b>	<b>120,09</b>	5
pds-30	<b>45</b>	<b>316,73</b>	46	322,87	5
pds-40	50	961,21	<b>49</b>	<b>933,71</b>	11
pds-50	53	1701,76	<b>50</b>	<b>1649,00</b>	6
pds-60	54	2959,94	<b>51</b>	<b>2798,58</b>	6
pds-70	54	4648,39	<b>52</b>	<b>4475,60</b>	5
pds-80	<b>51</b>	<b>6283,01</b>	52	6388,71	14
pds-90	<b>52</b>	<b>7875,59</b>	53	8001,88	9
pds-100	55	9211,17	<b>54</b>	<b>9055,51</b>	7
Watson-1	<b>80</b>	<b>9,08</b>	83	9,34	19
Watson-2	66	<b>14,18</b>	66	14,28	8

\* O método não resolveu o problema

Note que ATDS obtém redução no número de iterações em 18 problemas teste, além de resolver o problema dano3mip, que não é resolvido por PCx. A maior redução no número de iterações é de 27 iterações no problema lp22, correspondendo a uma taxa de aproximadamente 48%. A maior redução no tempo total é de 172,79 segundos no problema pds-70, correspondendo a uma taxa de redução de aproximadamente 4% do tempo total. No entanto, a maior taxa de redução do tempo total, que é de aproximadamente 44%, ocorreu no problema lp22 com redução de 4,04 segundos. Embora lp22 obtenha uma alta taxa redução de tempo total, este problema é relativamente pequeno. Problemas com maiores tempo de resolução, como pds-70, são os que fornecem as maiores reduções de tempo quando pelo menos uma iteração é reduzida por ATDS. Por outro lado, 10 problemas teste têm o número de iterações aumentado em no máximo duas iterações. Além disso, os resultados mostram que 16 problemas têm o tempo total aumentado a uma taxa aproximada de pelo menos 1%. A maior taxa de aumento do tempo total foi em torno de 18% no problema tp-palko. Nos problemas teste que não tiveram alteração no número de iterações, houve um aumento no tempo total de no máximo 2.5%. Na Figura 3, temos o perfil de desempenho em relação ao número de iterações e na Figura 4 em relação ao tempo total dos resultados da Tabela 2.

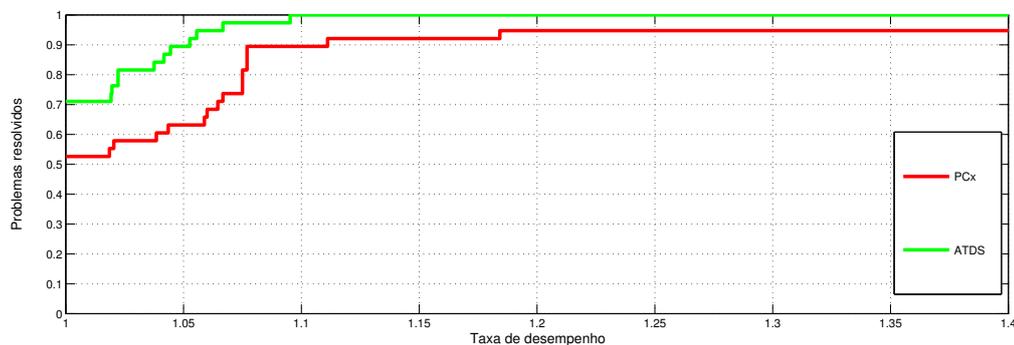


Figura 3 – Número de iterações.

Pela Figura 3, observamos que ATDS é o mais eficiente, resolvendo cerca de 71% dos problemas teste em um menor número de iterações. Além disso, é o mais robusto, visto que ATDS resolve todos os problemas teste. Em relação ao tempo total, podemos observar na Figura 4 que ATDS é também mais eficiente, pois encontrou a solução ótima mais rapidamente em torno de 58% dos problemas teste, enquanto que o PCx foi mais rápido em aproximadamente 42% dos problemas.

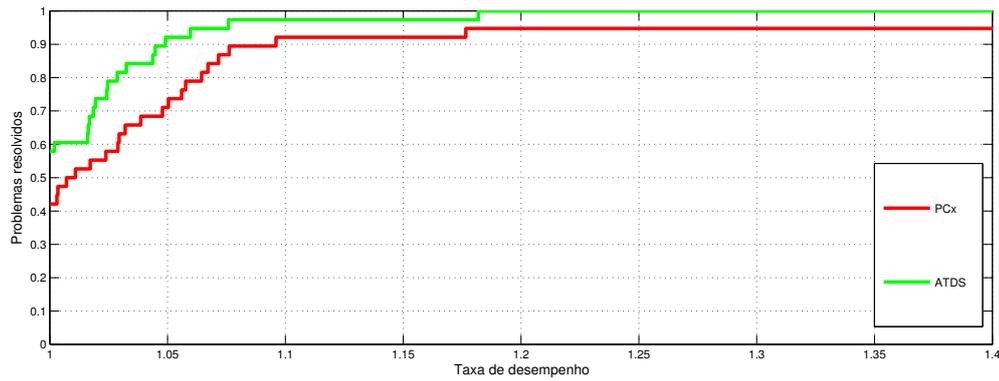


Figura 4 – Tempo total.

Observamos nos experimentos numéricos que na média o número de correções de centralidade é menor para o ATDS em comparação com o mesmo número para o PCx.

Portanto, de acordo com o perfil de desempenho apresentado nas Figuras 3 e 4, podemos afirmar que o ATDS é mais eficiente e robusto que o PCx considerando os problemas testados.

### 5.4.3 Comparação ATDS e o Algoritmo de ajustamento ótimo para $p$ coordenadas combinado com ATDS

Nesta subseção, apresentamos os resultados computacionais obtidos ao combinar o algoritmo de ajustamento ótimo com o ATDS, o qual referenciamos por ATDS-AJ. Na Tabela 4, é feita uma comparação dos resultados obtidos, em relação ao número total de iterações ( $k$ ), o tempo total de resolução ( $s$ ) e o número de iterações continuadas realizado ( $ic$ ). O valor de  $\omega$  em (3.63) para ATDS-AJ é  $\omega_1 = 0,94$ , pois assim ATDS-AJ obtém resultados mais competitivos com ATDS, que utiliza  $\omega_1 = 0,96$ .

Tabela 4 – Comparação entre ATDS e ATDS-AJ.

Problema	ATDS			ATDS-AJ		
	k	t(s)	ic	k	t(s)	ic
baxter	26	24,34	12	26	<b>23,33</b>	6
dano3mip	31	7,70	13	<b>30</b>	<b>7,42</b>	11
dbic1	46	12,77	22	<b>45</b>	<b>12,51</b>	11
dbir1	24	28,34	7	24	<b>28,32</b>	4
dbir2	26	35,91	8	<b>25</b>	<b>34,89</b>	5
ex3sta1	8	131,30	2	8	<b>130,39</b>	2
lp22	<b>29</b>	<b>5,20</b>	7	45	7,57	1
nsct1	<b>21</b>	<b>78,13</b>	7	22	81,54	6
nsct2	<b>26</b>	<b>99,22</b>	10	28	105,88	6
rlfdual	<b>14</b>	<b>6,05</b>	0	15	6,42	0
route	<b>18</b>	<b>3,02</b>	4	19	3,14	4
seymourl	16	22,73	8	16	<b>22,67</b>	8
slptsk	20	<b>34,20</b>	12	20	34,32	9
south31	<b>15</b>	<b>33,00</b>	10	16	34,92	5
rlfprim	9	5195,92	0	9	<b>5184,32</b>	0
fome11	47	31,96	4	<b>43</b>	<b>29,31</b>	3
fome12	40	54,77	4	<b>39</b>	<b>53,66</b>	4
fome13	<b>39</b>	<b>107,46</b>	4	42	114,73	3
fome20	<b>40</b>	<b>120,94</b>	5	41	123,04	3
fome21	<b>46</b>	<b>268,21</b>	4	47	272,81	3
Ts-palko	19	130,10	13	<b>18</b>	<b>123,96</b>	5
tp-6	23	<b>486,99</b>	17	23	487,59	17
degme	25	1308,58	23	<b>23</b>	<b>1209,60</b>	8
karted	15	<b>455,30</b>	13	15	455,31	7
df001	38	<b>12,97</b>	4	38	13,06	3
pds-06	29	2,48	10	29	<b>2,47</b>	8
pds-10	<b>31</b>	<b>12,12</b>	7	32	12,60	5
pds-20	<b>40</b>	<b>120,09</b>	5	41	122,31	3
pds-30	46	322,87	5	<b>44</b>	<b>310,88</b>	4
pds-40	49	933,71	11	49	<b>932,21</b>	5
pds-50	50	1649,00	6	50	<b>1645,82</b>	6
pds-60	51	2798,58	6	<b>50</b>	<b>2746,56</b>	6
pds-70	<b>52</b>	<b>4475,60</b>	5	56	4792,80	4
pds-80	52	6388,71	14	<b>51</b>	<b>6264,31</b>	5
pds-90	53	<b>8001,88</b>	9	53	8008,47	5
pds-100	54	9055,51	7	<b>52</b>	<b>8727,85</b>	4
Watson-1	83	9,34	19	<b>78</b>	<b>9,01</b>	10
Watson-2	66	<b>14,08</b>	8	<b>65</b>	14,21	5

Observe que no ATDS-AJ ocorre redução no número de iterações em 13 problemas teste, comparando com ATDS. A maior redução foi de 4 iterações no problema fome11, correspondendo a uma taxa de aproximadamente 8,5%. Além disso, este é o problema que apresenta a maior taxa de redução de tempo entre os 13 problemas que tiveram o tempo reduzido, com aproximadamente 8,3%. Embora com uma taxa menor de redução do tempo total, os problemas degme, pds-60, pds-80 e pds-100, que apresentam maiores tempo de resolução, são os que obtêm as maiores reduções de tempo, com aproximadamente 99, 52, 124 e 327 segundos, respectivamente. Por outro lado, 12 problemas testes tiveram o número de iterações e o tempo total de resolução aumentados. O maior aumento foi no problema lp22, de aproximadamente 55% no número de iterações e 45% no tempo total. Nos demais problemas testes, que tiveram o número de iterações inalterados, o aumento ou redução no tempo não foi superior a 1%. Note que ATDS-AJ obtém, na maioria dos problemas teste, um número menor de iterações continuadas realizadas comparado com ATDS.

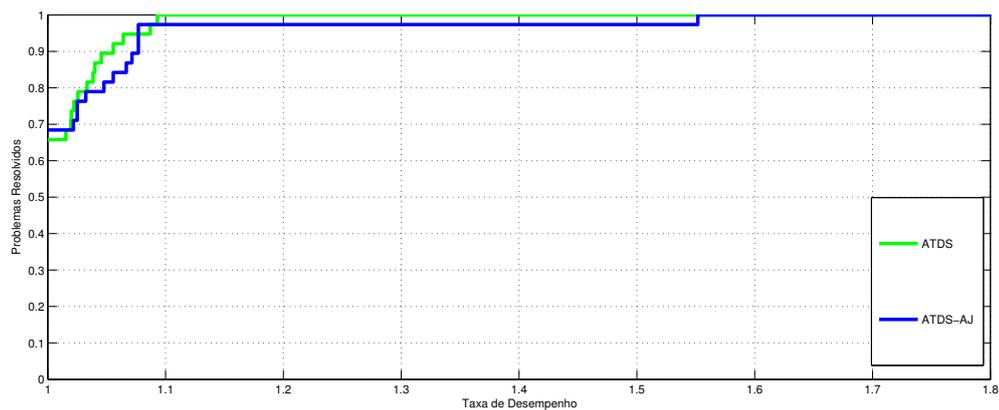


Figura 5 – Número de iterações.

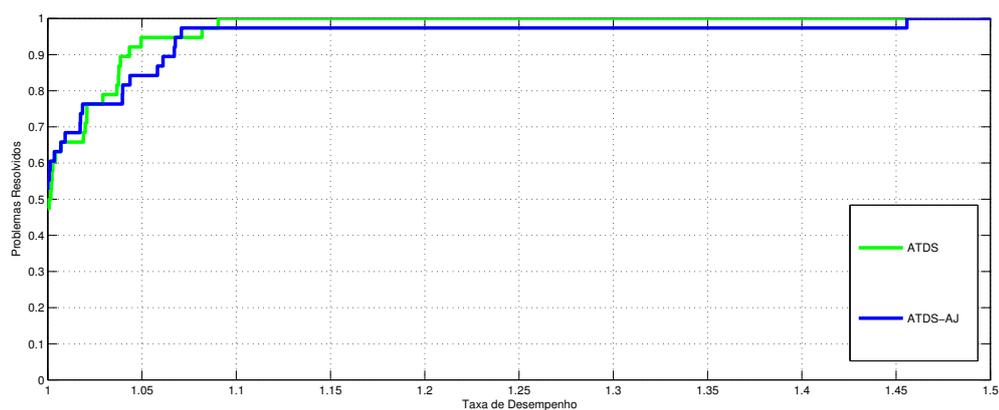


Figura 6 – Tempo total.

Pela Figura 5, observamos que a diferença de desempenho entre os dois métodos é muito pequena. Note que, ATDS-AJ é ligeiramente mais rápido se considerarmos a taxa de desempenho igual a 1. Esta relação se inverte para taxa de desempenho aproximadamente 1,05. Além disso, ATDS é ligeiramente mais robusto para esses problemas testes. A partir da Figura 6 podemos chegar às mesmas conclusões com relação ao tempo total.

# Capítulo 6

## Conclusões

Neste trabalho, aperfeiçoamos o método de pontos interiores preditor corretor com as múltiplas correções de centralidade. O maior esforço computacional, em cada iteração deste método, está na resolução de dois sistemas lineares para determinar a direção preditora corretora e sistemas lineares adicionais dependendo do número de correções de centralidades realizadas, com uma mesma matriz dos coeficientes. Utilizando eliminações de variáveis para resolver estes sistemas lineares, obtemos um sistema de equações normais. Uma abordagem muito utilizada para resolvê-lo é determinando a fatoração de Cholesky da matriz de coeficientes. Ao utilizar as múltiplas correções de centralidade, o método preditor corretor obtém uma direção de busca que possibilita que o método prossiga na vizinhança da trajetória central, que passos maiores sejam dados e que uma melhoria na convergência seja obtida. Dessa forma, muitos pesquisadores consideram essa variação dos métodos de pontos interiores a mais eficiente. No entanto, para muitos problemas a fatoração de Cholesky é muito cara, aumentando custo computacional em cada iteração do método de pontos interiores.

Propomos uma nova abordagem para a melhoria do método preditor corretor, além das múltiplas correções de centralidade, apresentamos a iteração continuada. A iteração continuada para o método de pontos interiores preditor corretor com as múltiplas correções de centralidade é apresentada, com o objetivo de reduzir o número de iterações e o tempo computacional que este método utiliza para resolver problemas de programação linear de grande porte. A iteração continuada é utilizada quando os tamanhos dos passos a serem dados em uma iteração são menores que o passo de Newton. Ela tem como propósito calcular uma direção continuada que, combinada com a direção preditora corretora, resulte em tamanhos de passos maiores para determinadas componentes, não violando a condição de ponto interior. Esta abordagem não garante que o ponto esteja próximo da trajetória central. Dessa forma, é utilizada antes que as múltiplas correções de centralidade sejam realizadas.

A direção continuada é mantida muito próxima da direção preditora corretora,

de forma que possibilite o aumento dos tamanhos dos passos primal e dual, reduzindo as infactibilidades, melhorando a convergência. Duas formas de utilizar as direções continuadas foram propostas, a direção moderada e a direção simples. A direção moderada é determinada satisfazendo aproximadamente o sistema linear que encontra a direção preditora corretora, anulando as componentes da direção que realizaram bloqueio. Além disso, a fatoração de Cholesky já calculada na iteração do método é utilizada para resolver sistemas lineares adicionais envolvidos com a matriz  $ADA^t$ . Por outro lado, a direção simples recebe os valores da direção preditora corretora calculada, anulando as componentes que realizaram bloqueio. Esta direção tem a vantagem de não exigir a resolução de sistemas lineares adicionais. Duas formas de utilizar as direções continuadas foram apresentadas, as formas avançada e atrasada, diferindo entre si pela combinação feita entre direção preditora corretora e a direção continuada. Uma análise dos resíduos foi realizada com as formas avançada e atrasada, impondo critérios para a utilização da direção continuada para melhoria do método de pontos interiores.

Os experimentos computacionais realizados mostraram que a forma atrasada da direção simples para a iteração continuada, ATDS, foi a que apresentou melhores resultados entre as formas de utilização e as direções continuadas apresentadas. Comparando PCx e ATDS, os resultados mostraram que ATDS é mais eficiente e robusto que PCx, pois determina a solução dos problemas testados rapidamente, além de resolver um problema teste não resolvido pelo PCx. Ao introduzir o algoritmo de ajustamento ótimo para  $p$  coordenadas ao ATDS, um método tão eficiente quanto o ATDS é obtido. Embora os problemas teste de maior tempo de resolução não forneçam a maior taxa de redução de tempo, estes apresentam uma economia de tempo total maior ao ter ao menos uma iteração reduzida.

Portanto, a iteração continuada pela forma atrasada da direção simples combinada com o método de pontos interiores preditor corretor com múltiplas correções resultou em um método de pontos interiores eficiente e robusto. Utilizando o algoritmo de ajustamento ótimo para  $p$  coordenadas, obtemos um método tão eficiente quanto o ATDS.

# Referências

- [1] Ilan Adler, Narendra Karmarkar, Mauricio G. C. Resende, and Geraldo Veiga. Data structures and programming techniques for the implementation of Karmarkar's algorithm. *ORSA Journal on Computing*, 1:84–106, 1989.
- [2] Ilan Adler, Mauricio G. C. Resende, Geraldo Veiga, and Narendra Karmarkar. An implementation of Karmarkar's algorithm for linear programming. *Mathematical Programming*, 44:297–335, 1989.
- [3] M.S. Bazaraa, J.J. Jarvis, and H.D. Sherali. *Linear Programming and Network Flows*. Wiley, 1990.
- [4] Lilian Ferreira Berti. Iteração continuada aplicada ao método de pontos interiores. Master's thesis, Campinas, Brasil, Abril, 2012. <http://www.bibliotecadigital.unicamp.br/document/?code=000856687>.
- [5] Dimitri P. Bertsekas. *Nonlinear Programming Second Edition*. Athena Scientific, Belmont, Massachusetts, 1999.
- [6] S. Bocanegra, F. F. Campos, and A. R. L. Oliveira. Using a hybrid preconditioner for solving large-scale linear systems arising from interior point methods. *Computational Optimization and Applications*, 36(1–2):149–164, 2007.
- [7] S. Bocanegra, J. Castro, and A. R. L. Oliveira. Improving an interior-point approach for large block-angular problems by hybrid preconditioners. *European Journal of Operational Research*, 231:263–273, 2013.
- [8] J.R. Bunch and B.N. Parlett. Direct methods for solving symmetric indefinite systems of linear equations. *SIAM Journal on Numerical Analysis*, 8:639–655, dec 1971.
- [9] Rainer E. Burkard, Stefan E. Karisch, and Franz Rendl. QAPLIB - A quadratic assignment problem library. *European Journal of Operational Research*, 55:115–119, 1991.
- [10] 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*, 28:240–248, 1990.

- 
- [11] V. V. Chvátal. *Linear Programming*. W. H. Freeman and Company, New York, USA, 1983.
- [12] M. Colombo and J. Gondzio. Further development of multiple centrality correctors for interior point methods. *Computational Optimization and Applications*, 41:277–305, 2008.
- [13] Joseph Czyzyk, S. Mehrotra, Michael Wagner, and Stephen J. Wright. PCx an interior point code for linear programming. *Optimization Methods & Software*, 11-2(1-4):397–430, 1999.
- [14] John E. Dennis and Robert B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. SIAM, Philadelphia, PA, 1996.
- [15] Elizabeth D. Dolan and Jorge J. Moré. Benchmarking optimization software with performance profiles. *Math. Program., Ser. A*, 91:201–213, 2002.
- [16] C. T. L. S. Ghidini, A. R. L. Oliveira, and J. Silva. Optimal adjustment algorithm for p coordinates and the starting point in interior point methods. *American Journal of Operations Research*, 1:191 – 202, 2011.
- [17] C. T. L. S. Ghidini, A. R. L. Oliveira, J. Silva, and M. I. Velazco. Combining a hybrid preconditioner and a optimal adjustment algorithm to accelerate the convergence of interior point methods. *Linear Algebra and its Applications*, 218:1267–1284, 2012.
- [18] J. P. M. Gonçalves. *A Family of Linear Programming Algorithms Based on the Von Neumann Algorithm*. PhD thesis, Lehigh University, Bethlehem, 2004.
- [19] 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*, 24:461–478, 2009.
- [20] J. Gondzio. Interior point methods 25 years later. *European Journal of Operational Research*, 218:587–601, 2012.
- [21] Jacek Gondzio. Multiple centrality corrections in a primal-dual method for linear programming. *Computational Optimization and Applications*, 6:137–156, 1996.
- [22] C. C. Gonzaga. Path-following methods for linear programming. *SIAM Review*, 34:167–224, 1992.
- [23] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4(4):373–395, 1984.
- [24] L. G. Khachiyan. Polynomial algorithms in linear programming. *Soviet Mathematics Doklady*, 20:191–194, 1979.

- [25] V. Klee and G. J. Minty. *How Good is the Simplex Algorithm?*, pages 159–175. Inequalities III (O. Shisha, ed.), Academic Press, New York, 1972.
- [26] M. Kojima, N. Megiddo, and S. Mizuno. A primal - dual infeasible interior point algorithm for linear programming. *Mathematical Programming*, 61:263–280, 1993.
- [27] 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, pages 29–47, 1989.
- [28] David G. Luenberger. *Linear and Nonlinear Programming*. Addison-Wesley, Reading, 1984.
- [29] 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.
- [30] S. Mehrotra. Implementations of affine scaling methods: Approximate solutions of systems of linear equations using preconditioned conjugate gradient methods. *ORSA Journal on Computing*, 4:103–118, 1992.
- [31] S. Mehrotra. On the implementation of a primal-dual interior point method. *SIAM Journal on Optimization*, 2(4):575–601, 1992.
- [32] Renato D. C. Monteiro, Ilan Adler, and Mauricio 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*, 15:191–214, 1990.
- [33] Esmond Ng and B. P. Peyton. Block sparse Cholesky algorithms on advanced uniprocessors computers. *SIAM Journal on Scientific Stat. Computing*, 14:1034–1056, 1993.
- [34] 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*, pages 928–929, 1988. (in Portuguese).
- [35] Aurelio R. L. Oliveira and D. C. Sorensen. A new class of preconditioners for large-scale linear systems from interior point methods for linear programming. *Linear Algebra and Its Applications*, 394:1–24, 2005.
- [36] Aurelio Ribeiro Leite Oliveira and Christiano Lyra. Implementação de um método de pontos interiores para programação linear. *SBA: Controle & Automação*, 3(2):370–382, 1991.

- 
- [37] J. Silva, R. M. O. Santos, and A. R. L. Oliveira. Uma nova escolha do parâmetro  $p$  para uma família de algoritmos simples em conjunto com métodos de pontos interiores. *Anais do XLVI Simpósio Brasileiro de Pesquisa Operacional*, pages 2728–2737, 2014.
- [38] Jair Silva. *Uma família de algoritmos para programação linear baseada no algoritmo de Von Neumann*. PhD thesis, IMECC – UNICAMP, Campinas SP, Abril, 2009. <http://libdigi.unicamp.br/document/?code=000443131>.
- [39] M. I. Velazco, A. R. L. Oliveira, and F. F. Campos. A note on hybrid preconditions for large scale normal equations arising from interior-point methods. *Optimization Methods and Software*, 25(2):321–332, 2010.
- [40] S. J. Wright. *Primal–Dual Interior–Point Methods*. SIAM Publications, SIAM, Philadelphia, PA, USA, 1997.

# Apêndice A

## Função Lagrangiana

Considere o problema:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & h_i(x) = 0 \quad i = 1, \dots, m. \end{aligned} \tag{A.1}$$

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

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

A função Lagrangiana  $l(x, \lambda) : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  associada ao problema (A.1) é dada por [5]:

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

**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  $\nabla h_1(x^*), \dots, \nabla h_m(x^*)$  são linearmente independentes. Então existe um único vetor  $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)^t$ , tal que:

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

Os elementos  $\lambda^*$  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  $M(x^*) = \{y \mid \nabla h_i(x^*)^t y = 0, i = 1, \dots, m\}$ .

Podemos escrever esta proposição de forma compacta, sendo que:

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

$$\nabla_x l(x^*, \lambda^*) = 0 \quad \text{e} \quad \nabla_\lambda l(x^*, \lambda^*) = 0, \tag{A.2}$$

$$y^t L(x^*) y \geq 0 \quad \forall y \in M(x^*), \tag{A.3}$$

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

A condição (A.2) é conhecida como condição necessária de primeira ordem e a condição (A.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:

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

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

# Apêndice B

## Perfil de Desempenho

O perfil de desempenho é uma ferramenta para avaliar e comparar o desempenho de métodos em otimização, introduzido por Dolan e Moré [15]. Considerando que temos  $n_s$  métodos e  $n_p$  problemas, podemos comparar os métodos em medidas diferentes de desempenho, por exemplo, tempo de processamento, número de iterações, entre outros, que denominamos custo. Considere  $S$  o conjunto de métodos e  $P$  conjunto de problemas, o desempenho do método  $s$  para resolver um problema  $p$  é comparado com o melhor desempenho de qualquer método  $s \in S$  com respeito a este problema, para isso é utilizada a taxa de desempenho dada por:

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

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

A função de distribuição acumulada  $\rho_s$  para a taxa de desempenho  $r_{p,s}$  é definida [15] por:

$$\rho_s : \mathbb{R} \longrightarrow [0, 1]$$

$$\rho_s(\tau) = \frac{1}{n_p} |B|,$$

em que  $|B|$  é a cardinalidade de  $B = \{p \in P | r_{p,s} \leq \tau\}$ . Então,  $\rho_s$  é a probabilidade do método  $s \in S$  resolver um problema a um fator  $\tau$ , da melhor taxa de desempenho.

Assim, para um dado  $\tau$ , o melhor método é o que possui o maior valor de  $\rho_s(\tau)$ . Se estamos interessados no método mais eficiente, necessitamos somente comparar os valores da função de cada método quando  $\tau = 1$ . Considere  $r_{p,s} \in [1, r_m]$ , em 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 dada por:

$$\rho_s^* = \lim_{\tau \rightarrow r_m^-} \rho_s(\tau).$$

Logo, se estamos interessados somente no método mais robusto, necessitamos comparar os valores de  $\rho_s^*$  de todos os métodos e escolher o de maior valor.