

LAC

Universidade Estadual de Campinas  
FACULDADE DE ENGENHARIA ELÉTRICA  
DEPARTAMENTO DE TELEMÁTICA  
Laboratório de Análise Convexa



## Algoritmos de pontos interiores e desigualdades matriciais lineares

Tese apresentada na Faculdade de Engenharia Elétrica da Universidade Estadual de  
Campinas, como parte dos requisitos exigidos para a obtenção do título de  
Mestre em Engenharia Elétrica.

**Maurício Carvalho de Oliveira**  
Engenheiro Eletricista — FEE/UNICAMP

em 21 de março de 1996 perante a banca examinadora

**José C. Geromel** Orientador

**Clóvis C. Gonzaga**  
**Alcir J. Monticelli**

Este exemplar foi aprovado para a defesa final da tese  
defendida por **Maurício Carvalho de Oliveira**

\_\_\_\_\_  
Juregador da Comissão  
Juregador em 21 de março de 1996

\_\_\_\_\_  
Orientador

UNIDADE	BC
N.º CHAMADA:	TUNICAMP
	OL 4. a
	E.
COMBO B0/	27588
ROC.	667196
C	<input type="checkbox"/>
D	<input checked="" type="checkbox"/>
RECO	R\$ 11,00
DATA	03/05/96
N.º CPD	C.M.0008780 4-7

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DA ÁREA DE ENGENHARIA — BAE — UNICAMP

Oliveira, Maurício Carvalho de  
OL4a Algoritmos de pontos interiores e desigualdades  
matriciais lineares / Maurício Carvalho de Oliveira.—  
Campinas, SP: [s.n.], 1996.

Orientador: José C. Geromel.

Dissertação (mestrado) - Universidade Estadual de  
Campinas, Faculdade de Engenharia Elétrica.

1. Programação (Matemática). 2. Métodos numéricos.  
3. Otimização matemática. 4. Algoritmos. I. Geromel,  
José C. II. Universidade Estadual de Campinas. Faculdade  
de Engenharia Elétrica. III. Título.

Tese : **Algoritmos de pontos interiores e  
desigualdades matriciais lineares**

Autor : **Maurício Carvalho de Oliveira**

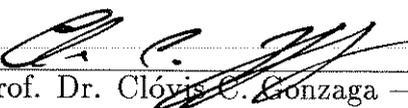
Orientador : **Prof. Dr. José C. Geromel**

Aprovada em 21 de março de 1996 pela banca examinadora



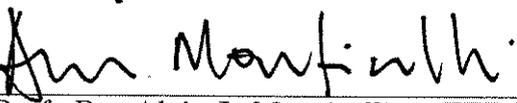
---

Prof. Dr. José C. Geromel (Presidente)



---

Prof. Dr. Clóvis C. Gonzaga — IM/UFSC



---

Prof. Dr. Alcir J. Monticelli — FEE/UNICAMP



A quem torna muito mais prazeroso  
(até mesmo) escrever este trabalho,  
Dani.

---



# Agradecimentos

Embora os créditos desta obra sejam atribuídos a mim, diversas pessoas contribuíram de maneira brilhante e intensa em sua realização. Antes de mais nada, espero ter podido reproduzir neste texto pequena parte da fascinante convivência com o sempre obstinado Prof. J. C. Geromel. Para antigos alunos e amigos não deve ser difícil notar sua presença ao longo de toda esta dissertação. Pelas sempre acaloradas e produtivas discussões, meu muito obrigado. Agradecimentos especiais cabem ao colega P. Gapski, pelo constante apoio e, em especial, pela leitura atenta dos primeiros capítulos desta obra, e ao Prof. P. Peres, pela eterna atenção. Abraço especial para A. de Oliveira, meu pai, pela paternal correção ortográfica e gramatical, além da inevitável tradução matematiqûês — português (Todos os eventuais erros foram adicionados por mim após a revisão!).

Agradecimento a todos os companheiros e colegas do Departamento de Telemática que conviveram comigo estes dois últimos anos.

Finalmente, agradeço a meus pais o indispensável apoio às aventuras de seu filho.



# Resumo

Esta dissertação é dedicada ao estudo dos mecanismos dos algoritmos de pontos interiores aplicados à resolução de problemas lineares sujeitos a restrições dadas na forma de desigualdades matriciais lineares. Abordam-se tanto aspectos teóricos quanto práticos. De aspecto teórico, encontram-se presentes análises de convergência e complexidade para diversos algoritmos *seguidores de trajetória*, *primais-duais* e *projetivos*, aliados às análises de alguns procedimentos críticos, como a resolução dos problemas de mínimos quadrados e a determinação do passo ótimo, aspectos eminentemente práticos. A título de ilustração, apresenta-se uma série de exemplos de problemas comumente encontrados em programação matemática e, em especial, problemas da área de controle ótimo formulados como LMI.

# Abstract

The subject of this thesis is the study of the interior point machinery applied to linear problems constrained by linear matrix inequalities (LMI). Both theoretical and practical issues are addressed. Of theoretical flavor, convergence and complexity of several *path following*, *primal-dual* and *projective* algorithms are analyzed; the analysis of some critical procedures, as solving mean-square problems and calculating the optimal step length, stand for the practical issues. For the sake of illustration, many mathematical programs and problems from optimal control theory are formulated as LMI.



# Conteúdo

<b>Agradecimentos</b>	<b>v</b>
<b>Resumo</b>	<b>vii</b>
<b>Abstract</b>	<b>vii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Preliminares . . . . .	1
1.2 Métodos de Barreira . . . . .	2
1.3 Estrutura do trabalho . . . . .	3
<b>2 Fundamentos teóricos</b>	<b>5</b>
2.1 O problema linear sujeito a restrições do tipo LMI . . . . .	5
2.2 O problema dual . . . . .	6
2.3 Uma barreira para LMI . . . . .	6
2.4 Formulação cônica . . . . .	9
2.5 Projeções . . . . .	12
2.6 Transformação simétrica . . . . .	13
2.7 O elipsóide de Dikin . . . . .	14
2.8 Uma barreira para o cone positivo . . . . .	16
2.9 Otimização restrita a variedades lineares . . . . .	17
2.10 O centro analítico . . . . .	19
<b>3 Algoritmos: aspectos conceituais</b>	<b>23</b>
3.1 Calculando o centro analítico . . . . .	23
3.1.1 O algoritmo básico . . . . .	23
3.1.2 A função apenada . . . . .	24
3.1.3 Medida da centralização . . . . .	25
3.1.4 O passo ótimo . . . . .	30
3.2 Algoritmos seguidores de trajetória . . . . .	32
3.2.1 Parametrização primal-dual . . . . .	32
3.2.2 Parametrização primal . . . . .	32
3.2.3 Parametrização dual . . . . .	34
3.2.4 O algoritmo seguidor de trajetória . . . . .	35
3.3 Algoritmos primais-duais . . . . .	40
3.3.1 A função potencial e a formulação primal-dual . . . . .	45

---

3.4	Algoritmo projetivo . . . . .	49
3.4.1	O problema de factibilidade . . . . .	50
3.4.2	O problema de otimização . . . . .	53
3.5	O problema de factibilidade . . . . .	58
<b>4</b>	<b>Algoritmos: aspectos práticos</b>	<b>61</b>
4.1	O cálculo das projeções . . . . .	61
4.2	O cálculo do passo ótimo . . . . .	64
4.3	O cálculo de $f_k^*$ . . . . .	69
4.4	Alguns experimentos computacionais . . . . .	69
<b>5</b>	<b>Algumas aplicações</b>	<b>73</b>
5.1	Abrangência da formulação LMI . . . . .	73
5.2	LMI e problemas de controle . . . . .	76
5.3	Notas finais e perspectivas . . . . .	82
	<b>Referências bibliográficas</b>	<b>85</b>

# Capítulo 1

## Introdução

### 1.1 Preliminares

A área de estudos que hoje se denomina *teoria de controle* ou *teoria de sistemas* vem se utilizando das mais diversas ferramentas matemáticas no estudo do comportamento de sistemas dinâmicos. O problema básico — determinar um controlador realizável em termos físicos e tecnológicos que estabilize um sistema dinâmico descrito por equações diferenciais ou a diferenças — vem se mostrando campo fértil para a aplicação de resultados que remontam ao cálculo operacional e se estendem até à análise funcional. Nos últimos anos, grande parte das pesquisas nesta área vem fazendo uso, de modo intenso e com sucesso, de ferramentas de programação matemática. Nesse contexto, a formulação dos mais diversos problemas de controle como problemas de otimização tem provido respostas a problemas que, há pouco tempo, eram considerados intratáveis.

Este trabalho se propõe estudar os algoritmos e mecanismos de pontos interiores aplicados à resolução de uma classe particular de problemas, a ser introduzida no segundo capítulo, que abrange uma fatia considerável dos problemas de controle que apresentam formulação convexa como, por exemplo, o problema de síntese de controladores em otimização  $\mathcal{H}_2$ , abordado brevemente na seção 5.2 do capítulo 5, dentre outros. As soluções para esses problemas vêm sendo obtidas, já há algum tempo e com certo sucesso<sup>1</sup>, mediante o emprego do algoritmo de planos de corte e do algoritmo do elipsóide. No entanto, o amplo espectro desses algoritmos, aplicáveis a todo e qualquer problema de otimização convexa, faz com que o tempo requerido para obtenção de uma solução torne-se muito grande à medida que aumentam as dimensões das variáveis e do conjunto de restrições do problema, o que limita seu uso a problemas de pequeno e médio porte. Por sua vez, a estrutura muito mais rígida e complexa dos algoritmos de pontos interiores permite obter soluções em tempo polinomial com um custo computacional bastante moderado.

Indo de certa maneira contra a linha atual de desenvolvimento desses métodos, em geral baseados diretamente no conceito de *autoconcordância* introduzido em Nesterov & Nemirovskii [14], este trabalho pretende discutir diversos algoritmos de pontos interiores como uma generalização, quase que imediata, dos resultados da bem-sucedida experiência em programação linear. Surpreendentemente, essa abordagem revela-se bastante frutífera e fornece interpretações mais sólidas, do ponto de vista do autor desta dissertação, dos

---

<sup>1</sup>Veja, por exemplo, Oliveira & Geromel [17].

mecanismos de convergências dos algoritmos em estudo.

A seguir, ainda nesta introdução, faz-se uma breve revisão dos mecanismos básicos dos algoritmos de pontos interiores, seguida de uma apresentação sucinta dos capítulos subsequentes.

## 1.2 Métodos de Barreira

Todos os atuais algoritmos de pontos interiores baseiam-se nos *métodos de barreira*, desenvolvidos na década de sessenta e extensamente discutidos no conhecido trabalho de Fiacco & McCormick [6], e que fazem parte, hoje em dia, de diversos livros tradicionais de otimização não-linear como, por exemplo, Luemberger [13] e Bazaraa *et alii* [2]. A estratégia utilizada pelos métodos de barreira para resolver o problema de otimização convexo

$$\begin{aligned} &\text{minimizar } f(x) \\ &\text{s.a. } g_i(x) \leq 0, \quad i = 1, \dots, m \end{aligned} \quad (1.1)$$

em que  $x \in \mathbb{R}^n$ ,  $f(\cdot)$  é diferenciável e tanto  $f(\cdot)$  quanto  $g_i(\cdot)$  são funções convexas, consiste em transformá-lo em um problema de otimização irrestrito mediante a eliminação do conjunto de restrições de desigualdade e a incorporação de uma *função de barreira* à função objetivo. Função de barreira é qualquer função  $B(x)$  que seja contínua, diferenciável e convexa no interior do conjunto

$$\Phi = \{x \mid g_i(x) \leq 0, \quad i = 1, \dots, m\} \quad (1.2)$$

e que tenda ao infinito para uma seqüência qualquer de pontos pertencentes a  $\Phi$  que tenda à sua fronteira. Assim, para um único conjunto convexo  $\Phi$ , devem existir diversas funções que possuam as características descritas acima e que, portanto, possam ser chamadas funções de barreira. Com qualquer uma dessas funções de barreira constrói-se o seguinte problema irrestrito

$$\min_x \{\beta f(x) + B(x)\} \quad (1.3)$$

em que  $\beta$  é uma constante positiva, em relação ao qual percebe-se que um valor suficientemente “grande” dessa constante  $\beta$  é capaz de localizar o ponto solução do problema (1.3) arbitrariamente próximo à única solução de (1.1). Portanto, a determinação de uma seqüência de soluções do problema convexo irrestrito (1.3) pertencentes a  $\Phi$  gerada por uma seqüência crescente de valores de  $\beta$  que tenda ao infinito, converge para a solução do problema original (1.1). A existência de eficientes métodos de resolução de problemas convexas irrestritos nos moldes de (1.3) garante o sucesso desse esquema. Em termos práticos, os custos computacionais associados às funções de barreira e o conhecido mau condicionamento numérico próximo à fronteira do conjunto  $\Phi$  constituem os maiores obstáculos. No que diz respeito ao mau condicionamento numérico, a singularidade introduzida pelas funções de barreira em pontos pertencentes à fronteira do conjunto  $\Phi$  costuma ser a principal fonte de problemas; paradoxalmente, essa singularidade assegura que as soluções dos problemas (1.3) mantenham-se sempre no interior do conjunto  $\Phi$  (o que, em última análise, provê a “eliminação” do conjunto de restrições no processo de otimização) enquanto nos casos de interesse<sup>2</sup> a solução do problema original (1.1) se encontra exatamente sobre a fronteira

<sup>2</sup>Caso contrário o conjunto de restrições não se encontra *ativo* no ponto solução e, portanto, pode ser ignorado.

de  $\Phi$ . A escolha do método de otimização irrestrita pode ajudar a diminuir sensivelmente esse problema e a experiência tem consagrado o método de Newton *modificado* como o mais adequado, pois apresenta boa taxa de convergência associada a uma certa “imunidade” a maus condicionamentos da matriz Hessiana.

Apesar do grande impacto que essas idéias representaram na época de seu desenvolvimento, problemas de ordem computacional, principalmente, fizeram com que esse tipo de abordagem caísse em desuso na década de setenta. Em meados da década de oitenta, o surgimento do método de Karmakar [12] como alternativa ao método simplex na resolução de problemas de programação linear, redespertou o interesse por essa classe de métodos de otimização. O principal atrativo dos algoritmos desenvolvidos a partir de então consiste na possibilidade de obtenção da solução de programas lineares em tempo polinomial e de maneira computacionalmente eficiente. Esta dissertação trata da extensão dessas propriedades a uma classe maior de problemas convexos.

### 1.3 Estrutura do trabalho

Incluindo-se esta introdução, esta dissertação é constituída de cinco capítulos, cujos conteúdos são descritos brevemente nas próximas linhas.

No segundo capítulo, define-se o *problema linear sujeito a restrições na forma de desigualdades matriciais lineares*, alvo de todos os algoritmos descritos neste trabalho. A já estabelecida sigla proveniente da denominação em inglês LMI (*Linear matrix inequalities*) é utilizada ao longo de todo o texto para denominar esse particularíssimo conjunto de restrições. No restante do capítulo, desenvolvem-se as ferramentas teóricas necessárias à construção de diversos algoritmos. Apesar de bastante comuns no contexto de programação linear, os conceitos utilizados na apresentação desses fundamentos não são encontrados explicitamente nos trabalhos que envolvem LMI, nos quais, em geral, recorre-se diretamente ao conceito de autoconcordância desenvolvido por Nesterov & Nemirovskii em [14].

Inicia-se o terceiro capítulo com a apresentação de alguns resultados de fundamental importância nas análises de convergência dos mecanismos de pontos interiores. Em seguida, descrevem-se detalhadamente três importantes classes de algoritmos de pontos interiores: os algoritmos seguidores de trajetória, os algoritmos primais-duais e os algoritmos projetivos. Por fim, discute-se brevemente o problema de obtenção de um ponto inicial estritamente factível.

Discutem-se, no quarto capítulo, detalhes referentes à implementação computacional de alguns procedimentos utilizados pelos algoritmos introduzidos no terceiro capítulo. No final do capítulo, resultados de um pequeno experimento computacional ilustram o desempenho desses procedimentos.

O quinto e último capítulo é dedicado à apresentação de exemplos de problemas de otimização no formato LMI. Apresentam-se a conversão de conhecidos problemas convexos a LMI e um exemplo de aplicação proveniente da área de controle.

Em todo o trabalho, sempre que possível, preferiu-se utilizar termos técnicos em português, apesar de as referências, em sua totalidade, encontrarem-se em língua inglesa. Porém, toda vez que um novo termo em português é introduzido, apresenta-se o correspondente termo em língua inglesa em itálico. Alguns poucos termos foram mantidos em inglês e, igualmente, aparecem sempre grafados em itálico. A notação procura ser a mais

coerente possível; associam-se as letras maiúsculas a matrizes e as letras minúsculas a vetores e escalares.

## Capítulo 2

# Fundamentos teóricos

### 2.1 O problema linear sujeito a restrições do tipo LMI

O problema básico a ser estudado nesse trabalho apresenta-se sempre no seguinte formato

$$\begin{aligned} & \text{minimizar} && c'x \\ & \text{s.a} && x \in \Phi \end{aligned} \tag{2.1}$$

em que  $x \in \mathbb{R}^n$  e o conjunto  $\Phi$  é definido por

$$\Phi := \{x \mid v'F(x)v \geq 0, \forall v \in \mathbb{R}^p\} \tag{2.2}$$

Fazendo-se uso do conceito de positividade para matrizes reais e simétricas, esse conjunto pode ser também descrito por

$$\Phi := \{x \mid F(x) \geq 0\} \tag{2.3}$$

A matriz  $F(x)$  consiste em uma LMI (do inglês, *Linear matrix inequalities*), isto é, uma combinação afim das matrizes simétricas  $F_{i=0,\dots,n} \in \mathbb{R}^{p \times p}$  na forma

$$F(x) := F_0 + \sum_{i=1}^n F_i x_i \tag{2.4}$$

Essa classe de problemas de otimização, conhecida pelo nome de programação definida positiva (do inglês, *PDP — Positive definite programming*), engloba grande número de conhecidos programas convexos como, por exemplo, programas lineares e programas quadráticos, dentre outros. Isso tornar-se-á claro no capítulo 5, no qual serão apresentados vários exemplos, juntamente com as ferramentas de conversão ao formato LMI.

A propriedade básica que permite o desenvolvimento de algoritmos eficientes para determinar a solução de (2.1) é, sem dúvida, a convexidade do conjunto  $\Phi$ . A garantia de que toda solução mínima local é de fato uma solução global para o problema em questão, aliada às propriedades que a convexidade impõe à matriz Hessiana de funções de barreira de domínios convexos, são as pedras fundamentais de todo algoritmo de pontos interiores eficiente. A convexidade do conjunto  $\Phi$  pode ser verificada através da seguinte identidade

$$v'F(\alpha x + (1 - \alpha)y)v = \alpha v'F(x)v + (1 - \alpha)v'F(y)v, \forall v \in \mathbb{R}^p \tag{2.5}$$

que assegura que  $F(x) \geq 0$  define um conjunto convexo. Finalmente, quanto à caracterização dos pontos pertencentes a  $\Phi$ , diz-se que  $x$  é estritamente factível ou pertencente ao *interior* de  $\Phi$  se  $x$  pertence ao subconjunto de  $\Phi$  dado por

$$\Phi^0 := \{x \mid F(x) > 0\} = \text{int}(\Phi) \quad (2.6)$$

Caso  $x$  pertença a  $\Phi$  mas não pertença a  $\Phi^0$ , diz-se que  $x$  é um ponto na fronteira de  $\Phi$  ou um ponto de fechamento. Como antecipado no capítulo 1, a busca de soluções por algoritmos de pontos interiores dá-se pela determinação de uma seqüência de pontos estritamente factíveis, isto é, obrigatoriamente contida em  $\Phi^0$ .

## 2.2 O problema dual

Pode-se associar ao problema (2.1) a seguinte *função lagrangeana*

$$L(x, Z) := c'x - \text{tr}(ZF(x)) = \sum_{i=1}^n [c_i - \text{tr}(ZF_i)]x_i - \text{tr}(ZF_0) \quad (2.7)$$

donde  $\max_{\{Z \geq 0\}} \min_x L(x, Z)$  define o problema

$$\begin{array}{ll} \text{minimizar} & \text{tr}(ZF_0) \\ \text{s.a} & \left\{ \begin{array}{l} \text{tr}(ZF_i) = c_i, \text{ para } i = 1, \dots, n \\ Z \geq 0 \end{array} \right. \end{array} \quad (2.8)$$

chamado problema *dual*, em oposição ao problema *primal* (2.1). As relações entre esses dois problemas serão discutidas mais adiante, na seção 2.4.

## 2.3 Uma barreira para LMI

O primeiro passo para se tentar resolver (2.1), seguindo a linha dos métodos de barreira, consiste em escolher uma função contínua, convexa e diferenciável, que aja como barreira para pontos no interior do conjunto<sup>1</sup> convexo  $\Phi$ . Uma função que apresenta tais características é dada por

$$B(x) := \begin{cases} -\log \det F(x), & \text{para } F(x) > 0 \\ \infty, & \text{em qualquer outra condição} \end{cases} \quad (2.9)$$

Note-se que  $B(x)$  é finita apenas em pontos estritamente factíveis, isto é, em pontos que pertencem ao conjunto  $\Phi^0$ , tendendo ao infinito para qualquer ponto na fronteira de  $\Phi$ . Várias propriedades desta função, a serem exploradas posteriormente, são apresentadas a seguir.

**Lema 2.1** *Dado  $B(x)$ , uma função de barreira para o domínio convexo  $\Phi$ , definida em (2.9), os seguintes fatos são verdadeiros*

<sup>1</sup>Note-se que esse conjunto é definido por uma função não diferenciável, a não ser para escolhas particulares das matrizes  $F_i$ .

i) O valor da barreira para pontos pertencentes à reta definida por  $(x + \gamma h)$ ,  $x \in \Phi^0$ ,  $h \in \mathbb{R}^n$  e  $\gamma \in \mathbb{R}$ , é dado por

$$B(x + \gamma h) = B(x) - \sum_{i=1}^p \log [1 + \gamma \lambda_i(\Psi(h))], \quad (2.10)$$

em que  $\lambda_i(\cdot)$  denota o  $i$ ésimo autovalor de  $(\cdot)$  e  $\Psi(h)$  é a matriz definida por

$$\Psi(h) := F(x)^{-\frac{1}{2}} [F(h) - F_0] F(x)^{-\frac{1}{2}} \quad (2.11)$$

ii) As derivadas direcionais primeira, segunda e terceira de  $B(x)$ , tomadas no sentido do vetor  $h$ , são dadas por

$$DB(x, h) = -\operatorname{tr}(\Psi(h)) \quad (2.12)$$

$$D^2B(x, h) = \operatorname{tr}(\Psi^2(h)) \quad (2.13)$$

$$D^3B(x, h) = -2 \operatorname{tr}(\Psi^3(h)) \quad (2.14)$$

e obedecem às seguintes desigualdades

$$|DB(x, h)| \leq (D^2B(x, h))^{1/2} \quad (2.15)$$

$$|D^3B(x, h)| \leq 2 (D^2B(x, h))^{3/2} \quad (2.16)$$

iii) Os componentes do vetor gradiente e da matriz Hessiana de  $B(x)$  são dados por

$$\nabla B(x)_i = -\operatorname{tr}(F(x)^{-\frac{1}{2}} F_i F(x)^{-\frac{1}{2}}) \quad (2.17)$$

$$\nabla^2 B(x)_{i,j} = \operatorname{tr}(F(x)^{-\frac{1}{2}} F_i F(x)^{-\frac{1}{2}} F(x)^{-\frac{1}{2}} F_j F(x)^{-\frac{1}{2}}) \quad (2.18)$$

iv) Se  $F(x) \geq 0$  representa a interseção de  $k$  conjuntos definidos por LMI, ou seja,

$$\Phi = \left\{ x \mid x \in \bigcap_{i=1}^k \{F_i(x) \geq 0\} \right\} \quad (2.19)$$

$B(x)$  é dada por

$$B(x) = \sum_{i=1}^k B_i(x) \quad (2.20)$$

em que  $B_i(x)$ ,  $i = 1, \dots, k$ , é a função de barreira  $B(\cdot)$  para o conjunto definido por  $F_i(x) \geq 0$ .

**Prova:** Começando pelo ponto i), a prova deste lema será feita por pontos.

i) O valor de  $B(x + \gamma h)$ , para  $F(x) > 0$ , pode ser desenvolvido em

$$\begin{aligned} B(x + \gamma h) &= -\log \det F(x + \gamma h) \\ &= -\log \det \{F(x) + \gamma [F(h) - F_0]\} \\ &= -\log \det F(x) - \log \det [\mathbf{I} + \gamma \Psi(h)] \\ &= B(x) - \sum_{i=1}^p \log [1 + \gamma \lambda_i(\Psi(h))] \end{aligned}$$

ii) Derivando-se esta última expressão em relação a  $\gamma$  obtém-se

$$\frac{dB(x + \gamma h)}{d\gamma} = - \sum_{i=1}^p \frac{\lambda_i(\Psi(h))}{1 + \gamma \lambda_i(\Psi(h))}$$

donde o valor da derivada direcional primeira pode ser obtido por meio de

$$\left. \frac{dB(x + \gamma h)}{d\gamma} \right|_{\gamma=0} = - \sum_{i=1}^p \lambda_i(\Psi(h)) = - \text{tr}(\Psi(h))$$

Derivando-se, mais uma vez, em relação a  $\gamma$

$$\frac{d^2 B(x + \gamma h)}{d\gamma^2} = \sum_{i=1}^p \frac{\lambda_i^2(\Psi(h))}{[1 + \gamma \lambda_i(\Psi(h))]^2}$$

donde, como anteriormente, é obtido o valor da derivada direcional segunda

$$\left. \frac{d^2 B(x + \gamma h)}{d\gamma^2} \right|_{\gamma=0} = \sum_{i=1}^p \lambda_i^2(\Psi(h)) = \text{tr}(\Psi(h)\Psi(h))$$

Repetindo-se esse procedimento para a derivada direcional terceira

$$\frac{d^3 B(x + \gamma h)}{d\gamma^3} = -2 \sum_{i=1}^p \frac{\lambda_i^3(\Psi(h))}{[1 + \gamma \lambda_i(\Psi(h))]^3}$$

tem-se

$$\left. \frac{d^3 B(x + \gamma h)}{d\gamma^3} \right|_{\gamma=0} = -2 \sum_{i=1}^p \lambda_i^3(\Psi(h)) = -2 \text{tr}(\Psi(h)\Psi(h)\Psi(h))$$

As expressões acima atribuem à função de barreira  $B(x)$  o título de *estritamente convexa* nos casos de interesse. Para que a derivada segunda de  $B(x)$  seja igual a zero, todos os autovalores de  $\Psi(h)$  devem valer zero, o que somente ocorre quando  $F(h) - F_0 = \mathbf{0}$ , ou seja, se  $F_0$  puder ser escrito como uma combinação linear das matrizes  $F_{i=1, \dots, n}$ . Nesse caso, porém, a solução do problema (2.1) ou é ilimitada, ou se reduz à origem  $x = \mathbf{0}$ .

Quanto às desigualdades das derivadas, a primeira delas pode ser justificada mediante o emprego do seguinte argumento

$$\begin{aligned} |DB(x, h)| &= \left| \sum_{i=1}^p \lambda_i(\Psi(h)) \right| \\ &\leq \sum_{i=1}^p |\lambda_i(\Psi(h))| \\ &\leq \left( \sum_{i=1}^p \lambda_i^2(\Psi(h)) \right)^{1/2} \\ &= (D^2 B(x, h))^{1/2} \end{aligned}$$

O mesmo raciocínio aplicado à segunda desigualdade implica

$$\begin{aligned} |D^3B(x, h)| &= 2|\operatorname{tr}(\Psi^3(h))| \\ &\leq 2\sum_{i=1}^p |\lambda_i(\Psi(h))|^3 \\ &\leq 2\left(\sum_{i=1}^p \lambda_i^2(\Psi(h))\right)^{3/2} \\ &= 2(D^2B(x, h))^{3/2} \end{aligned}$$

Essas desigualdades garantem que  $B(x)$  seja uma barreira *autoconcordante* (do inglês, *self-concordant*). Todas as propriedades de convergência dos métodos abordados neste trabalho podem ser entendidas como decorrentes deste conceito. Essa abordagem, prioritariamente algébrica e por demais generalizada, encobre, no entanto, grande parte do denso conteúdo geométrico associado a problemas formulados nos moldes de (2.1) e (2.8). Porém, vale a pena ressaltar que esse conceito generaliza os mecanismos dos algoritmos de pontos interiores para domínios convexos genéricos. Uma abordagem completa e profunda desse assunto pode ser encontrada em Nesterov & Nemirovskii [14].

iii) As expressões para o vetor gradiente e para a matriz Hessiana de  $B(x)$  podem ser obtidas como casos particulares das expressões anteriores, desde que se defina uma direção  $h_i$  tal que a reta  $(x + \gamma h_i)$  represente apenas variações da  $i$ ésima componente de  $x$ . Nesse caso  $F(x + \gamma h_i) = (F(x) + \gamma F_i)$  e  $\Psi(h_i)$  se reduz a  $F(x)^{-\frac{1}{2}} F_i F(x)^{-\frac{1}{2}}$ , o que comprova diretamente a expressão obtida para o vetor gradiente. No caso da matriz Hessiana, manipulações como as efetuadas na prova do ponto anterior, tomando-se o cuidado de substituir a segunda derivada na direção  $h_i$  por uma derivada parcial em relação a uma segunda direção  $h_j$ , levam ao resultado apresentado. As expressões são praticamente idênticas e serão, portanto, omitidas.

iv) A interseção de  $k$  conjuntos definidos por LMI pode ser representada por

$$F(x) = \begin{bmatrix} F_1(x) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & F_k(x) \end{bmatrix} \geq 0$$

donde o restante da prova é trivial.

□

## 2.4 Formulação cônica

Os problemas (2.1) e (2.8) podem ser reescritos de maneira a ressaltar aspectos geométricos interessantes. A semelhança com o problema de programação linear torna-se aparente, o que permite vislumbrar novas perspectivas para a solução do problema original. Algumas definições preliminares se fazem necessárias. No espaço vetorial cujos elementos são todas

as matrizes simétricas pertencentes a  $\mathfrak{R}^{p \times p}$ , o produto interno de duas matrizes  $X$  e  $Y$  é dado por

$$\langle X, Y \rangle = \langle Y, X \rangle := \text{tr}(XY) \quad (2.21)$$

e a norma induzida por essa definição de produto interno, que, nesse caso, em que as matrizes são simétricas, coincide com a chamada norma de Frobenius, é calculada, para uma dada matriz  $X$ , através da seguinte expressão

$$\|X\| := \sqrt{\text{tr}(XX)} = \sqrt{\text{tr}(X^2)} \quad (2.22)$$

Definindo-se as matrizes  $D := F_0$  e  $C$  tal que

$$\langle C, X \rangle = c'x, \quad (2.23)$$

para toda matriz pertencente à variedade linear  $X = F_0 + \sum_{i=1}^n F_i x_i$ , isto é,

$$\sum_{i=1}^n (\langle C, F_i \rangle - c_i) x_i + \langle C, F_0 \rangle = 0 \quad (2.24)$$

relação que é sempre satisfeita caso<sup>2</sup>

$$\langle C, F_i \rangle = c_i \text{ e } \langle C, D \rangle = 0, \quad (2.25)$$

o problema primal (2.1) pode ser reescrito

$$\begin{aligned} &\text{minimizar } \langle C, X \rangle \\ &\text{s.a. } X \in \{\mathcal{K} \cap (\mathcal{L} + D)\} \end{aligned} \quad (2.26)$$

em que  $\mathcal{K}$  denota o cone das matrizes definidas positivas, isto é,  $\mathcal{K} := \{X \mid X \geq 0\}$ ,  $\mathcal{L}$  denota o subespaço linear gerado por todas as possíveis combinações lineares das matrizes  $F_{i=1, \dots, n}$ ,  $\mathcal{L} := \{X \mid X = \sum_{i=1}^n F_i \alpha_i, \alpha_i \in \mathfrak{R}\}$ , e  $(\mathcal{L} + D)$  representa a variedade linear gerada pela adição da matriz  $D$  às matrizes pertencentes ao subespaço  $\mathcal{L}$ . A transcrição de (2.1) para (2.26) é imediata, observadas as definições anteriores, e a razão do nome cônica torna-se evidente: o conjunto de restrições passa a ser dado pela interseção de uma variedade linear com o cone das matrizes definidas positivas. Este fato é semelhante ao que ocorre nos programas lineares, onde o conjunto poliedral de restrições pode ser sempre escrito na forma  $\{\{x \geq 0\} \cap \{Ax = b\}\}$ , isto é, a interseção da variedade linear  $\{Ax = b\}$  com o ortante ou cone dos vetores positivos  $\{x \geq 0\}$ . Porém, o aspecto mais surpreendente desta formulação consiste na caracterização do problema dual (2.8). Observando-se que as restrições  $\text{tr}(ZF_i) = c_i$  implicam

$$\left\langle Z, \sum_{i=1}^n F_i x_i \right\rangle = \sum_{i=1}^n c_i x_i \quad (2.27)$$

que, por sua vez, para qualquer matriz  $W \in \mathcal{L}$ , pode ser escrita

$$\langle Z - C, W \rangle = 0 \quad (2.28)$$

---

<sup>2</sup>Se o número de componentes independentes da matriz  $C$ ,  $(p(p+1)/2)$ , for maior que o número original de variáveis mais um,  $(n+1)$ , uma matriz  $C$  que atenda a essas características sempre poderá ser obtida.

o problema dual pode ser visto como

$$\begin{aligned} & \text{minimizar} && \langle D, Z \rangle \\ & \text{s.a} && Z \in \{ \mathcal{K} \cap (\mathcal{L}^\perp + C) \} \end{aligned} \quad (2.29)$$

em que  $(\mathcal{L}^\perp + C)$  denota a variedade linear gerada pela adição do vetor  $C$  ao subespaço constituído por todos os vetores ortogonais a  $\mathcal{L}$ . Dessa forma, fica claro que o problema dual é também um problema linear sujeito a restrições do tipo LMI que respeita uma espécie de simetria em relação ao problema primal semelhante à encontrada em programas lineares: as variáveis primal e dual pertencem a variedades lineares geradas por subespaços ortogonais, cujo vetor de deslocamento em relação à origem para o problema primal é o vetor gradiente do objetivo linear do problema dual, e vice-versa. Outra semelhança: o produto escalar de matrizes primal e dual factíveis é dado por

$$\eta := \langle X, Z \rangle = \langle X, C \rangle + \langle Z, D \rangle \quad (2.30)$$

fato que pode ser verificado tendo-se em mente que  $\langle C, D \rangle = 0$ . Note-se que  $\eta$  é o chamado *gap* de dualidade do par de problemas primal (2.26) e dual (2.29). No contexto de programação linear, essa relação é conhecida pelo nome de folga complementar (do inglês, *complementary slackness*) e suas implicações, para o caso cônico, estão reunidas no seguinte lema.

**Lema 2.2** *Dadas as matrizes  $X$  e  $Z$ , matrizes factíveis, respectivamente, para os problemas (2.26) e (2.29), a seguinte relação é sempre válida*

$$\langle X, C \rangle \geq -\langle Z, D \rangle \quad (2.31)$$

*isto é, o valor do objetivo dual é limitado superiormente pelo valor do objetivo primal de qualquer ponto primal factível. Caso existam matrizes  $X$  estritamente primal factível e  $Z$  estritamente dual factível, essa relação é respeitada na igualdade apenas para as matrizes  $X^*$  e  $Z^*$ , respectivamente, pontos ótimos primal e dual, ou seja*

$$\eta^* = \langle X^*, Z^* \rangle = 0 \quad (2.32)$$

**Prova:** Como  $X \geq 0$ ,  $Z \geq 0$  e (2.30) se aplicam, tem-se

$$\langle X, Z \rangle = \text{tr}(XZ) \geq 0 \rightarrow \langle X, C \rangle + \langle Z, D \rangle \geq 0$$

Uma prova para a segunda parte do lema pode ser encontrada no parágrafo 30 em Rockafellar [21] ou na seção 4.2 em Nesterov & Nemirovskii [14].  $\square$

Esse lema provê respaldo extra às condições necessárias à obtenção de soluções via algoritmos de pontos interiores, pois a existência de pontos factíveis interiores para os problemas (2.1) e (2.8) também é condição imprescindível para que a função de barreira (2.9) seja bem definida. Os problemas para os quais isso não ocorre devem poder ser detectados pelos métodos de obtenção de um ponto inicial estritamente factível<sup>3</sup>.

<sup>3</sup>Veja seção 3.5.

Por último, é importante ressaltar que o conjunto  $\{x \mid F(x) \geq 0\}$ , definido no espaço dos vetores  $x$  de dimensão  $n$ , e o conjunto  $\{X \mid X \in \mathcal{K} \cap (\mathcal{L} + D)\}$ , definido no espaço das matrizes reais e simétricas  $X$  de dimensão  $p \times p$ , são isomorfos segundo a relação  $X = F(x)$ . Assim, deste ponto em diante, a letra grega  $\Phi$  será utilizada para identificá-los de maneira ambígua, o mesmo acontecendo em relação aos seus respectivos *interiores*, que passarão a ser denotados por  $\Phi^0$ . Embora possa causar uma certa confusão, isso é feito de modo a simplificar a notação dos problemas que possam ser simultaneamente formulados em termos vetoriais e matriciais.

## 2.5 Projeções

Nesta seção e nas subseqüentes, os conceitos geométricos de projeção e transformação simétrica, dentre outros, serão definidos em relação a entes do espaço vetorial associado com a formulação cônica.

A projeção de um dado vetor  $W$  sobre o subespaço linear  $\mathcal{L}$ , definido na seção anterior, consiste na solução de um problema de otimização do tipo mínimos quadrados. Assim, o vetor projetado  $W_p$  é obtido mediante

$$W_p = \arg \min_{X \in \mathcal{L}} \|W - X\| \quad (2.33)$$

e o vetor  $W_p^\perp$ , definido por  $W_p^\perp = W - W_p$ , é ortogonal a  $W_p$  e a qualquer outro vetor pertencente a  $\mathcal{L}$ . As condições de otimalidade para o problema (2.33) são expressas por

$$\begin{cases} \sum_{i=1}^n \langle F_j, F_i \rangle w_p^i = \langle W, F_j \rangle, j = 1, \dots, n \\ W_p = \sum_{i=1}^n F_i w_p^i \end{cases} \quad (2.34)$$

Reescrevendo-se as  $n$  primeiras equações em termos de uma matriz  $M \in \mathfrak{R}^{n \times n}$  cujos componentes são definidos como

$$M_{i,j} = M_{j,i} := \langle F_j, F_i \rangle \quad (2.35)$$

a matriz projetada  $W_p$  é dada pela expressão

$$W_p = \sum_{i=1}^n F_i \sum_{j=1}^n (M^{-1})_{i,j} \langle W, F_j \rangle = P_{\mathcal{L}} W \quad (2.36)$$

Note-se que  $P_{\mathcal{L}}$  é um operador linear. Fica claro, pela definição dada anteriormente do vetor  $W_p^\perp$ , que o operador complemento ortogonal é dado por  $P_{\mathcal{L}}^\perp = \mathbf{I} - P_{\mathcal{L}}$ .

A idéia de projeção sugere um critério simples para a caracterização de pontos factíveis para o problema dual (2.29) expressa no lema seguinte.

**Lema 2.3** *Uma matriz  $Z \geq 0$  é factível para o problema dual (2.29) se, e somente se,  $P_{\mathcal{L}} Z = P_{\mathcal{L}} C$ .*

**Prova:** Se  $P_{\mathcal{L}} Z = P_{\mathcal{L}} C$ , então  $P_{\mathcal{L}}(Z - C) = 0$ , donde  $Z \in (\mathcal{L}^\perp + C)$ . Como  $Z \geq 0$ , então  $Z$  é dual factível. A necessidade é imediata.  $\square$

## 2.6 Transformação simétrica

A matriz  $\tilde{X}$  resultante da *transformação simétrica* da matriz  $X$  por uma matriz  $E^{-1} > 0$  é obtida, por definição, como o resultado da seguinte operação

$$\tilde{X} = E^{-\frac{1}{2}} X E^{-\frac{1}{2}} \quad (2.37)$$

donde  $X$  pode ser sempre obtido mediante uma transformação da mesma natureza aplicada a  $\tilde{X}$  definida pela matriz  $E > 0$ , isto é,  $X = E^{\frac{1}{2}} \tilde{X} E^{\frac{1}{2}}$ . Se  $X$  pertencer a  $(\mathcal{L} + D)$ , a matriz transformada  $\tilde{X}$  pertencerá à variedade linear  $(\tilde{\mathcal{L}} + \tilde{D})$  definida por

$$(\tilde{\mathcal{L}} + \tilde{D}) = \left\{ \tilde{X} \mid \tilde{X} = \sum_{i=1}^n \tilde{F}_i \alpha_i + \tilde{D}, \alpha_i \in \Re \right\} \quad (2.38)$$

em que  $\tilde{F}_i = E^{-\frac{1}{2}} F_i E^{-\frac{1}{2}}$  e  $\tilde{D} = E^{-\frac{1}{2}} D E^{-\frac{1}{2}}$ . Além do mais, notando-se que

$$\langle C, X \rangle = \text{tr} \left( C E^{\frac{1}{2}} \tilde{X} E^{\frac{1}{2}} \right) = \text{tr} \left( E^{\frac{1}{2}} C E^{\frac{1}{2}} \tilde{X} \right) = \langle \tilde{C}, \tilde{X} \rangle \quad (2.39)$$

em que  $\tilde{C} = E^{\frac{1}{2}} C E^{\frac{1}{2}}$ , o problema primal (2.26) pode ser reescrito em função das matrizes transformadas

$$\begin{aligned} & \text{minimizar} \quad \langle \tilde{C}, \tilde{X} \rangle \\ & \text{s.a.} \quad \tilde{X} \in \{ \mathcal{K} \cap (\tilde{\mathcal{L}} + \tilde{D}) \} \end{aligned} \quad (2.40)$$

Nesse caso, a notação  $\tilde{C}$  indica que a matriz  $C$  sofre uma transformação simétrica definida pela matriz  $E$  e não por  $E^{-1}$ . Permitindo-se essa ambiguidade na notação torna-se possível denotar de maneira uniforme as transformações simétricas sofridas pelos problemas primal e dual, explicadas a seguir. A natureza da matriz transformada, primal ou dual, torna clara qual transformação deve ser efetivamente utilizada. Observando-se que, para uma matriz  $Z$  factível em relação ao problema dual (2.29),

$$\langle D, Z \rangle = \langle E^{\frac{1}{2}} \tilde{D} E^{\frac{1}{2}}, Z \rangle = \langle \tilde{D}, E^{\frac{1}{2}} Z E^{\frac{1}{2}} \rangle \quad (2.41)$$

e

$$\begin{aligned} \left\langle Z - C, \sum_{i=1}^n F_i \alpha_i \right\rangle &= \left\langle Z - C, \sum_{i=1}^n E^{\frac{1}{2}} \tilde{F}_i E^{\frac{1}{2}} \alpha_i \right\rangle \\ &= \left\langle E^{\frac{1}{2}} (Z - C) E^{\frac{1}{2}}, \sum_{i=1}^n \tilde{F}_i \alpha_i \right\rangle = 0 \end{aligned} \quad (2.42)$$

as relações anteriores permitem redefinir o problema dual como

$$\begin{aligned} & \text{minimizar} \quad \langle \tilde{D}, \tilde{Z} \rangle \\ & \text{s.a.} \quad \tilde{Z} \in \{ \mathcal{K} \cap (\tilde{\mathcal{L}}^\perp + \tilde{C}) \} \end{aligned} \quad (2.43)$$

em que  $\tilde{Z} = E^{\frac{1}{2}} Z E^{\frac{1}{2}}$  define a relação entre as variáveis duais  $\tilde{Z}$  transformada e  $Z$  não transformada. O exposto acima está resumidamente expresso no seguinte lema.

**Lema 2.4** Dado o par de matrizes  $(X, Z)$  respectivamente factíveis em relação aos problemas primal (2.26) e dual (2.29), e uma matriz  $E > 0$ , o par de matrizes

$$(\tilde{X}, \tilde{Z}) = (E^{-\frac{1}{2}} X E^{-\frac{1}{2}}, E^{\frac{1}{2}} Z E^{\frac{1}{2}}) \quad (2.44)$$

é factível em relação aos problemas transformados definidos em (2.40) e (2.43).

De maneira análoga, a projeção de uma matriz  $W$  no subespaço transformado  $\tilde{\mathcal{L}}$  é dada por

$$\tilde{W}_p = \arg \min_{\tilde{X} \in \tilde{\mathcal{L}}} \|W - \tilde{X}\| \quad (2.45)$$

e pode ser obtida mediante

$$\tilde{W}_p = \sum_{i=1}^n \tilde{F}_i \sum_{j=1}^n (\tilde{M}^{-1})_{i,j} \langle W, \tilde{F}_j \rangle = P_{\tilde{\mathcal{L}}} W \quad (2.46)$$

em que a matriz  $\tilde{M}$  é dada por

$$\tilde{M}_{i,j} = \tilde{M}_{j,i} = \langle \tilde{F}_i, \tilde{F}_j \rangle = \langle E^{-\frac{1}{2}} F_j E^{-\frac{1}{2}}, E^{-\frac{1}{2}} F_i E^{-\frac{1}{2}} \rangle \quad (2.47)$$

Como é sempre possível decompor uma matriz  $E > 0$  no produto de duas matrizes quadradas não obrigatoriamente simétricas na forma  $E = T'T$ , por exemplo, através da decomposição de Cholesky, as transformações do tipo

$$\tilde{X} = T X T' \quad (2.48)$$

também constituem transformações simétricas. De uma maneira geral, todas as expressões obtidas com o auxílio da transformação  $E^{\frac{1}{2}} X E^{\frac{1}{2}}$  podem ser facilmente reformuladas nesta última forma, o que, em muitos casos, pode representar uma considerável economia de esforço computacional. Apesar disso, haverá um predomínio da primeira forma ao longo de todo o texto a fim de uniformizar a notação e facilitar a leitura. Ainda em relação à notação, quando houver possibilidade de confusão quanto à matriz que define uma transformação simétrica, esta deverá aparecer explicitamente junto à matrix transformada. Como exemplo, o subespaço transformado  $\tilde{\mathcal{L}}$  cujas matrizes básicas sofrem transformações do tipo  $(\tilde{F}_i)_E = E^{\frac{1}{2}} F_i E^{\frac{1}{2}}$  ou  $(\tilde{F}_i)_{T'T} = T F_i T'$  será denotado por  $\tilde{\mathcal{L}}_E$  ou  $\tilde{\mathcal{L}}_{T'T}$ , respectivamente.

## 2.7 O elipsóide de Dikin

A região definida pela desigualdade

$$\varphi(X_0, r) = \{X \mid \|X_0 - X\| < r\} \quad (2.49)$$

é denominada, por analogia ao caso de vetores no espaço tridimensional, uma *esfera de raio  $r$  centrada em  $X_0$* . Uma propriedade interessante pode ser identificada para o caso particular de esferas centradas em  $\mathbf{I}$ , isto é, quando  $X_0$  é a matriz identidade. O fato é que toda matriz  $X$  pertencente a uma esfera cujo raio seja menor do que um, isto é,  $X \in \varphi(\mathbf{I}, r < 1)$ , é estritamente definida positiva, e isso decorre da seguinte observação: se  $X \in \varphi(\mathbf{I}, r < 1)$  e  $Y = \mathbf{I} - X$ , então  $\|Y\| = \|\mathbf{I} - X\| = (\sum_{i=1}^p \lambda_i^2(Y))^{1/2} < 1$  o que implica  $\max_i |\lambda_i(Y)| < 1$  que,

**Lema 2.4** Dado o par de matrizes  $(X, Z)$  respectivamente factíveis em relação aos problemas primal (2.26) e dual (2.29), e uma matriz  $E > 0$ , o par de matrizes

$$(\tilde{X}, \tilde{Z}) = (E^{-\frac{1}{2}} X E^{-\frac{1}{2}}, E^{\frac{1}{2}} Z E^{\frac{1}{2}}) \quad (2.44)$$

é factível em relação aos problemas transformados definidos em (2.40) e (2.43).

De maneira análoga, a projeção de uma matriz  $W$  no subespaço transformado  $\tilde{\mathcal{L}}$  é dada por

$$\tilde{W}_p = \arg \min_{\tilde{X} \in \tilde{\mathcal{L}}} \|W - \tilde{X}\| \quad (2.45)$$

e pode ser obtida mediante

$$\tilde{W}_p = \sum_{i=1}^n \tilde{F}_i \sum_{j=1}^n (\tilde{M}^{-1})_{i,j} \langle W, \tilde{F}_j \rangle = P_{\tilde{\mathcal{L}}} W \quad (2.46)$$

em que a matriz  $\tilde{M}$  é dada por

$$\tilde{M}_{i,j} = \tilde{M}_{j,i} = \langle \tilde{F}_i, \tilde{F}_j \rangle = \langle E^{-\frac{1}{2}} F_j E^{-\frac{1}{2}}, E^{-\frac{1}{2}} F_i E^{-\frac{1}{2}} \rangle \quad (2.47)$$

Como é sempre possível decompor uma matriz  $E > 0$  no produto de duas matrizes quadradas não obrigatoriamente simétricas na forma  $E = T'T$ , por exemplo, através da decomposição de Cholesky, as transformações do tipo

$$\tilde{X} = T X T' \quad (2.48)$$

também constituem transformações simétricas. De uma maneira geral, todas as expressões obtidas com o auxílio da transformação  $E^{\frac{1}{2}} X E^{\frac{1}{2}}$  podem ser facilmente reformuladas nesta última forma, o que, em muitos casos, pode representar uma considerável economia de esforço computacional. Apesar disso, haverá um predomínio da primeira forma ao longo de todo o texto a fim de uniformizar a notação e facilitar a leitura. Ainda em relação à notação, quando houver possibilidade de confusão quanto à matriz que define uma transformação simétrica, esta deverá aparecer explicitamente junto à matrix transformada. Como exemplo, o subespaço transformado  $\tilde{\mathcal{L}}$  cujas matrizes básicas sofrem transformações do tipo  $(\tilde{F}_i)_E = E^{\frac{1}{2}} F_i E^{\frac{1}{2}}$  ou  $(\tilde{F}_i)_{T'T} = T F_i T'$  será denotado por  $\tilde{\mathcal{L}}_E$  ou  $\tilde{\mathcal{L}}_{T'T}$ , respectivamente.

## 2.7 O elipsóide de Dikin

A região definida pela desigualdade

$$\varphi(X_0, r) = \{X \mid \|X_0 - X\| < r\} \quad (2.49)$$

é denominada, por analogia ao caso de vetores no espaço tridimensional, uma *esfera de raio  $r$  centrada em  $X_0$* . Uma propriedade interessante pode ser identificada para o caso particular de esferas centradas em  $\mathbf{I}$ , isto é, quando  $X_0$  é a matriz identidade. O fato é que toda matriz  $X$  pertencente a uma esfera cujo raio seja menor do que um, isto é,  $X \in \varphi(\mathbf{I}, r < 1)$ , é estritamente definida positiva, e isso decorre da seguinte observação: se  $X \in \varphi(\mathbf{I}, r < 1)$  e  $Y = \mathbf{I} - X$ , então  $\|Y\| = \|\mathbf{I} - X\| = (\sum_{i=1}^p \lambda_i^2(Y))^{1/2} < 1$  o que implica  $\max_i |\lambda_i(Y)| < 1$  que,

por sua vez, garante que  $X = \mathbf{I} - Y$  seja definida positiva. O contrário também é verdadeiro, ou seja, se  $\|X\| < 1$  então  $Y > 0 \rightarrow X < \mathbf{I}$ .

Caso se aplique uma transformação simétrica a  $\varphi(X_0, r)$ , as matrizes pertencentes a essa região delimitarão, no espaço transformado, um *elipsóide* centrado em  $\tilde{X}_0$ . Esta idéia permite formular o seguinte lema.

**Lema 2.5** *Toda matriz  $X$  pertencente ao elipsóide definido para  $X_0 > 0$  por*

$$\tilde{\varphi}(X_0, r) = \left\{ X \mid \|\mathbf{I} - X_0^{-\frac{1}{2}} X X_0^{-\frac{1}{2}}\| < r \right\} \quad (2.50)$$

*cujos raio seja menor do que um e estritamente definida positiva. O maior elipsóide que contém apenas matrizes definidas positivas, ou seja, o de raio  $r = 1$ , é chamado elipsóide de Dikin.*

*Por sua vez, se  $\|X_0^{-\frac{1}{2}} X X_0^{-\frac{1}{2}}\| < 1$  então  $X_0 - X > 0$ .*

**Prova:** A prova desse lema decorre diretamente da discussão anterior sobre as esferas centradas em  $\mathbf{I}$ . Assim, quando  $r < 1$ , a matriz  $X_0^{-\frac{1}{2}} X X_0^{-\frac{1}{2}}$  é sempre estritamente definida positiva, donde decorre diretamente que  $X > 0$ , já que  $X_0 > 0$ . Note-se que esse resultado pode ser interpretado como a aplicação de uma transformação simétrica às matrizes  $X$  e  $X_0$  em relação à matriz  $X_0^{-1}$ , de tal modo que a matriz transformada  $\tilde{X}$  pertença a uma esfera centrada em  $\mathbf{I}$ . A prova da segunda parte do lema é também imediata, pois  $X_0^{-\frac{1}{2}} X X_0^{-\frac{1}{2}} < \mathbf{I}$  o que implica  $X_0 - X > 0$ , para  $X_0 > 0$ .  $\square$

As seguintes definições de produto interno *transformado*

$$\langle X, Y \rangle_E := \langle E^{\frac{1}{2}} X E^{\frac{1}{2}}, E^{\frac{1}{2}} Y E^{\frac{1}{2}} \rangle = \text{tr}(X E Y E) \quad (2.51)$$

e sua norma induzida associada

$$\|X\|_E := \|E^{\frac{1}{2}} X E^{\frac{1}{2}}\| = \sqrt{\text{tr}(X E X E)} \quad (2.52)$$

fazem com que esferas centradas em  $X_0$ , definidas em relação ao produto interno transformado,

$$\varphi_{X_0^{-1}}(X_0, r) = \|X_0 - X\|_{X_0^{-1}} < r \quad (2.53)$$

coincidam com elipsóides centrados em  $X_0$  definidos em relação ao produto interno usual.

Pelo lema acima é possível obter uma região elipsoidal em torno de  $X_0$  onde a positividade das matrizes pertencentes a  $\mathcal{L}$  está assegurada. Caso se busque esta mesma propriedade sobre uma reta que passe por  $X_0$ , a existência de limitantes mais estritos é garantida pelo próximo lema.

**Lema 2.6** *Dado  $X_0 > 0$  e uma direção  $H$  qualquer, os valores de  $\gamma$  que garantem que a matriz  $Y = X_0 + \gamma H$  seja definida positiva pertencem ao intervalo*

$$\frac{1}{-\lambda_M(\Psi(H))} \leq \gamma \leq \frac{1}{-\lambda_m(\Psi(H))} \quad (2.54)$$

em que  $\Psi(H) = X_0^{-\frac{1}{2}} H X_0^{-\frac{1}{2}}$ ,  $\lambda_m(\cdot)$  e  $\lambda_M(\cdot)$  denotam, respectivamente, o menor autovalor negativo e o maior autovalor positivo de  $(\cdot)$ . Caso  $\Psi(H)$  não apresente autovalores positivos,  $-1/\lambda_M(\Psi(H))$  deve ser entendido como menos infinito. O mesmo se aplica para  $-1/\lambda_m(\Psi(H))$ , entendido como infinito, caso  $\Psi(H)$  não apresente autovalores negativos.

**Prova:** Como  $X_0 > 0$ , pode-se escrever, para que  $Y$  pertença à fronteira de  $\mathcal{K}$

$$\begin{aligned}\det Y &= \det(X_0 + \gamma H) \\ &= \det X_0 \det(\mathbf{I} + \gamma \Psi(H)) = 0\end{aligned}$$

donde o determinante de  $Y$  será nulo quando  $\gamma$  for igual ao oposto do inverso de um dos autovalores de  $\Psi(H)$ . A fronteira do cone  $\mathcal{K}$  é alcançada para o menor valor positivo e para o maior valor negativo de  $\gamma$ , que fornecem os limitantes acima. Caso  $H$  aponte para uma direção em que o cone  $\mathcal{K}$  é aberto, então  $\Psi(H)$  apresentará autovalores somente positivos ou negativos e um dos limitantes deve ser entendido como infinito.  $\square$

Esse resultado também pode ser estendido para o conjunto vetorial  $\Phi$  e, nesse caso, indica quais os valores de  $\gamma$  que mantêm  $(x_0 + \gamma h)$  ainda interiores a  $\Phi$ . A transcrição do resultado é imediata, desde que se considere  $\Psi(h)$ , definido no lema 2.1, em vez de  $\Psi(H)$ , nas expressões acima.

## 2.8 Uma barreira para o cone positivo

Na formulação cônica, tanto para o problema primal quanto para o problema dual, o conjunto de factibilidade se resume à interseção do cone  $\mathcal{K}$  com uma variedade linear. Nesse caso, para se obter uma solução por meio de algoritmos de pontos interiores, deve-se definir uma função de barreira para o cone  $\mathcal{K}$  e utilizar técnicas de otimização restrita a variedades lineares. No que se refere à escolha da função de barreira para  $\mathcal{K}$ , a função

$$B_{\mathcal{K}}(X) := \begin{cases} -\log \det X, & \text{para } X \in \mathcal{K}^0 \\ \infty, & \text{em qualquer outra condição} \end{cases} \quad (2.55)$$

análoga à função de barreira vetorial definida na seção 2.3, é uma excelente candidata, principalmente sabendo-se que várias das propriedades discutidas no lema 2.1, permanecem válidas. Esse é o assunto do próximo lema.

**Lema 2.7** *Dado  $B_{\mathcal{K}}(X)$ , uma função de barreira para o conjunto convexo  $\mathcal{K}$ , o cone das matrizes definidas positivas, definida em (2.55), os seguintes fatos são verdadeiros*

- i) *O valor da barreira para pontos pertencentes à variedade linear  $(X + \gamma H)$ ,  $X \in \mathcal{K}^0$ ,  $H$  simétrica pertencente a  $\mathfrak{R}^{p \times p}$  e  $\gamma \in \mathfrak{R}$ , é dado por*

$$B_{\mathcal{K}}(X + \gamma H) = B_{\mathcal{K}}(X) - \sum_{i=1}^p \log [1 + \gamma \lambda_i(\Psi(H))], \quad (2.56)$$

em que  $\Psi(H)$  é uma matriz definida por

$$\Psi(H) := X^{-\frac{1}{2}} H X^{-\frac{1}{2}} \quad (2.57)$$

- ii) *As derivadas direcionais primeira, segunda e terceira de  $B_{\mathcal{K}}(X)$ , tomadas no sentido da matriz  $H$ , são dadas por*

$$DB_{\mathcal{K}}(X, H) = -\operatorname{tr}(\Psi(H)) \quad (2.58)$$

$$D^2 B_{\mathcal{K}}(X, H) = \operatorname{tr}(\Psi^2(H)) \quad (2.59)$$

$$D^3 B_{\mathcal{K}}(X, H) = -2 \operatorname{tr}(\Psi^3(H)) \quad (2.60)$$

e obedecem às seguintes desigualdades

$$|DB_{\mathcal{K}}(X, H)| \leq \left(D^2 B_{\mathcal{K}}(X, H)\right)^{1/2} \quad (2.61)$$

$$|D^3 B_{\mathcal{K}}(X, H)| \leq 2 \left(D^2 B_{\mathcal{K}}(X, H)\right)^{3/2} \quad (2.62)$$

iii) A matriz gradiente de  $B_{\mathcal{K}}(X)$  é dada por

$$\nabla B_{\mathcal{K}}(X) = -X^{-1} \quad (2.63)$$

e o mapeamento da matriz  $H \in \mathfrak{R}^{p \times p} \rightarrow \mathfrak{R}^{p \times p}$  definido pela Hessiana de  $B_{\mathcal{K}}(X)$  é dado por

$$\nabla^2 B_{\mathcal{K}}(X)H = X^{-1}HX^{-1} \quad (2.64)$$

iv) Uma transformação simétrica não afeta as variações da função  $B_{\mathcal{K}}(\cdot)$ , isto é, dados  $X, Y$  e  $E$  matrizes estritamente definidas positivas

$$B_{\mathcal{K}}(X) - B_{\mathcal{K}}(Y) = B_{\mathcal{K}}\left(E^{\frac{1}{2}}XE^{\frac{1}{2}}\right) - B_{\mathcal{K}}\left(E^{\frac{1}{2}}YE^{\frac{1}{2}}\right) \quad (2.65)$$

**Prova:** A prova dos pontos *i*) e *ii*) segue o mesmo esquema da prova apresentada para o lema 2.1 e será omitida. Note-se, mais uma vez, que as expressões das derivadas garantem que  $B_{\mathcal{K}}(X)$  é uma função estritamente convexa, pois  $D^2 B_{\mathcal{K}}(X)$  vale zero somente quando  $H = \mathbf{0}$ , e as desigualdades apresentadas no ponto *ii*) garantem que  $B_{\mathcal{K}}(X)$  seja autoconcordante. O ponto *iii*) pode ser justificado a partir da expansão em séries de potência da função  $B_{\mathcal{K}}(X)$  em torno de  $X$  que, por sua vez, decorre diretamente do ponto *ii*). Assim, para  $X > 0$  e  $H$  “pequeno”<sup>4</sup>

$$\begin{aligned} B_{\mathcal{K}}(X + H) &= \log \det(X + H)^{-1} \\ &\approx B_{\mathcal{K}}(X) + \langle -X^{-1}, H \rangle + \frac{1}{2} \langle X^{-1}HX^{-1}, H \rangle \end{aligned}$$

Finalmente, o ponto *iv*) pode ser facilmente verificado, levando-se em conta que

$$B_{\mathcal{K}}\left(E^{\frac{1}{2}}XE^{\frac{1}{2}}\right) = -\log \det\left(E^{\frac{1}{2}}XE^{\frac{1}{2}}\right) = B_{\mathcal{K}}(X) + B_{\mathcal{K}}(E)$$

□

## 2.9 Otimização restrita a variedades lineares

Continuando-se a discussão iniciada na seção anterior, serão discutidos nesta seção alguns aspectos relativos à otimização de funções matriciais cujas variáveis estejam restritas a variedades lineares. O problema em questão consiste em

$$\begin{aligned} &\text{minimizar } f(X) \\ &\text{s.a } X \in \Phi^0 \end{aligned} \quad (2.66)$$

em que  $f(X)$  é uma função matricial contínua e diferenciável em  $\Phi^0$ . Num ponto  $X_0 \in \Phi^0$ , como  $f(X)$  é contínua e diferenciável, é possível determinar a chamada direção de máximo decrescimento (do inglês, *steepest-descent direction*) utilizando-se o gradiente de  $f(X)$ . A determinação dessa direção é discutida no próximo lema.

<sup>4</sup>Em termos da norma definida na seção 2.4.

**Lema 2.8** Em  $X_0 \in \Phi^0$ , a direção de máximo decrescimento da função  $f(X)$  é dada pela matriz  $H \in \mathcal{L}$  que minimiza a linearização da função  $f(X)$  em torno de  $X_0$  numa esfera centrada em  $X_0$ , ou seja, a solução do problema

$$\arg \min_{H \in \mathcal{L}} \{ \langle \nabla f(X_0), H \rangle \mid \|H\| = r \} \quad (2.67)$$

que, para qualquer valor de  $r > 0$ , está sempre no sentido da matriz

$$-P_{\mathcal{L}} \nabla f(X_0) \quad (2.68)$$

**Prova:** A função lagrangeana associada ao problema (2.67) para uma esfera de raio  $r_0$  é

$$L(H, \mu) := \langle \nabla f(X_0), H \rangle + \mu (\langle H, H \rangle - r_0^2)$$

que, para  $H \in \mathcal{L}$ , implica as seguintes condições de otimalidade de primeira ordem

$$\begin{cases} \langle F_i, \nabla f(X_0) \rangle + 2\mu \langle F_i, H \rangle = 0, i = 1, \dots, n \\ \langle H, H \rangle = r_0^2 \end{cases}$$

e de segunda ordem

$$2\mu M > 0$$

em que  $M$  é a matriz definida em (2.35) que, por construção, é sempre definida positiva, donde  $\mu > 0$ . Por sua vez, as  $n$  primeiras equações das condições de otimalidade de primeira ordem implicam

$$\langle F_i, \nabla f(X) + 2\mu H \rangle = 0, i = 1, \dots, n \rightarrow (\nabla f(X) + 2\mu H) \in \mathcal{L}^\perp$$

donde

$$P_{\mathcal{L}} H = H = -\frac{1}{2\mu} P_{\mathcal{L}} \nabla f(X)$$

Note-se que, para qualquer valor de  $r_0$ , a matriz  $H$  apresenta-se sempre no sentido da matriz  $-P_{\mathcal{L}} \nabla f(X)$  multiplicada pelo escalar positivo  $1/2\mu$ , o que completa a prova.  $\square$

A determinação da direção de máximo decrescimento pressupõe a escolha de regiões esféricas, sobre as quais se minimiza a linearização de  $f(X)$ . Essa escolha, além de prover uma solução simples para o problema, é feita de modo a não privilegiar qualquer direção, o que é razoável quando  $X$  pode ser qualquer matriz pertencente à variedade linear  $(\mathcal{L} + D)$ . Porém, além de pertencer a  $(\mathcal{L} + D)$ ,  $X$  deve também pertencer ao cone  $\mathcal{K}$ . Ora, a minimização da linearização de  $f(X)$  numa região esférica próxima à fronteira do cone  $\mathcal{K}$  não leva em conta o fato de não se poder cruzá-la, o que motiva a utilização de outras regiões de minimização. Nesse caso, uma região de minimização, bastante interessante do ponto de vista teórico, seria a família de elipsóides centrados em  $X_0 > 0$  definida por  $\tilde{\varphi}(X_0, r)$ . Procurar uma direção de máximo decrescimento sobre esses elipsóides que, conforme visto na seção (2.7), contêm o elipsóide de Dikin, que sempre *tangencia* o cone  $\mathcal{K}$  em  $p$  pontos, significaria ponderar as matrizes pertencentes à esfera original por um fator proporcional à sua “positividade”, medida pela norma transformada definida pela matriz  $X_0^{-1}$ . Do ponto

de vista prático, uma transformação simétrica definida por  $X_0^{-1}$  é capaz de mapear esses elipsóides em esferas centradas em  $\mathbf{I}$ , o que permite que a determinação da direção de máximo decrescimento seja efetuada com base no resultado do lema acima. Detalhando-se um pouco mais, para  $H \in \mathcal{L}$  e a transformação simétrica definida por  $X_0^{-1}$ , as relações

$$\|H\|_{X_0^{-1}} = \|\tilde{H}\| = r \quad (2.69)$$

e

$$\langle \nabla f(X_0), H \rangle = \left\langle X_0^{\frac{1}{2}} \nabla f(X_0) X_0^{\frac{1}{2}}, \tilde{H} \right\rangle = \langle \tilde{\nabla} f(X_0), \tilde{H} \rangle \quad (2.70)$$

permitem que, a fim de se determinar a direção de máximo decrescimento sobre a família de elipsóides  $\tilde{\varphi}(X_0, r)$ , o seguinte problema seja resolvido

$$\arg \min_{\tilde{H} \in \tilde{\mathcal{L}}} \left\{ \langle \tilde{\nabla} f(X_0), \tilde{H} \rangle \mid \|\tilde{H}\| = r \right\} \quad (2.71)$$

cuja solução, conforme o lema 2.8, está sempre no sentido de  $\tilde{H} = -P_{\tilde{\mathcal{L}}} \tilde{\nabla} f(X_0)$ , donde  $H$  pode ser obtido através da transformação inversa  $H = X_0^{\frac{1}{2}} \tilde{H} X_0^{\frac{1}{2}}$ .

## 2.10 O centro analítico

Se o conjunto  $\Phi^0$  for limitado e diferente de vazio, isto é, se todo  $X \in \Phi^0 \rightarrow \|X\| < \infty$ , então o *centro analítico*  $\bar{X}$  de  $\Phi$  pode ser definido como a solução do seguinte problema

$$\begin{aligned} & \text{minimizar } B_{\mathcal{K}}(X) \\ & \text{s.a } X \in \Phi^0 \end{aligned} \quad (2.72)$$

Note-se que, como  $B_{\mathcal{K}}(X)$  é uma função estritamente convexa para  $X \in \Phi^0$ ,  $\bar{X}$  é único. A condição de otimalidade associada a (2.72) é dada por

$$-P_{\mathcal{L}} \nabla B_{\mathcal{K}}(\bar{X}) = P_{\mathcal{L}} \bar{X}^{-1} = \mathbf{0} \quad (2.73)$$

O problema que define o centro analítico, por si só, não apresenta nenhuma relação com o problema de minimização (2.26), ou seja,  $\bar{X}$  depende apenas de  $(\mathcal{L} + D)$  e não da matriz objetivo  $C$ . Entretanto, uma versão ligeiramente modificada, na qual o problema (2.72) é resolvido num hiperplano definido pelo objetivo  $C$  é capaz de fornecer valiosas informações. Esse problema é

$$\begin{aligned} & \text{minimizar } B_{\mathcal{K}}(X) \\ & \text{s.a } \begin{cases} X \in \Phi^0 \\ \langle C, X \rangle = \rho \end{cases} \end{aligned} \quad (2.74)$$

e suas condições de otimalidade podem ser obtidas por meio da função lagrangeana

$$L(X, \mu) := -\log \det X + \mu (\langle C, X \rangle - \rho) \quad (2.75)$$

em que  $X \in (\mathcal{L} + D)$  implica

$$\begin{cases} \mu \langle F_i, C \rangle - \langle F_i, X^{-1} \rangle = 0, i = 1, \dots, n \\ \langle C, X \rangle = \rho \end{cases} \quad (2.76)$$

cuja solução, para um dado  $\rho$ , é denotada por  $\bar{X}(\rho)$ . É interessante notar que as condições de otimalidade de segunda ordem desse problema se reduzem a testar a positividade da matriz Hessiana de  $B_{\mathcal{K}}(X)$  que, segundo o lema 2.7, é sempre definida positiva, o que implica  $\mu > 0$ . A curva parametrizada em  $\rho$  descrita pelos pontos  $\bar{X}(\rho)$  é chamada de *trajetória central*. Pontos nessa curva apresentam propriedades bastante interessantes e uma delas é o assunto do próximo lema.

**Lema 2.9** *Dado um ponto  $\bar{X}(\rho)$  pertencente à trajetória central associada ao conjunto  $\Phi$ , a matriz  $\bar{Z}(\rho) = \bar{X}^{-1}(\rho)/\mu$ , em que  $\mu > 0$  é o multiplicador de lagrange definido em (2.75), é factível em relação ao problema dual definido em (2.29) e o gap de dualidade associado ao par  $(\bar{X}(\rho), \bar{Z}(\rho))$  é dado por  $\eta(\rho) = p/\mu$ .*

**Prova:** A prova desse lema decorre diretamente das condições de otimalidade (2.76). As  $n$  primeiras equações podem ser reescritas

$$\langle F_i, \mu C - \bar{X}^{-1}(\rho) \rangle = 0, i = 1, \dots, n$$

o que implica  $(\mu C - \bar{X}^{-1}(\rho)) \in \mathcal{L}^\perp$  donde, para  $\mu > 0$ ,

$$P_{\mathcal{L}} \frac{\bar{X}^{-1}(\rho)}{\mu} = P_{\mathcal{L}} C \quad (2.77)$$

que, à luz do lema 2.3, assegura que  $\bar{Z}(\rho)$  é factível para o problema dual (2.29). Assim, pode-se calcular o *gap* de dualidade

$$\eta(\rho) = \langle \bar{X}(\rho), \bar{Z}(\rho) \rangle = \frac{\text{tr}(\mathbf{I})}{\mu} = \frac{p}{\mu}$$

□

Note-se que também é possível associar ao problema dual (2.29) uma trajetória central  $\bar{Z}(\lambda)$  parametrizada em termos do objetivo dual  $\langle D, Z \rangle = \lambda$ . Como a barreira para o problema dual é dada por  $B_{\mathcal{K}}(Z)$ , os resultados serão idênticos aos obtidos no lema anterior, o que permite concluir que  $\bar{Z}(\rho)$  é também um ponto pertencente à trajetória central dual  $\bar{Z}(\lambda)$  para  $\lambda = p/\mu - \rho$ .

A importância do lema 2.9 reside no seguinte fato: pontos pertencentes a uma trajetória central, quer seja primal, quer seja dual, estão associadas a um determinado *gap* de dualidade. Ora, se for possível gerar uma seqüência<sup>5</sup> de valores para  $\rho$  tal que  $\eta(\rho) \rightarrow 0$ , a seqüência de pontos pertencentes à trajetória central correspondente, obtida mediante sucessivas resoluções do problema de *centralização* (2.74), convergirá para a solução do problema (2.26). Esta é a idéia por trás dos métodos chamados seguidores de trajetória (do inglês, *path-following*). Estes métodos serão descritos no capítulo 3.

Por fim, o conceito de trajetória central pode ser estendido para a barreira associada ao problema primal (2.1)  $B(x)$ . O problema equivalente a (2.74) é dado por

$$\begin{array}{ll} \text{minimizar} & B(x) \\ \text{s.a.} & \begin{cases} x \in \Phi^0 \\ c'x = \rho \end{cases} \end{array} \quad (2.78)$$

<sup>5</sup>O mesmo é válido para uma seqüência de valores para  $\lambda$ .

cuja condições de otimalidade podem ser obtidas por meio da função lagrangeana

$$L(x, \mu) := -\log \det F(x) + \mu (c'x - \rho) \quad (2.79)$$

e são dadas por

$$\begin{cases} \operatorname{tr}(F_i F(x)^{-1}) = \mu c_i, i = 1, \dots, n \\ c'x = \rho \end{cases} \quad (2.80)$$

Como anteriormente, a trajetória central pode ser definida como a curva parametrizada em  $\rho$  constituída pelos pontos  $\bar{x}(\rho)$  que satisfazem às condições de otimalidade. As  $n$  primeiras equações das condições de otimalidade coincidem, porém, com as restrições do problema dual (2.8), o que permite definir uma matriz dual  $\bar{Z}(\rho) = F(\bar{x}(\rho))^{-1}/\mu$  factível e obter os mesmos resultados do lema 2.9. Esta incrível “coincidência” nada mais é do que a expressão de que problemas de otimização envolvendo o objetivo linear definido em (2.1) associado à barreira (2.9) ou envolvendo o objetivo definido em (2.26) associado à barreira (2.55) restrita à variedade linear  $(\mathcal{L} + D)$ , são completamente equivalentes e suas soluções ótimas estão relacionadas por meio da identidade  $X^* = F(x^*)$ .

Como uma última observação, note-se que a denominação “centro analítico” pode ser mais bem compreendida nesse contexto vetorial, uma vez que o acréscimo de restrições redundantes em  $F(x) \geq 0$  é capaz de alterar a posição do centro  $\bar{x}(\rho)$  sem alterar a forma do conjunto  $\Phi$ , ou seja, o centro está ligado à formulação “analítica”  $F(x)$  das restrições, e não à forma do conjunto  $\Phi$ . O mesmo já não ocorre em relação aos centros  $\bar{X}(\rho)$ , pois um acréscimo de restrições redundantes efetivamente altera a forma do conjunto  $\Phi$ , que pertence, por definição, ao subespaço vetorial gerado pelas próprias restrições.



## Capítulo 3

# Algoritmos: aspectos conceituais

### 3.1 Calculando o centro analítico

Nesta seção apresenta-se um algoritmo para resolução do problema (2.66) introduzido na seção 2.9. Aplicando-se esse algoritmo a funções adequadas, definidas ao longo desse capítulo, encontram-se de modo eficiente os centros analíticos associados ao conjunto  $\Phi$ .

#### 3.1.1 O algoritmo básico

O algoritmo seguinte baseia-se inteiramente na discussão do capítulo anterior acerca da determinação da direção de máximo decrescimento sobre regiões elípticas.

**Algoritmo 1** Dado  $X_0 \in \Phi^0$  e  $f(X) : \mathbb{R}^{p \times p} \rightarrow \mathbb{R}$  contínua e diferenciável

1. Aplicar a transformação simétrica:  $\tilde{\nabla} f(X_k) = X_k^{\frac{1}{2}} \nabla f(X_k) X_k^{\frac{1}{2}}$ ,
2. Calcular a direção de máximo decrescimento para o gradiente transformado:  $\tilde{H}_k = \Psi(H_k) = -P_{\tilde{L}} \tilde{\nabla} f(X_k)$ ,
3. Calcular a matriz  $Y_k$  via:  $Y_k = \mathbf{I} + \gamma_k \Psi(H_k)$ , selecionando-se  $\gamma_k$  de tal modo que  $Y_k$  ainda pertença a  $\Phi$ ,
4. Atualizar  $X_{k+1} = X_k^{\frac{1}{2}} Y_k X_k^{\frac{1}{2}}$ .
5. Parar segundo critério de convergência. Caso contrário, incrementar  $k$  e voltar ao primeiro passo.

□

O algoritmo 1 foi proposto no contexto de programação linear por Dikin, em 1967, aplicado à própria função objetivo linear, isto é  $f(X) = \langle C, X \rangle$ . Empregava-se o passo  $\gamma_k = 1/\|\Psi(H_k)\|$  que, conforme o lema 2.5, assegura que  $\|\gamma_k \Psi(H_k)\| < 1$  e, portanto, que  $Y_k$  se mantém sempre definida positiva. Nessa forma, esse algoritmo ficou conhecido pelo nome, em inglês, *affine-scaling algorithm*. Ainda no contexto de programação linear e aplicado à função objetivo linear, vários pesquisadores utilizaram esse algoritmo estabelecendo um passo igual a uma percentagem do limitante positivo para  $\gamma_k$  fornecido pelo lema 2.6. Um

método completamente equivalente a este (porém baseado em um argumento diverso) foi proposto por Oliveira & Geromel em [18] para a resolução do problema (2.26). Esse tipo de implementação já apresenta um desempenho razoável.

Quando aplicado a funções não lineares o algoritmo acima é conhecido pelo nome, em inglês, *scaling steepest descent algorithm*. Ao longo desse capítulo, consideram-se outras funções  $f(X)$ .

### 3.1.2 A função apenada

Chama-se função apenada (do inglês, *penalized function*) a uma combinação linear da função objetivo e da função de barreira, intensamente explorada no trabalho de Fiacco & McCormick [6] e já apresentada no capítulo 1. Para a função de barreira (2.55) a função apenada associada ao problema (2.26) é dada por

$$f_\beta(X) = \beta \langle C, X \rangle + B_\kappa(X) \quad (3.1)$$

cujo gradiente vale

$$\nabla f_\beta(X) = \beta C - X^{-1} \quad (3.2)$$

Para um valor fixo do parâmetro  $\beta$ , o mínimo em  $\Phi$  dessa função estritamente convexa ocorre num ponto  $\bar{X}(\beta)$  no qual a projeção do gradiente de  $f_\beta(X)$  em  $\mathcal{L}$  é nula, isto é

$$P_{\mathcal{L}} \nabla f_\beta(\bar{X}(\beta)) = 0 \rightarrow P_{\mathcal{L}} \frac{\bar{X}^{-1}(\beta)}{\beta} = P_{\mathcal{L}} C \quad (3.3)$$

Uma vez que essa expressão se iguala à condição (2.77) com  $\beta = \mu$ ,  $\bar{X}(\beta)$  também é um ponto na trajetória central de  $\Phi$ . Além do mais, conforme o lema 2.9, o fato de  $\beta$  estar diretamente relacionado com o multiplicador de Lagrange  $\mu$  permite associar ao ponto  $\bar{X}(\beta)$  o *gap* de dualidade  $\eta(\beta) = p/\beta$ . Essa equivalência entre a trajetória central e a trajetória definida pela função apenada faz com que uma seqüência de pontos  $\bar{X}(\beta_l)$  em que o parâmetro  $\beta_l$  é atualizado pela regra  $\beta_{l+1} = \beta_l/\theta$ ,  $\theta \in (0, 1)$ , convirja para a solução de (2.26) de tal forma que  $\eta(\beta_{l+1}) = \theta\eta(\beta_l)$ , ou seja, a cada solução do problema de centralização  $\bar{X}(\beta_{l+1})$  o *gap* de dualidade é dividido por uma razão constante  $\theta$ .

Obviamente pode-se utilizar o algoritmo 1 na determinação dos pontos  $\bar{X}(\beta_l)$ . Nesse caso, é de interesse que o passo  $\gamma_k$  seja escolhido não apenas de modo a manter a matriz  $Y_k$  definida positiva mas também a fim de minimizar a função  $f_{\beta_k}(Y_k)$ , o que pode ser realizado por um procedimento de busca unidimensional, assunto da seção 3.1.4. Pode-se também utilizar como alternativa ao algoritmo 1 na solução dos problemas convexos de centralização um método do tipo Newton modificado. Em relação à função apenada, esta segunda possibilidade, embora conceitualmente diversa, é completamente equivalente à primeira, como mostra o seguinte lema.

**Lema 3.1** *O algoritmo 1, para a função apenada  $f_\beta(X)$ , é completamente equivalente ao método de Newton modificado aplicado a essa mesma função.*

**Prova:** A direção de Newton  $H$  é a direção que minimiza a aproximação quadrática da função  $f_\beta(X)$ , o que, segundo o lema 2.7, consiste em resolver o seguinte problema

$$H = \arg \min_{H \in \mathcal{L}} \left\{ \langle \nabla f_\beta(X), H \rangle + \frac{1}{2} \langle X^{-1} H, X^{-1} H \rangle \right\}$$

Definindo-se  $\tilde{H} = X^{-\frac{1}{2}} H X^{-\frac{1}{2}}$ , o problema anterior transforma-se em

$$\min_{\tilde{H} \in \tilde{\mathcal{L}}} \left\{ \left\langle X^{\frac{1}{2}} \nabla f_{\beta}(X) X^{\frac{1}{2}}, \tilde{H} \right\rangle + \frac{1}{2} \left\langle \tilde{H}, \tilde{H} \right\rangle \right\}$$

cuja condição de otimalidade é dada por

$$\tilde{H} = -P_{\tilde{\mathcal{L}}} \tilde{\nabla} f_{\beta}(X)$$

ou seja, a direção de máximo decrescimento sobre os elipsóides centrados em  $X$ , assunto discutido na seção 2.9, o que permite concluir que as iterações na forma

$$X_{k+1} = X_k + \gamma_k H_k$$

são idênticas às efetuadas pelo algoritmo 1.

Note-se que, de maneira equivalente, a direção  $H$  pode ser encontrada por meio do seguinte problema de mínimos quadrados

$$H = \arg \min_{H \in \mathcal{L}} \left\{ \left\| X^{\frac{1}{2}} \nabla f_{\beta}(X) X^{\frac{1}{2}} + X^{-\frac{1}{2}} H X^{-\frac{1}{2}} \right\| \right\}$$

□

Cabe ressaltar que essa equivalência total só existe porque o termo  $(\beta \langle C, X \rangle)$  não contribui para a Hessiana de  $f_{\beta}(X)$ . Outro ponto interessante é que o método de Newton modificado aplicado à minimização da função apenas matricial análoga à função apenas matricial dada por

$$f_{\beta}(x) = \beta c'x + B(x) \tag{3.4}$$

também equivale ao algoritmo 1, o que pode ser facilmente verificado a partir das expressões para o vetor gradiente e a matriz Hessiana fornecidos pelo lema 2.1. Por isso, no decorrer desse capítulo, são apresentadas provas dos diversos resultados apenas para o caso matricial; sempre que a obtenção dos análogos vetoriais requerer manipulações algébricas um pouco mais complexas serão feitas as devidas ressalvas.

### 3.1.3 Medida da centralização

Utilizando-se o algoritmo 1 para encontrar pontos pertencentes à trajetória central da função apenas  $f_{\beta}(X)$ , uma estimativa da distância do ponto obtido na  $k$ ésima iteração,  $X_k$ , ao ponto central  $\bar{X}(\beta)$  pode ser obtida em função do escalar  $\delta(X_k)$ , definido como o resultado da aplicação da seguinte função ao ponto  $X_k$

$$\delta(X) := \|X^{-\frac{1}{2}} H X^{-\frac{1}{2}}\| = \|\tilde{H}\| = \|H\|_{X^{-1}} \tag{3.5}$$

em que  $H$  é a direção de Newton avaliada no ponto  $X$ . O valor de  $\delta(X)$  como medida da centralização pode ser mais bem compreendido por meio do próximo teorema.

**Teorema 3.2** *Dados  $X \in \Phi$  e  $\beta > 0$ , as seguintes afirmações são verdadeiras*

*i) Se  $\delta(X) < 1$ , o ponto  $Y = X + H$  é factível e  $\delta(Y) \leq \delta^2(X)$ .*

ii) Se  $\delta(X) < 1$ , uma estimativa da distância do ponto  $X$  ao ponto pertencente à trajetória central  $\bar{X}(\beta)$  é dada por

$$f_\beta(X) - f_\beta(\bar{X}(\beta)) \leq \frac{\delta^2(X)}{1 - \delta^2(X)} \quad (3.6)$$

iii) Se  $\delta(X) < 1$ , então

$$\|X - \bar{X}(\beta)\|_{X^{-1}} \leq \frac{\delta(X)}{1 - \delta(X)} \quad (3.7)$$

iv) Dado  $\gamma = 1/(1 + \delta(X))$ , o ponto  $Y = X + \gamma H$  é factível e

$$f_\beta(Y) - f_\beta(X) \leq \log(1 + \delta(X)) - \delta(X) \quad (3.8)$$

**Prova:** A prova desse teorema será feita por pontos.

i) Na prova desse ponto, assumir-se-á que  $X = \mathbf{I}$ . Essa imposição não implica perda de generalidade, uma vez que é sempre possível definir um problema equivalente através da transformação simétrica definida pela matriz  $X^{-1}$ , conforme discutido na seção 2.6, a qual mapeia o ponto  $X \in \Phi$  genérico na matriz  $\mathbf{I}$ . Assim, fica claro, segundo o lema 2.5, que o ponto  $Y = X + H = \mathbf{I} + H$  é factível pois, por hipótese,  $\delta(X) = \delta(\mathbf{I}) = \|H\| < 1$ . Além disso  $H = -P_{\mathcal{L}} \nabla f_\beta(\mathbf{I}) = P_{\mathcal{L}}(\mathbf{I} - \beta C)$  e, portanto, o vetor  $\nabla f_\beta(\mathbf{I})$  pode ser escrito como  $\nabla f_\beta(\mathbf{I}) = -H + Z_X$ , em que  $Z_X \in \mathcal{L}^\perp$  é o complemento ortogonal de  $\nabla f_\beta(\mathbf{I})$ . Por sua vez,  $\tilde{H}_Y = -P_{\mathcal{L}} \tilde{\nabla} f_\beta(Y)$ , donde  $\delta(Y) = \|\tilde{H}_Y\|$  é dado por

$$\delta(Y) = \min_{Z_Y \in \mathcal{L}^\perp} \|Y^{\frac{1}{2}} (\nabla f_\beta(Y) - Z_Y) Y^{\frac{1}{2}}\|$$

Em especial, para  $Z_Y = Z_X$

$$\begin{aligned} \delta(Y) &\leq \|Y^{\frac{1}{2}} (\nabla f_\beta(Y) - Z_X) Y^{\frac{1}{2}}\| \\ &= \|Y^{\frac{1}{2}} (\nabla f_\beta(Y) - \nabla f_\beta(\mathbf{I}) - H) Y^{\frac{1}{2}}\| \\ &= \|Y^{\frac{1}{2}} (-Y^{-1} + \mathbf{I} - Y + \mathbf{I}) Y^{\frac{1}{2}}\| \\ &= \|-Y^2 + 2Y - \mathbf{I}\| \\ &= \|(Y - \mathbf{I})^2\| \\ &= \|H^2\| \\ &= \sqrt{\text{tr}(H^2 H^2)} \\ &\leq \sqrt{(\text{tr}(H^2))^2} \\ &= \delta^2(X) \end{aligned}$$

ii) Se  $Y = X + H$  e  $\delta(X) < 1$ , então, devido à convexidade da função  $f_\beta$ ,

$$f_\beta(Y) - f_\beta(X) = f_\beta(X + H) - f_\beta(X) \geq \langle \nabla f_\beta(X), H \rangle$$

Como  $H$  é uma direção de Newton,  $\langle \nabla f_\beta(X), H \rangle = -\langle \nabla^2 f_\beta(X)H, H \rangle = -\delta^2(X)$ , o que permite reescrever a inequação anterior como

$$f_\beta(Y) - f_\beta(X) \geq -\delta^2(X)$$

A expressão final é obtida aplicando-se esse resultado ao algoritmo 1 com  $\gamma = 1$ . Tomando-se  $X = X_k$  e  $Y = X_{k+1}$  e supondo-se que  $\delta(X_0) < 1$ , de tal forma que  $\delta(X_{k+1}) \leq \delta^2(X_k)$ , tem-se

$$f_\beta(X_k) - f_\beta(X_{k+1}) \leq \delta^2(X_k)$$

portanto

$$\begin{aligned} f_\beta(X_0) - f_\beta(X_k) &= f_\beta(X_0) - f_\beta(X_1) + f_\beta(X_1) - f_\beta(X_2) + \dots \\ &\quad \dots - f_\beta(X_{k-1}) + f_\beta(X_{k-1}) - f_\beta(X_k) \\ &\leq \delta^2(X_0) + \delta^2(X_1) + \dots + \delta^2(X_k) \\ &\leq \delta^2(X_0) + \delta^4(X_0) + \dots + \delta^{2^k}(X_0) \\ &\leq \delta^2(X_0) + \delta^4(X_0) + \dots \\ &= \frac{\delta^2(X_0)}{1 - \delta^2(X_0)} \end{aligned}$$

que fornece a estimativa apresentada, uma vez que  $f_\beta(X_k)$  tende a  $f_\beta(\bar{X}(\beta))$  quando  $k$  tende ao infinito.

iii) Assumindo-se novamente que  $X = X_0 = \mathbf{I}$ , deseja-se provar que

$$\|\bar{X}(\beta) - X_0\| \leq \delta(X_0) + \delta^2(X_0) + \dots$$

Inicialmente, provar-se-á, por indução, que

$$\|X_k - X_0\| \leq \delta(X_0) + \delta^2(X_0) + \dots + \delta^{2^k-1}(X_0)$$

em que  $X_k$  é o ponto gerado pelo algoritmo 1 com  $\gamma = 1$  na  $k$ ésima iteração. Para  $k = 1$

$$\|X_1 - \mathbf{I}\| = \|H_0\| = \delta(X_0)$$

Na  $k$ ésima iteração  $X_{k+1} - X_k = H_k$ ,  $X_k > 0$  e  $\delta(X_k) < 1$ , portanto

$$\begin{aligned} \|X_{k+1} - X_k\| &= \|H_k\| \\ &= \|X_k^{\frac{1}{2}} \tilde{H}_k X_k^{\frac{1}{2}}\| \\ &= \text{tr} \left( X_k^{\frac{1}{2}} \tilde{H}_k X_k \tilde{H}_k X_k^{\frac{1}{2}} \right)^{\frac{1}{2}} \\ &\leq \lambda_M^{\frac{1}{2}}(X_k) \text{tr} \left( \tilde{H}_k X_k \tilde{H}_k \right)^{\frac{1}{2}} \\ &\leq \lambda_M(X_k) \text{tr} \left( \tilde{H}_k^2 \right)^{\frac{1}{2}} \\ &= \lambda_M(X_k) \|\tilde{H}_k\| \\ &\leq \lambda_M(X_k) \delta^{2^k}(X_0) \end{aligned}$$

Entretanto, associando-se à hipótese da indução que  $\lambda_M(X_k - \mathbf{I}) = \lambda_M(X_k) - 1 \leq \|X_k - X_0\|$ , obtém-se

$$\lambda_M(X_k) \leq 1 + \delta(X_0) + \delta^2(X_0) + \cdots + \delta^{2^k-1}(X_0)$$

Unindo-se esses dois últimos resultados chega-se a

$$\|X_{k+1} - X_k\| \leq \delta^{2^k}(X_0) + \delta^{2^{k+1}}(X_0) + \cdots + \delta^{2^{k+1}-1}(X_0)$$

que verifica a hipótese para  $k + 1$  já que

$$\begin{aligned} \|X_{k+1} - X_0\| &\leq \|X_{k+1} - X_k\| + \|X_k - X_0\| \\ &\leq \delta(X_0) + \delta^2(X_0) + \cdots + \delta^{2^{k+1}-1}(X_0) \end{aligned}$$

Finalmente, como  $X_k$  é um ponto arbitrariamente próximo a  $\bar{X}(\beta)$ , então

$$\begin{aligned} \|\bar{X}(\beta) - X_0\| &\leq \delta(X_0) + \delta^2(X_0) + \cdots \\ &\leq \frac{\delta(X_0)}{1 - \delta(X_0)} \end{aligned}$$

iv) Utilizando-se o lema 2.7 obtém-se, para  $Y = X + \gamma H$ ,

$$\begin{aligned} f_\beta(Y) - f_\beta(X) &= f_\beta(X + \gamma H) - f_\beta(X) \\ &= \gamma \langle \beta C, H \rangle - \sum_{i=1}^p \log [\mathbf{I} + \gamma \lambda_i(\tilde{H})] \\ &= \gamma \langle \nabla f_\beta(X), H \rangle + \gamma \langle X^{-1}, H \rangle - \sum_{i=1}^p \log [\mathbf{I} + \gamma \lambda_i(\tilde{H})] \\ &= -\gamma \delta^2(X) + \sum_{i=1}^p \gamma \lambda_i(\tilde{H}) - \sum_{i=1}^p \log [\mathbf{I} + \gamma \lambda_i(\tilde{H})] \end{aligned}$$

Expandindo-se cada termo logarítmico em série de potências tem-se que

$$\begin{aligned} f_\beta(Y) - f_\beta(X) &= -\gamma \delta^2(X) + \sum_{i=1}^p \gamma \lambda_i(\tilde{H}) - \\ &\quad - \sum_{i=1}^p \gamma \lambda_i(\tilde{H}) + \sum_{k=2}^{\infty} \frac{(-1)^k}{k} \sum_{i=1}^p (\gamma \lambda_i(\tilde{H}))^k \\ &\leq -\gamma \delta^2(X) + \sum_{k=2}^{\infty} \frac{1}{k} \left( \sum_{i=1}^p (\gamma \lambda_i(\tilde{H}))^2 \right)^{k/2} \\ &\leq -\gamma \delta(X) [1 + \delta(X)] - \log(1 - \gamma \delta(X)) \end{aligned}$$

Verifica-se que  $\gamma = 1/(1 + \delta(X))$  maximiza a desigualdade anterior e provê

$$f_\beta(Y) - f_\beta(X) \leq -\delta(X) + \log(1 + \delta(X))$$

Para completar a prova basta verificar que  $\|\gamma \tilde{H}\| = \gamma \delta(X) = \delta(X)/(1 + \delta(X)) < 1$  o que, segundo o lema 2.5, mantém  $Y$  factível. Note-se, também, que essa mesma condição é que permite efetuar a expansão em série de potências dos termos logarítmicos.

□

Baseado nos resultados desse teorema é possível formula-se a seguinte estratégia de escolha do passo  $\gamma_k$  para o algoritmo 1 aplicado à função apenas

$$\gamma_k := \begin{cases} 1, & \text{se } \delta(X_k) < 1 \\ 1/(1 + \delta(X_k)), & \text{se } \delta(X_k) \geq 1 \end{cases} \quad (3.9)$$

que assegura convergência quadrática do algoritmo na região em que  $\delta(X) < 1$ . Pontos nessa região são chamados pontos *quase centrais*. Fora dessa região, o valor da função decresce conforme o ponto *iv*).

Assim como para os pontos centrais, é possível associar a pontos quase centrais pontos duais factíveis. Esse assunto é discutido no próximo lema.

**Lema 3.3** *Dados o parâmetro  $\beta > 0$  e o ponto quase central  $X$ , o ponto*

$$Z = X^{-\frac{1}{2}} (\mathbf{I} - \tilde{H}) X^{-\frac{1}{2}} / \beta \quad (3.10)$$

em que  $\tilde{H}$  é a direção de Newton avaliada no ponto  $X$  e transformada por  $X^{-1}$ , é um ponto dual factível e o gap de dualidade associado ao par  $(X, Z)$  é

$$\eta(X, \beta) = \frac{p - \langle \mathbf{I}, \tilde{H} \rangle}{\beta} \leq \frac{p + \delta(X)\sqrt{p}}{\beta} \quad (3.11)$$

**Prova:** Pelo lema 2.5,  $Z > 0$ , pois  $\delta(X) = \|\tilde{H}\| < 1$ . Por sua vez, pelo lema 2.4,  $Z$  será factível se, e somente se, a matriz

$$\tilde{Z} = X^{\frac{1}{2}} Z X^{\frac{1}{2}} = (\mathbf{I} - \tilde{H}) / \beta$$

for factível em relação ao conjunto dual no espaço transformado por  $X^{-1}$ . Como

$$P_{\tilde{Z}} \tilde{Z} = (P_{\tilde{Z}} \mathbf{I} - P_{\tilde{Z}} \tilde{H}) / \beta$$

e  $P_{\tilde{Z}} \tilde{H} = \tilde{H} = -\beta P_{\tilde{Z}} \tilde{C} + P_{\tilde{Z}} \mathbf{I}$ , então  $P_{\tilde{Z}} \tilde{Z} = P_{\tilde{Z}} \tilde{C}$ , o que permite concluir, pelo lema 2.3, que  $Z$  é factível em relação ao problema dual. Isso permite calcular o gap de dualidade

$$\eta(X, \beta) = \langle X, Z \rangle = \langle \mathbf{I}, \tilde{Z} \rangle = (p - \langle \mathbf{I}, \tilde{H} \rangle) / \beta \leq (p + \delta(X)\sqrt{p}) / \beta$$

pois  $|\langle \mathbf{I}, \tilde{H} \rangle| \leq \|\mathbf{I}\| \|\tilde{H}\|$ . □

Em termos práticos, a noção de ponto quase central é formalizada pela regra  $\delta(X) \leq \varepsilon$ , para valores de  $\varepsilon$  arbitrários, desde que menores do que um. Nesse trabalho, um ponto será considerado quase central se  $\delta(X) < \varepsilon = 0.5$ .

### 3.1.4 O passo ótimo

Apesar da importância dos resultados do teorema 3.2 na análise de convergência dos métodos de pontos interiores, é desejável que a seleção do passo  $\gamma_k$  seja feita de modo a minimizar a função  $f_\beta(X_k + \gamma_k H_k)$ , assunto do próximo teorema.

**Teorema 3.4** *Dados  $X \in \Phi$ , a matriz  $\Psi(H) = \tilde{H}$ , correspondente à direção de Newton  $H$  avalida no ponto  $X$ , e  $\delta(X)$ , o valor do escalar  $\gamma$  que minimiza a função  $f_\beta(X + \gamma H)$  é o inverso do maior autovalor positivo da matriz  $\Gamma$  dada por*

$$\Gamma := \frac{1}{\delta^2(X)} \Lambda \Lambda' - \text{diag}(\Lambda) \quad (3.12)$$

em que  $\Lambda = [ \lambda_1(\Psi(H)) \ \cdots \ \lambda_p(\Psi(H)) ]$ . Essa escolha de  $\gamma$  será denotada por  $\gamma^*$ .

**Prova:** De acordo com o lema 2.7, a função  $f_\beta(X + \gamma H)$  é dada por

$$f_\beta(X + \gamma H) = f_\beta(X) - \sum_{i=1}^p \log [1 + \gamma \lambda_i(\Psi(H))] + \gamma \langle \beta C, H \rangle$$

cuja derivada em relação a  $\gamma$  vale

$$\frac{df_\beta(X + \gamma H)}{d\gamma} = - \sum_{i=1}^p \frac{\lambda_i(\Psi(H))}{1 + \gamma \lambda_i(\Psi(H))} + \langle \beta C, H \rangle$$

Definindo-se  $\mu := 1/\gamma$  e igualando-se a expressão anterior a zero em busca de otimalidade, obtém-se

$$\sum_{i=1}^p \frac{\lambda_i^2(\Psi(H))}{\mu + \lambda_i(\Psi(H))} - \sum_{i=1}^p \lambda_i(\Psi(H)) + \langle \beta C, H \rangle = 0$$

que, mediante o emprego da expressão da derivada direcional primeira de  $f_\beta(X)$  na direção de Newton  $H$

$$\langle \nabla f_\beta(X), H \rangle = - \sum_{i=1}^p \lambda_i(\Psi(H)) + \langle \beta C, H \rangle = - \langle \nabla^2 f_\beta(X) H, H \rangle = -\delta^2(X) < 0$$

torna-se equivalente a

$$\frac{1}{\delta^2(X)} \sum_{i=1}^p \frac{\lambda_i^2(\Psi(H))}{\mu + \lambda_i(\Psi(H))} - 1 = \frac{1}{\delta^2(X)} \Lambda' [\mu \mathbf{I} + \text{diag}(\Lambda)]^{-1} \Lambda - 1 = 0 \quad (3.13)$$

Assim, a matrix  $\Gamma$  pode ser obtida aplicando-se a essa última relação uma conhecida propriedade do determinante que permite escrever

$$\det \left( \frac{1}{\delta^2(X)} \Lambda \Lambda' [\mu \mathbf{I} + \text{diag}(\Lambda)]^{-1} - \mathbf{I} \right) = 0$$

ou, do modo como apresentado no enunciado do teorema,

$$\det \left( \mu \mathbf{I} + \text{diag}(\Lambda) - \frac{1}{\delta^2(X)} \Lambda \Lambda' \right) = \det(\mu \mathbf{I} - \Gamma) = 0$$

A questão “qual autovalor de  $\Gamma$  corresponde ao passo desejado?” pode ser facilmente respondida com base na seguinte análise: como  $\delta^2(X)$  é positivo, a derivada direcional primeira tem sinal negativo, logo, como  $f_\beta(X)$  é estritamente convexa, o passo correto  $\gamma^*$  deve ser o menor valor positivo possível, ou seja, o inverso do maior autovalor positivo de  $\Gamma$ .

É interessante observar que  $\Gamma = -\text{diag}(\Lambda) + \Lambda\Lambda'/\delta^2(X) \geq -\text{diag}(\Lambda)$  donde  $\gamma^* = 1/\lambda_M(\Gamma) \leq -1/\lambda_m(\Psi(H))$ , portanto,  $X + \gamma^*H \in \Phi$ , uma vez que  $\gamma^*$  obedece aos limitantes impostos pelo lema 2.6.  $\square$

É ainda possível formular o seguinte corolário.

**Corolário 3.5** *Mantidas as condições de validade do teorema 3.4,*

$$\gamma^+ = \frac{1}{1 - \lambda_m(\tilde{H})} \quad (3.14)$$

*é um valor de passo sub-ótimo. Entende-se por passo sub-ótimo o seguinte fato*

$$f_\beta(X) \geq f_\beta(X + \gamma^+H) \geq f_\beta(X + \gamma^*H) \quad (3.15)$$

**Prova:** Segundo o teorema 3.4,

$$\begin{aligned} \gamma^* &= \frac{1}{\lambda_M(\Gamma)} \\ &\geq \frac{1}{\lambda_M(\frac{1}{\delta^2(X)}\Lambda\Lambda') + \lambda_M(-\text{diag}(\Lambda))} \\ &= \frac{1}{\frac{1}{\delta^2(X)} \sum_{i=1}^p \lambda_i^2(\tilde{H}) - \lambda_m(\tilde{H})} \end{aligned}$$

Porém, segundo o lema 2.7,

$$\delta^2(X) = D^2 B_{\mathcal{K}}(X, H) = \sum_{i=1}^p \lambda_i^2(\tilde{H})$$

o que fornece

$$\gamma^+ = \frac{1}{1 - \lambda_m(\tilde{H})} \leq \gamma^*$$

$\square$

Note-se que a seguinte relação é sempre verdadeira

$$\frac{1}{1 + \delta(X)} \leq \frac{1}{1 + \max_i |\lambda_i(\tilde{H})|} \leq \gamma^+ \leq \gamma^* \quad (3.16)$$

e os análogos vetoriais do teorema 3.4 e do corolário 3.1.4 são imediatos, bastando substituir-se  $\Psi(H)$  por  $\Psi(h)$ , conforme definido no lema 2.1.

## 3.2 Algoritmos seguidores de trajetória

Fazendo uso das propriedades discutidas até o momento, inicia-se nesta seção com os algoritmos *seguidores de trajetória* o estudo de algumas classes de métodos para a resolução do problema (2.26). A idéia básica por detrás dos algoritmos seguidores de trajetória consiste na escolha de funções que definam centros analíticos segundo um parâmetro proporcional ao objetivo primal, ao objetivo dual ou ao objetivo primal-dual. A partir dessas funções, resolve-se uma série de problemas de centralização segundo uma seqüência de atualização dos parâmetros livres que reduza de maneira uniforme o *gap* de dualidade. Inicia-se pela parametrização primal-dual.

### 3.2.1 Parametrização primal-dual

A função que parametriza centros analíticos com o parâmetro livre proporcional ao *gap* de dualidade definido para pontos centrais é a função apenas  $f_\beta(X)$ , introduzida no início deste capítulo. Na seção 3.1.2, há uma discussão acerca da variação do parâmetro  $\beta$  e sua relação direta com o *gap* de dualidade  $\eta(\beta)$ . Esse resultado é o ponto chave na construção dos métodos seguidores de trajetória e será repetido aqui.

Um ponto  $\bar{X}(\beta)$  pertencente à trajetória central definida pela função apenas está associado a uma matriz dual factível  $\bar{Z}(\beta) = \bar{X}^{-1}(\beta)/\beta$ . Esse par de matrizes define um *gap* de dualidade  $\eta(\beta) = p/\beta$ . Assim, escolhido um valor para o parâmetro  $\theta \in (0, 1)$ , a seqüência de pontos  $\bar{X}(\beta_l)$  em que

$$\beta_{l+1} = \frac{\beta_l}{\theta} \quad (3.17)$$

converge para a solução global de (2.26) pois

$$\eta(\beta_{l+1}) = \theta\eta(\beta_l) \quad (3.18)$$

converge para zero quando  $l$  tende ao infinito.

### 3.2.2 Parametrização primal

A segunda parametrização possível está associada à função *central*

$$f_\lambda(X) = -q \log(\lambda - \langle C, X \rangle) + B_{\mathcal{K}}(X) \quad (3.19)$$

introduzida por Huard em [11], em que o valor de  $q$  deve ser feito sempre maior ou igual a  $p$  (a dimensão da LMI) e o parâmetro  $\lambda$  representa um limite superior para o objetivo primal, daí o nome parametrização primal. Essa função é estritamente convexa e apresenta, dado um valor fixo a  $\lambda$ , um ponto de mínimo  $\bar{X}(\lambda)$ . O gradiente de  $f_\lambda(X)$  é dado por

$$\nabla f_\lambda(X) = \frac{q}{\lambda - \langle C, X \rangle} C - X^{-1} \quad (3.20)$$

e permite caracterizar o ponto  $\bar{X}(\lambda)$  mediante

$$P_{\mathcal{L}} \nabla f_\lambda(\bar{X}(\lambda)) = 0 \rightarrow P_{\mathcal{L}} \bar{X}^{-1}(\lambda) = \frac{q}{\lambda - \langle C, \bar{X}(\lambda) \rangle} P_{\mathcal{L}} C \quad (3.21)$$

Novamente, como na seção 3.1.2, o ponto  $\bar{X}(\lambda)$  pertence à trajetória central e pode ser relacionado com o parâmetro primal-dual via  $\beta = q / (\lambda - \langle C, \bar{X}(\lambda) \rangle)$  e, conseqüentemente, com o *gap* de dualidade  $\eta(\lambda) = p (\lambda - \langle C, \bar{X}(\lambda) \rangle) / q$ . Porém, nesse caso, como  $\eta(\lambda)$  é também função de  $\bar{X}(\lambda)$ , uma regra de atualização inspirada na parametrização primal-dual, isto é

$$\lambda_{l+1} = \theta \lambda_l + (1 - \theta) \langle C, \bar{X}(\lambda_l) \rangle \quad (3.22)$$

implica

$$\eta(\lambda_{l+1}) > \theta \eta(\lambda_l) \quad (3.23)$$

pois  $\eta(\lambda_{l+1}) = p (\lambda_{l+1} - \langle C, \bar{X}(\lambda_{l+1}) \rangle) / q > p (\lambda_{l+1} - \langle C, \bar{X}(\lambda_l) \rangle) / q$ . Todavia, embora a relação (3.23) não garanta um decrescimento uniforme do *gap* de dualidade, essa atualização é ainda “produtiva”, no sentido do próximo lema.

**Lema 3.6** *Dados a regra de atualização (3.22), o ponto central  $\bar{X}(\lambda_l)$  e o ponto solução do problema (2.26),  $X^*$ , então*

$$\frac{\lambda_{l+1} - \langle C, X^* \rangle}{\lambda_l - \langle C, X^* \rangle} \leq \frac{p + q\theta}{p + q} \leq \frac{1 + \theta}{2} \quad (3.24)$$

e

$$\eta(\lambda_l) \leq \lambda_l - \langle C, X^* \rangle \leq \frac{q + p}{p} \eta(\lambda_l) \quad (3.25)$$

**Prova:** Definindo-se  $d = \langle C, \bar{X}(\lambda_l) \rangle - \eta(\lambda_l) \leq \langle C, X^* \rangle$ ,

$$\lambda_l - d = \lambda_l - \langle C, \bar{X}(\lambda_l) \rangle + p (\lambda_l - \langle C, \bar{X}(\lambda_l) \rangle) / q$$

Por sua vez, utilizando-se a regra de atualização (3.22), tem-se

$$\lambda_{l+1} - d = \theta (\lambda_l - \langle C, \bar{X}(\lambda_l) \rangle) + p (\lambda_l - \langle C, \bar{X}(\lambda_l) \rangle) / q$$

Sabendo-se que  $q \geq p$  e utilizando-se as expressões acima obtém-se

$$\begin{aligned} \frac{\lambda_{l+1} - \langle C, X^* \rangle}{\lambda_l - \langle C, X^* \rangle} &= \frac{(\lambda_{l+1} - d) - (\langle C, X^* \rangle - d)}{(\lambda_l - d) - (\langle C, X^* \rangle - d)} \\ &= \frac{\frac{p+q\theta}{p+q}(\lambda_l - d) - (\langle C, X^* \rangle - d)}{(\lambda_l - d) - (\langle C, X^* \rangle - d)} \\ &\leq \frac{p + q\theta}{p + q} \\ &\leq \frac{1 + \theta}{2} \end{aligned}$$

pois  $(\langle C, X^* \rangle - d) > 0$  e  $(p + q\theta)/(p + q) \in (0, 1)$ .

A segunda desigualdade pode ser provada observando-se que  $q \geq p$  implica

$$\eta(\lambda_l) = p (\lambda_l - \langle C, \bar{X}(\lambda_l) \rangle) / q \leq \lambda_l - \langle C, X^* \rangle$$

e, portanto, o termo à direita está sujeito a

$$\lambda_l - \langle C, X^* \rangle \leq \lambda_l - \langle C, \bar{X}(\lambda_l) \rangle + \eta(\lambda_l) = \frac{q}{p} \eta(\lambda_l) + \eta(\lambda_l)$$

□

Assim, apesar dessa regra de atualização do parâmetro  $\lambda$  prover uma redução no *gap* de dualidade menor do que a obtida no caso da parametrização primal-dual, ainda assim há uma efetiva redução do objetivo primal a cada etapa de centralização. Note-se também que o decrescimento do objetivo primal pode ser colocado em função do parâmetro  $q$ .

De maneira análoga, a função vetorial

$$f_\lambda(x) = -q \log(\lambda - c'x) + B(x) \quad (3.26)$$

possui as mesmas propriedades descritas para o caso matricial.

Resta saber como calcular os pontos centrais  $\bar{X}(\lambda)$ . Para a função apenada, provou-se que o algoritmo 1 é completamente equivalente ao método de Newton. A função central, no entanto, possui um termo não linear além da barreira  $B_{\mathcal{K}}(X)$  que contribui para a Hessiana, donde não se pode aplicá-lo diretamente à função  $f_\lambda(X)$  sem que haja alguma degradação no desempenho<sup>1</sup>. Entretanto, é possível restabelecer essa equivalência definindo-se o subespaço  $\mathcal{L}_c$  associado à LMI

$$F_c(x) = \begin{bmatrix} F(x) & \mathbf{0} \\ \mathbf{0} & \text{diag}([\lambda - c'x] \cdots [\lambda - c'x]) \end{bmatrix} \quad (3.27)$$

em que a matriz diagonal extra contém  $q$  cópias do termo  $(\lambda - c'x)$ . Desse modo, a função central  $f_\lambda(x)$  torna-se equivalente à função de barreira (2.9) associada à LMI  $\{F_c(x) \geq 0\}$ . No caso matricial, isso significa que o algoritmo 1 apresenta total equivalência com o método de Newton para a função  $f_\lambda(X)$ , desde que as projeções no subespaço  $\mathcal{L}$  sejam substituídas por projeções no subespaço  $\mathcal{L}_c$ .

### 3.2.3 Parametrização dual

Finalmente, a terceira e última parametrização está associada à função *potencial*

$$f_\rho(X) = q \log(\langle C, X \rangle - \rho) + B_{\mathcal{K}}(X) \quad (3.28)$$

introduzida por Karmakar em [12], em que o valor de  $q$  deve ser feito sempre maior ou igual a  $p$  e o parâmetro  $\rho$  representa um limite inferior para o objetivo primal, ou seja, um objetivo dual factível. Essa função é o resultado da soma de uma função convexa, a barreira, com um termo côncavo. Como o termo côncavo é uma função monotônica que tende ao infinito apenas se  $\langle C, X \rangle = \rho$ ,  $f_\rho(X)$  sempre apresenta um ponto de mínimo  $\bar{X}(\rho)$  pertencente a  $\Phi$ . Como para as demais parametrizações, esse ponto pode ser caracterizado mediante a determinação da projeção do gradiente

$$\nabla f_\rho(X) = \frac{q}{\langle C, X \rangle - \rho} C - X^{-1} \quad (3.29)$$

no subespaço  $\mathcal{L}$ , o que implica

$$P_{\mathcal{L}} \nabla f_\rho(\bar{X}(\rho)) = 0 \rightarrow P_{\mathcal{L}} \bar{X}^{-1}(\rho) = \frac{q}{\langle C, \bar{X}(\rho) \rangle - \rho} P_{\mathcal{L}} C \quad (3.30)$$

Logo, o ponto  $\bar{X}(\rho)$  pertence à trajetória central de  $\Phi$  e está relacionado ao parâmetro primal-dual via  $\beta = q / (\langle C, \bar{X}(\rho) \rangle - \rho)$  e ao *gap* de dualidade  $\eta(\rho) = p(\rho - \langle C, \bar{X}(\rho) \rangle) / q$ .

<sup>1</sup>Veja mais adiante a discussão sobre a função potencial.

Assim como na parametrização primal,  $\eta(\rho)$  é também função de  $\bar{X}(\rho)$  e constrói-se a regra de atualização de  $\rho$

$$\rho_{l+1} = \theta\rho_l + (1 - \theta) \langle C, \bar{X}(\rho_l) \rangle \quad (3.31)$$

Essa regra implica

$$\eta(\rho_{l+1}) < \theta\eta(\rho_l) \quad (3.32)$$

ou seja, a redução do *gap* é maior do que a obtida pela parametrização primal-dual. Isso pode ser justificado por

$$\eta(\rho_{l+1}) = p (\langle C, \bar{X}(\rho_{l+1}) \rangle - \rho_{l+1}) / q < p (\langle C, \bar{X}(\rho_l) \rangle - \rho_{l+1}) / q \quad (3.33)$$

Há, no entanto, a exigência de que o valor de  $\theta$  gere um parâmetro  $\rho_{l+1}$  que ainda seja um objetivo dual factível. Verifica-se que isso sempre ocorre caso  $\theta \in (n/q, 1)$ . Se for possível conhecer o valor ótimo do objetivo, pode-se *estender* o valor de  $\theta$  além destes limites, respeitando-se o máximo valor possível para  $\theta$  dado por

$$(\langle C, \bar{X}(\rho_l) \rangle - \langle C, X^* \rangle) / (\langle C, \bar{X}(\rho_l) \rangle - \rho_l) \quad (3.34)$$

Isso ocorre, por exemplo, quando se considera como objetivo o *gap* de dualidade, assunto que será tratado mais adiante na seção 3.3. Em alguns casos, pode-se ainda utilizar uma estimativa do valor ótimo, representada por um objetivo dual factível. Mais detalhes sobre esse assunto podem ser encontrados em Nemirovskii & Nesterov [14] na seção 4.3.

Quanto à maneira de se calcular o ponto central associado ao parâmetro dual, utilizar-se-á novamente o algoritmo 1. Neste caso, porém, não há como escapar da influência do termo não linear sobre a Hessiana e o algoritmo 1 se transforma num algoritmo do tipo *quase* Newton, isto é, a direção  $H_k$  é calculada como a direção de Newton da seguinte função

$$fl_\rho(X) = q \log(\langle C, X_k \rangle - \rho) + \langle \nabla V(X_k), X - X_k \rangle + B_{\mathcal{K}}(X) \quad (3.35)$$

em que  $\nabla V(X_k)$  representa o valor do gradiente do termo côncavo  $V(X) = q \log(\langle C, X \rangle - \rho)$  avaliado no ponto  $X_k$ . Isso significa que o valor da Hessiana de  $fl_\rho(X)$  é igual ao valor da Hessiana da função de barreira  $B_{\mathcal{K}}(X)$ , e que a equivalência com o conceito de projeção no espaço transformado é novamente completa. Note-se que a convergência desse procedimento está assegurada pela relação  $f_\rho(X) \leq fl_\rho(X)$ . Assim, minimizando-se  $fl_\rho(X)$ , cujas propriedades de convergência são garantidas pelo teorema 3.2, obtém-se um decréscimo ainda maior da função potencial  $f_\rho(X)$ . Esse assunto será tratado com mais detalhes na seção 3.3.

### 3.2.4 O algoritmo seguidor de trajetória

Uma vez introduzidas as três parametrizações de interesse, formalizar-se-á o algoritmo seguidor de trajetória. Utilizar-se-á o parâmetro  $\alpha$  a fim de denotar de maneira genérica qualquer um dos três possíveis parâmetros,  $\beta$ ,  $\lambda$  ou  $\rho$ .

**Algoritmo 2** *Dados o parâmetro  $\alpha_0$  associado ao ponto pertencente à trajetória central  $X_0 = \bar{X}(\alpha_0)$  e  $\epsilon > 0$*

1. Atualizar o parâmetro  $\alpha$  segundo

$$\begin{cases} \text{Primal-dual:} & \beta_{l+1} = \beta_l/\theta, & \theta \in (0, 1) \\ \text{Primal:} & \lambda_{l+1} = \theta\lambda_l + (1-\theta)\langle C, X_l \rangle, & \theta \in (0, 1) \\ \text{Dual:} & \rho_{l+1} = \theta\rho_l + (1-\theta)\langle C, X_l \rangle, & \theta \in (n/q, 1) \end{cases} \quad (3.36)$$

2. Calcular  $X_{l+1} = \bar{X}(\alpha_{l+1})$

3. Parar se  $\eta(\alpha_{l+1}) < \epsilon$ . Caso contrário, incrementar  $l$  e voltar ao primeiro passo.

□

Vale a pena discutir um pouco mais o critério de parada. É desejável que o algoritmo acima seja interrompido em um ponto cujo valor do objetivo esteja próximo do valor ótimo. Mais formalmente falando, para um  $\epsilon > 0$

$$\langle C, X_l \rangle - \langle C, X^* \rangle < \epsilon \quad (3.37)$$

Como geralmente não é possível saber de antemão o valor do objetivo ótimo, utiliza-se um critério indireto de parada, representado pelo valor do *gap* de dualidade associado a pontos pertencentes à trajetória central. Para todo  $X_l$  pertencente à trajetória central  $\eta(X_l) \geq \langle C, X_l \rangle - \langle C, X^* \rangle$  e, para cada uma das três parametrizações introduzidas, esse critério de parada se traduz em

$$\begin{cases} \text{Primal-dual:} & \beta_l > p/\epsilon \\ \text{Primal:} & \lambda_l - \langle C, X_l \rangle < q\epsilon/p \\ \text{Dual:} & \langle C, X_l \rangle - \rho_l < q\epsilon/p \end{cases} \quad (3.38)$$

No próximo lema obtém-se uma estimativa do número de iterações desse algoritmo. A análise leva em conta apenas o número de iterações medidas pelo índice  $l$  sem avaliar o número de iterações gastos para se efetuar o passo 2 (cálculo do ponto central) já que esse assunto receberá um tratamento à parte mais adiante.

**Lema 3.7** *Se o gap de dualidade associado ao ponto central inicial  $X_0 = \bar{X}(\alpha_0)$  for tal que  $\eta(\alpha_0) < 1/\epsilon$ ,  $\epsilon < 1$ , então o número de iterações do algoritmo 2 é limitado por  $O(\log(\epsilon)/\log(\tau))$ . Para as parametrizações primal-dual e dual  $\tau := \theta$  e para a parametrização primal  $\tau := (1 + \theta)/2$ .*

**Prova:** A prova desse lema provém diretamente das discussões sobre as regras de atualização feitas nas seções dedicadas a cada uma das três parametrizações. Para as parametrizações primal-dual e dual a seguinte relação é sempre válida

$$\eta(\alpha_{l+1}) \leq \theta\eta(\alpha_l)$$

donde, na iteração de número  $l$ , impondo-se  $\tau = \theta < 1$ , tem-se

$$\eta(\alpha_l) \leq \tau^l \eta(\alpha_0) \leq \frac{\tau^l}{\epsilon}$$

Como o algoritmo não é interrompido enquanto

$$\frac{\tau^l}{\epsilon} \geq \eta(\alpha_l) > \epsilon$$

pode-se obter um limitante para  $l$  aplicando-se a essa última expressão a função logarítmica, o que resulta em  $l \log(\tau) - \log(\epsilon) \geq \log(\epsilon)$ , ou

$$l \geq 2 \frac{\log(\epsilon)}{\log(\theta)} = O(\log(\epsilon)/\log(\theta))$$

Para a parametrização primal, pelo lema 3.6,

$$\lambda_{l+1} - \langle C, X^* \rangle \leq \frac{1+\theta}{2} (\lambda_{l+1} - \langle C, X^* \rangle)$$

donde, na  $l$ ésima iteração, impondo-se  $\tau = (1+\theta)/2 < 1$  e utilizando-se a segunda parte do lema 3.6, tem-se

$$\eta(\lambda_l) \leq \frac{q+p}{p} \tau^l \eta(\alpha_0) \leq \frac{q+p}{p} \frac{\tau^l}{\epsilon}$$

que pode ser desenvolvida segundo o mesmo raciocínio empregado anteriormente, observando-se que  $O(q) = p$ , o que conclui a prova.  $\square$

Conforme discutido nas seções anteriores, os pontos centrais são calculados por meio do algoritmo 1. Como há sempre a possibilidade de se transformar os problemas formulados com o auxílio das parametrizações primal e dual em problemas de minimização de uma função apenas equivalente, o estudo da complexidade do algoritmo 1 será efetuado apenas em relação a essa função. Resultados semelhantes para as demais parametrizações podem ser derivados de maneira análoga. O próximo lema baseia-se inteiramente no teorema 3.2 e é capaz de fornecer valiosas informações.

**Lema 3.8** *Dados o parâmetro  $\beta_l$ , um ponto quase central  $X(\beta_l) \in \Phi^0$  de tal forma que  $\delta(X(\beta_l)) \leq 0.5$  e a regra de atualização  $\beta_{l+1} = (1+\mu)\beta_l$ , com  $\mu > 0$ , o algoritmo 1 é capaz de encontrar um ponto  $Y(\beta_{l+1})$  de tal forma que  $\delta(Y(\beta_{l+1})) \leq 0.5$*

- i) em apenas uma iteração caso  $\mu < 0.1/\sqrt{p}$  ou*
- ii) em no máximo  $O(p\mu^2/(1+\mu))$  iterações, caso contrário.*

**Prova:** A prova será feita por pontos.

- i) Assume-se, mais uma vez, que  $X(\beta_l) = \mathbf{I}$ . Por hipótese,  $\delta(X(\beta_l)) = \|\beta_l P_{\mathcal{L}} C - P_{\mathcal{L}} \mathbf{I}\| \leq 0.5$ , donde*

$$\beta_l \|P_{\mathcal{L}} C\| \leq \|P_{\mathcal{L}} \mathbf{I}\| + 0.5 \leq \sqrt{p} + 0.5 \leq 1.5\sqrt{p} \quad (3.39)$$

Assim, dados  $X_0 = X(\beta_l) = \mathbf{I}$ ,  $\beta_{l+1} = (1+\mu)\beta_l$  e  $\gamma = 1$ , a primeira iteração do algoritmo 1 tem a forma

$$X(\beta_{l+1}) = X_1 = X_0 + H = X_0 - P_{\mathcal{L}}(\beta_{l+1} C - \mathbf{I})$$

e o valor de  $\delta(X_0)$  é dado por

$$\delta(X_0) = \|H\| = \|\beta_l P_{\mathcal{L}} C - P_{\mathcal{L}} \mathbf{I} + \mu \beta_l P_{\mathcal{L}}\| \leq \|\beta_l P_{\mathcal{L}} C - P_{\mathcal{L}} \mathbf{I}\| + \mu \beta_l \|P_{\mathcal{L}} C\|$$

Substituindo-se  $\mu < 0.1/\sqrt{p}$  juntamente com (3.39), tem-se que

$$\delta(X_0) \leq \delta(X(\beta_l)) + \frac{0.1}{\sqrt{p}} \beta_l \|P_{\mathcal{L}} C\| \leq 0.5 + 0.15 \leq 0.7$$

portanto, aplicando-se o resultado do item *i*) do teorema 3.2

$$\delta(Y(\beta_{l+1})) \leq \delta^2(X_0) \leq (0.7)^2 < 0.5$$

*ii*) Segundo o item *iv*) do teorema 3.2, se  $\delta(X_k) > 0.5$ , então  $f_{\beta}(X_{k+1}) - f_{\beta}(X_k) \leq - (0.5 - \log(1.5)) \leq -0.094$  e

$$k \geq \frac{f_{\beta}(X_0) - f_{\beta}(X_k)}{0.094} \leq \bar{k} = 11 (f_{\beta}(X_0) - f_{\beta}(\bar{X}(\beta)))$$

Assim, o algoritmo 1 gera um ponto  $Y(\beta_{l+1}) = X_k$  que satisfaz à relação  $\delta(Y(\beta_{l+1})) < 0.5$  em no máximo  $\bar{k} = O(f_{\beta}(X_0) - f_{\beta}(\bar{X}(\beta)))$  iterações. Assim, impondo-se  $X_0 = X(\beta_l)$ , pode-se calcular um limitante para o número de iterações a partir de

$$f_{\beta_{l+1}}(X_0) - f_{\beta_{l+1}}(\bar{X}(\beta_{l+1})) = \Delta_1 + \Delta_2$$

em que  $\Delta_1 = f_{\beta_{l+1}}(X_0) - f_{\beta_{l+1}}(\bar{X}(\beta_l))$  e  $\Delta_2 = f_{\beta_{l+1}}(\bar{X}(\beta_l)) - f_{\beta_{l+1}}(\bar{X}(\beta_{l+1}))$ .

Calculando-se  $\Delta_1$

$$\begin{aligned} \Delta_1 &= f_{\beta_{l+1}}(X_0) - f_{\beta_{l+1}}(\bar{X}(\beta_l)) \\ &= f_{\beta_l}(X_0) - f_{\beta_l}(\bar{X}(\beta_l)) + \mu \beta_l \langle C, X_0 - \bar{X}(\beta_l) \rangle \\ &\leq f_{\beta_l}(X_0) - f_{\beta_l}(\bar{X}(\beta_l)) + \mu \beta_l \|C\| \|X_0 - \bar{X}(\beta_l)\| \end{aligned}$$

donde, utilizando-se os resultados *i*) e *iii*) do teorema 3.2 e de (3.39), tem-se

$$\Delta_1 \leq \frac{\delta^2(X(\beta_l))}{1 - \delta^2(X(\beta_l))} + 1.5\mu\sqrt{p} \frac{\delta(X(\beta_l))}{1 - \delta(X(\beta_l))} = O(\mu\sqrt{p})$$

pois  $\delta(X(\beta_l)) \leq 0.5$ .

Calculando-se  $\Delta_2$

$$\begin{aligned} \Delta_2 &= f_{\beta_{l+1}}(\bar{X}(\beta_l)) - f_{\beta_{l+1}}(\bar{X}(\beta_{l+1})) \\ &= f_{\beta_l}(\bar{X}(\beta_l)) - f_{\beta_l}(\bar{X}(\beta_{l+1})) + \mu \beta_l \langle C, \bar{X}(\beta_l) - \bar{X}(\beta_{l+1}) \rangle \end{aligned}$$

Como os pontos  $\bar{X}(\beta_l)$  e  $\bar{X}(\beta_{l+1})$  são pontos que pertencem à trajetória central, pode-se associá-los, conforme lema 2.9, aos seguintes limitantes do objetivo dual  $d(\beta_l)$  e  $d(\beta_{l+1})$

$$\langle C, \bar{X}(\beta_l) \rangle = \eta(\beta_l) + d(\beta_l) \text{ e } \langle C, \bar{X}(\beta_{l+1}) \rangle = \eta(\beta_{l+1}) + d(\beta_{l+1})$$

Levando-se em conta que  $d(\beta_l) \leq d(\beta_{l+1})$  e  $f_{\beta_l}(\bar{X}(\beta_l)) \leq f_{\beta_l}(\bar{X}(\beta_{l+1}))$  obtém-se, finalmente, uma estimativa para  $\Delta_2$

$$\begin{aligned}
\Delta_2 &\leq \mu\beta_l \langle C, \bar{X}(\beta_l) - \bar{X}(\beta_{l+1}) \rangle \\
&\leq \mu\beta_l (\eta(\beta_l) - \eta(\beta_{l+1}) + d(\beta_l) - d(\beta_{l+1})) \\
&\leq \mu\beta_l \left( \frac{p}{\beta_l} - \frac{p}{\beta_{l+1}} \right) \\
&\leq \mu\beta_l \left( \frac{p\mu}{\beta_l(1+\mu)} \right) \\
&\leq \frac{p\mu^2}{1+\mu}
\end{aligned}$$

A partir daí, chega-se ao resultado final

$$\begin{aligned}
\bar{k} &= O \left( f_{\beta_{l+1}}(X_0) - f_{\beta_{l+1}}(\bar{X}(\beta_{l+1})) \right) \\
&= O(\Delta_1 + \Delta_2) \\
&\leq O(\mu\sqrt{p}) + \frac{p\mu^2}{1+\mu} \\
&= O \left( \frac{p\mu^2}{1+\mu} \right)
\end{aligned}$$

□

Assim, com o auxílio do lema 3.8, pode-se estudar a complexidade total dos algoritmos seguidores de trajetória. Esses métodos dividem-se em duas classes: os de passo curto (do inglês, *short-step*) e os de passo largo (do inglês, *large-step*). Inicia-se pelos de passo curto.

**Lema 3.9 Métodos de passo curto:** *Se o ponto quase central inicial  $X_0$  é tal que  $\delta(X_0) \leq 0.5$  e  $\eta(\alpha_0) < 1/\epsilon$ ,  $\epsilon < 1$ , utilizando-se na regra de atualização o parâmetro  $\theta = 1/(1 + 0.1/\sqrt{p})$ , o algoritmo 2 converge em  $O(\sqrt{p}\log(1/\epsilon))$  iterações do algoritmo 1.*

**Prova:** A prova decorre diretamente dos lemas 3.7 e 3.8. Do primeiro lema, observando-se que  $\log(1+a) \geq a/(1+a)$ , obtém-se

$$\begin{aligned}
l &= O \left( \frac{\log(1/\epsilon)}{\log(1 + 0.1/\sqrt{p})} \right) \\
&= O \left( \frac{(1 + 0.1/\sqrt{p}) \log(1/\epsilon)}{0.1/\sqrt{p}} \right) \\
&= O \left( \frac{(\sqrt{p} + 0.1) \log(1/\epsilon)}{0.1} \right) \\
&= O(\sqrt{p}\log(1/\epsilon))
\end{aligned}$$

Como o ponto  $i$ ) do segundo lema diz que o algoritmo 1 é capaz de gerar um ponto *quase* central em apenas uma iteração, o número total de iterações é igual  $l$ . Resta provar se é

possível associar ao ponto *quase* central  $X_l$  um *gap* de dualidade que satisfaça o critério de convergência. Pelo lema 3.3

$$\eta(X_l, \beta_l) \leq \frac{p + 0.5\sqrt{p}}{\beta_l} \leq 2\frac{p}{\beta_l} = 2\eta(\beta_l)$$

que, conforme o lema 3.7 torna-se menor do que  $\epsilon$  em no máximo  $l$  iterações.  $\square$

Segundo a prova apresentada para o ponto *i*) do lema 3.8, a garantia de que seja necessária apenas uma iteração do algoritmo 1 para se calcular um ponto *quase* central provém do seguinte fato

$$\delta(X_0) = \|\beta_{l+1}P_{\tilde{\mathcal{L}}}\tilde{C} - P_{\tilde{\mathcal{L}}}\mathbf{I}\| \leq \sqrt{0.5} \quad (3.40)$$

Ora, essa expressão pode ser resolvida para a igualdade a fim de se obter o maior valor de  $\beta_{l+1}$  que ainda assegure a convergência em um passo de tal forma que o limitante  $O(\sqrt{p}\log(1/\epsilon))$  dado no lema anterior ainda permaneça válido. Essa regra de atualização do parâmetro  $\beta$  é chamada de *o maior passo curto* (do inglês, *largest short step*).

Quando  $\theta < 1/(1 + 0.1/\sqrt{p})$  obtêm-se os chamados métodos de passo largo.

**Lema 3.10 Métodos de passo largo:** *Se o ponto quase central inicial  $X_0$  é tal que  $\delta(X_0) \leq 0.5$  e  $\eta(\alpha_0) < 1/\epsilon$ ,  $\epsilon < 1$ , utilizando-se na regra de atualização um parâmetro  $\theta = 1/(1 + \mu)$ ,  $\mu > 0.1/\sqrt{p}$ , o algoritmo 2 converge em  $O(\mu p \log(1/\epsilon))$  iterações do algoritmo 1.*

**Prova:** Do lemas 3.7 obtém-se

$$\begin{aligned} l &= O\left(\frac{\log(1/\epsilon)}{\log(1 + \mu)}\right) \\ &= O\left(\frac{(1 + \mu)\log(1/\epsilon)}{\mu}\right) \end{aligned}$$

Como o ponto *ii*) do lema 3.8 limita o número de iterações do algoritmo 1 a  $O(p\mu^2/(1 + \mu))$ , o número total de iterações é então limitado a  $O(\mu p \log(1/\epsilon))$ . As considerações tecidas ao final do último lema continuam válidas.  $\square$

Quando  $\mu = O(1)$  o lema acima assegura uma complexidade global de  $O(p\log(1/\epsilon))$ , maior portanto do que a obtida com passos curtos. Caso  $\mu = \nu/\sqrt{p}$ , com  $\nu = O(1)$ , obtém-se a chamada estratégia de *passo não muito curto* (do inglês, *not-so-short step*), que garante uma complexidade de  $O(\sqrt{p}\log(1/\epsilon))$ .

### 3.3 Algoritmos primais-duais

Utilizando-se a notação

$$A \oplus B = \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & B \end{bmatrix} \quad (3.41)$$

e segundo a equação 2.30, a solução do problema

$$\begin{aligned} &\text{minimizar } \langle F, W \rangle \\ &\text{s.a } W \in \{\mathcal{K} \cap (\mathcal{L} + D)\} \oplus \{\mathcal{K} \cap (\mathcal{L}^\perp + C)\} \end{aligned} \quad (3.42)$$

em que

$$F = C \oplus D \text{ e } W = X \oplus Z \quad (3.43)$$

cuja função objetivo equivale ao *gap* de dualidade  $\langle X, Z \rangle$ , resolve simultaneamente os problemas (2.26) e (2.29)

De maneira análoga à formulação exclusivamente primal, pode-se definir uma trajetória central primal-dual a partir do problema

$$\begin{aligned} & \text{minimizar} && B_{\mathcal{K}}(X) + B_{\mathcal{K}}(Z) \\ & \text{s.a} && \begin{cases} X \in \{\mathcal{K}^0 \cap (\mathcal{L} + D)\} \\ Z \in \{\mathcal{K}^0 \cap (\mathcal{L}^\perp + C)\} \\ \langle X, Z \rangle = \langle C, X \rangle + \langle D, Z \rangle = \eta \end{cases} \end{aligned} \quad (3.44)$$

e do seguinte lema.

**Lema 3.11** *Um par de pontos primal e dual factível,  $(\bar{X}(\eta), \bar{Z}(\eta))$ , pertence à trajetória central primal-dual se, e somente se, a seguinte condição for satisfeita*

$$\frac{p}{\eta} \bar{X}(\eta) \bar{Z}(\eta) = \mathbf{I} \quad (3.45)$$

**Prova:** Construindo-se a função lagrangeana

$$L(X, Z, \mu) := -\log \det X - \log \det Z + \mu(\langle X, Z \rangle - \eta)$$

as condições de otimalidade para esse problema são dadas por

$$\mu \bar{Z}(\eta) = \bar{X}^{-1}(\eta)$$

e

$$\mu \bar{X}(\eta) = \bar{Z}^{-1}(\eta)$$

que, juntamente com as restrições de factibilidade do problema, implicam  $\bar{X}^{-1}(\eta)/\mu \in (\mathcal{L}^\perp + C)$  e  $\bar{Z}^{-1}(\eta)/\mu \in (\mathcal{L} + D)$ . Como a função objetivo é estritamente convexa, essas duas condições devem corresponder a um único par de pontos  $\bar{X}(\eta) \in \{\mathcal{K}^0 \cap (\mathcal{L} + D)\}$  e  $\bar{Z}(\eta) \in \{\mathcal{K}^0 \cap (\mathcal{L}^\perp + C)\}$ , consequentemente, o valor de  $\mu$  pode ser obtido por meio da restrição

$$\langle \bar{X}(\eta), \bar{Z}(\eta) \rangle = \frac{p}{\mu} = \eta$$

Essa condição é também necessária uma vez que o par solução é único.  $\square$

Assim, pode-se resolver esse problema linear primal-dual na variável  $W$  por meio de qualquer um dos já discutidos métodos seguidores de trajetória. Para tal, basta que a atualização dos parâmetros livres seja feita em função da redução do *gap* de dualidade na forma

$$\eta_{l+1} = \theta \eta_l \quad (3.46)$$

Como na formulação exclusivamente primal, consegue-se isso observando-se as relações entre os gradientes das funções auxiliares e o lema acima. Por exemplo, para a correspondente função apenada

$$\begin{aligned} f_\beta(W) = f_\beta(X, Z) &= \beta \langle F, W \rangle + B_{\mathcal{K}}(W) \\ &= \beta \langle X, Z \rangle + B_{\mathcal{K}}(X) + B_{\mathcal{K}}(Z) \end{aligned} \quad (3.47)$$

cujo gradiente vale

$$\begin{aligned} \nabla f_\beta(W) &= \nabla f_\beta(X, Z)_X \oplus \nabla f_\beta(X, Z)_Z \\ &= (\beta C - X^{-1}) \oplus (\beta D - Z^{-1}) \\ &= (\beta Z - X^{-1}) \oplus (\beta X - Z^{-1}) \end{aligned} \quad (3.48)$$

a regra de atualização é, novamente,

$$\beta_{l+1} = \frac{\beta_l}{\theta} = \frac{p/\theta}{\langle \bar{X}(\beta_l), \bar{Z}(\beta_l) \rangle} \quad (3.49)$$

Em relação à determinação das projeções exigidas pelo algoritmo 1, alguns comentários se fazem necessários. A direção de Newton  $H_W = H_X \oplus H_Z$  é obtida como a solução de dois problemas de mínimos quadrados independentes

$$H_X = \arg \min_{H_X \in \mathcal{L}} \left\{ \|X^{\frac{1}{2}} \nabla f_\beta(X, Z)_X X^{\frac{1}{2}} + X^{-\frac{1}{2}} H_X X^{-\frac{1}{2}}\| \right\} \quad (3.50)$$

e

$$H_Z = \arg \min_{H_Z \in \mathcal{L}^\perp} \left\{ \|Z^{\frac{1}{2}} \nabla f_\beta(X, Z)_Z Z^{\frac{1}{2}} + Z^{-\frac{1}{2}} H_Z Z^{-\frac{1}{2}}\| \right\} \quad (3.51)$$

Conforme discutido na seção 3.1.2, esses problemas são equivalentes à determinação das projeções

$$H_X = -X^{\frac{1}{2}} P_{\tilde{\mathcal{L}}_{X^{-1}}} \left( \beta X^{\frac{1}{2}} Z X^{\frac{1}{2}} - \mathbf{I} \right) X^{\frac{1}{2}} \quad (3.52)$$

e

$$H_Z = -Z^{\frac{1}{2}} P_{\tilde{\mathcal{L}}^\perp_{(Z^{-1})}} \left( \beta Z^{\frac{1}{2}} X Z^{\frac{1}{2}} - \mathbf{I} \right) Z^{\frac{1}{2}} \quad (3.53)$$

em que  $\mathcal{L}$  sofre uma transformação definida por  $X^{-1}$  enquanto  $\mathcal{L}^\perp$  é transformado por  $Z^{-1}$ . Como, na prática, a dimensão de  $\mathcal{L}^\perp$  é bem maior do que a dimensão de  $\mathcal{L}$ , pode ser vantajoso avaliar esta projeção em  $\mathcal{L}^\perp_{Z^{-1}}$  calculando-se o complemento ortogonal<sup>2</sup> da projeção em  $\tilde{\mathcal{L}}_Z$ . Em todo caso, é sempre necessário avaliar dois operadores projeção diferentes, o que significa praticamente dobrar o esforço computacional requerido por cada iteração<sup>3</sup>. Além do mais, observando-se a expressão do gradiente de  $f_\beta(W)$ , percebe-se que  $H_X$  e  $H_Z$  podem ser colocados em função apenas de  $X$  e  $Z$ , respectivamente, o que significa que é possível visualizar essas direções como provenientes de dois problemas completamente desacoplados! Considerando-se que os lemas 2.3 e 3.3 fornecem matrizes duais factíveis a partir de matrizes primais centrais ou quase centrais, as iterações em  $Z$  tornam-se, para todos os efeitos, dispensáveis.

<sup>2</sup>Veja seções 2.5 e 2.6.

<sup>3</sup>Veja seção 4.1.

Esse fato motiva a busca por outras direções de minimização que requeiram menos esforço computacional. A idéia básica consiste em utilizar o conceito de *direção de máximo decrescimento* após aplicar transformações simétricas definidas pelas matrizes  $D^{-1}$  e  $D$  às matrizes primal e dual, respectivamente, de tal forma que ambas as direções  $H_X$  e  $H_Z$  possam ser determinadas por meio da avaliação de projeções em  $\tilde{\mathcal{L}}_{D^{-1}}$ . Dois casos são obtidos trivialmente: a transformação primal, em que  $D = X$ , e a transformação dual, em que  $D = Z^{-1}$ . Nesses casos, uma das direções sempre coincide com a direção de Newton enquanto a outra é apenas uma *direção de máximo decrescimento* transformada. Uma outra opção, um pouco mais elaborada, baseia-se no resultado do próximo lema.

**Lema 3.12** *A matriz*

$$T = X^{-\frac{1}{2}} U \Lambda^{\frac{1}{4}} \quad (3.54)$$

em que  $U$  e  $\Lambda$  são, respectivamente, a matriz ortogonal e a matriz diagonal provenientes da decomposição de Schur da matriz  $X^{\frac{1}{2}} Z X^{\frac{1}{2}}$ , diagonaliza o gradiente da função  $f_\beta(X, Z)$  na forma

$$\nabla f_\beta(X, Z) = T \left( \beta \Lambda^{\frac{1}{2}} - \Lambda^{-\frac{1}{2}} \right) T' \oplus (T')^{-1} \left( \beta \Lambda^{\frac{1}{2}} - \Lambda^{-\frac{1}{2}} \right) T^{-1} \quad (3.55)$$

**Prova:** A partir das matrizes estritamente definidas positivas  $X$  e  $Z$ , calcula-se a decomposição de Schur  $U \Lambda U' = X^{\frac{1}{2}} Z X^{\frac{1}{2}}$ , em que  $U U' = U' U = \mathbf{I}$  e  $\Lambda$  é uma matriz diagonal definida positiva. Assim, o primeiro componente do gradiente de  $f_\beta(X, Z)$  é dado por

$$\begin{aligned} \nabla f_\beta(X, Z)_X &= X^{-\frac{1}{2}} \left( \beta X^{\frac{1}{2}} Z X^{\frac{1}{2}} - \mathbf{I} \right) X^{-\frac{1}{2}} \\ &= X^{-\frac{1}{2}} \left( \beta U \Lambda U' - U U' \right) X^{-\frac{1}{2}} \\ &= X^{-\frac{1}{2}} \left( \beta U \Lambda^{\frac{1}{4}} \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{4}} U' - U \Lambda^{\frac{1}{4}} \Lambda^{-\frac{1}{2}} \Lambda^{\frac{1}{4}} U' \right) X^{-\frac{1}{2}} \\ &= T \left( \beta \Lambda^{\frac{1}{2}} - \Lambda^{-\frac{1}{2}} \right) T' \end{aligned}$$

Lembrando-se que  $\left( X^{\frac{1}{2}} Z X^{\frac{1}{2}} \right)^{-1} = X^{-\frac{1}{2}} Z^{-1} X^{-\frac{1}{2}} = U \Lambda^{-1} U'$ , finaliza-se a prova aplicando-se o mesmo raciocínio ao segundo componente do gradiente

$$\begin{aligned} \nabla f_\beta(X, Z)_Z &= X^{\frac{1}{2}} \left( \beta \mathbf{I} - X^{-\frac{1}{2}} Z^{-1} X^{-\frac{1}{2}} \right) X^{\frac{1}{2}} \\ &= X^{\frac{1}{2}} \left( \beta U \Lambda^{-\frac{1}{4}} \Lambda^{\frac{1}{2}} \Lambda^{-\frac{1}{4}} U' - U \Lambda^{-\frac{1}{4}} \Lambda^{-\frac{1}{2}} \Lambda^{-\frac{1}{4}} U' \right) X^{\frac{1}{2}} \\ &= (T')^{-1} \left( \beta \Lambda^{\frac{1}{2}} - \Lambda^{-\frac{1}{2}} \right) T^{-1} \end{aligned}$$

□

Segundo esse lema e conforme a discussão no final da seção 2.6, isso significa que a matriz  $T'T$  define uma transformação simétrica que permite determinar as direções de máximo decrescimento

$$H_X = - (T')^{-1} P_{\tilde{\mathcal{L}}_{T'T}} \left( \beta \Lambda^{\frac{1}{2}} - \Lambda^{-\frac{1}{2}} \right) T^{-1} \quad (3.56)$$

e

$$H_Z = - T P_{\tilde{\mathcal{L}}_{(T'T)^{-1}}} \left( \beta \Lambda^{\frac{1}{2}} - \Lambda^{-\frac{1}{2}} \right) T' \quad (3.57)$$

Desse modo, o esforço computacional associado ao cálculo das direções  $H_X$  e  $H_Z$  reduz-se ainda mais, pois, uma vez que os operandos das projeções são idênticos e as matrizes

que definem as transformações primal e dual são inversas uma da outra, determina-se  $H_Z$  aplicando-se uma transformação simétrica à matriz complemento ortogonal da projeção já obtida no cálculo de  $H_X$ . Esta transformação é equivalente ao chamado *primal-dual scaling* em programação linear, e será denominada transformação primal-dual. Cabe ressaltar que todas essas direções são equivalentes, caso o par  $(X, Z)$  pertença à trajetória central.

Em relação à medida de centralização, caso  $H_X$  e  $H_Z$  sejam calculados como as direções de Newton primal e dual, é possível aplicar os resultados do teorema 3.2 ao escalar  $\delta(W) = \delta(X, Z) = \|W^{-\frac{1}{2}} H_W W^{-\frac{1}{2}}\|$ . Alternativamente, para as demais direções de decrescimento, pode-se utilizar o escalar

$$\tilde{\delta}(X, Z) := \|\beta X^{\frac{1}{2}} Z X^{\frac{1}{2}} - \mathbf{I}\| = \left( \sum_{i=1}^p \lambda_i^2(\beta X Z - \mathbf{I}) \right)^{\frac{1}{2}} \quad (3.58)$$

que, segundo o lema 3.11, vale zero apenas para pontos pertencentes à trajetória central. Alguns resultados referentes a este escalar estão reunidos no próximo lema.

**Lema 3.13** *Dados o parâmetro  $\beta > 0$  e o par primal-dual factível  $(X, Z)$ , as seguintes afirmações são verdadeiras*

i)  $f_\beta(X, Z) \geq p + p \log \beta$

ii) e, se  $\tilde{\delta}(X, Z) < 1$ , então  $f_\beta(X, Z) \leq p + p \log \beta - \tilde{\delta}(X, Z) - \log(1 - \tilde{\delta}(X, Z))$

**Prova:** A função apenas  $f_\beta(X, Z)$  pode ser escrita como

$$\begin{aligned} f_\beta(X, Z) &= \text{tr}(\beta X Z) - \log \det(X Z) \\ &= \sum_{i=1}^p \lambda_i(\beta X Z) - \sum_{i=1}^p \log \lambda_i(\beta X Z) + p \log \beta \end{aligned}$$

Denotando-se por  $\mu_i$  o  $i$ ésimo autovalor da matriz  $(\beta X Z - \mathbf{I})$ , pode-se reescrever essa última expressão na forma

$$f_\beta(X, Z) = p + p \log \beta + \sum_{i=1}^p \mu_i - \sum_{i=1}^p \log(\mu_i + 1)$$

O primeiro ponto do lema pode ser provado observando-se que, devido ao fato de as matrizes  $X$  e  $Z$  serem definidas positivas,  $\mu_i > -1$ , logo,  $\mu_i - \log(\mu_i + 1) \geq 0$  e

$$f_\beta(X, Z) \geq p + p \log \beta$$

Na prova do segundo ponto, observando-se que  $|\mu_i| \leq \tilde{\delta}(X, Z) < 1$ , torna-se possível expandir cada termo logarítmico em série de potências, o que leva a

$$\begin{aligned} f_\beta(X, Z) &= p + p \log \beta + \sum_{i=1}^p \mu_i - \sum_{i=1}^p \mu_i + \sum_{k=2}^{\infty} \frac{(-1)^k}{k} \sum_{i=1}^p \mu_i^k \\ &\leq p + p \log \beta + \sum_{k=2}^{\infty} \frac{1}{k} \left( \sum_{i=1}^p \mu_i^2 \right)^{k/2} \\ &\leq p + p \log \beta - \tilde{\delta}(X, Z) - \log(1 - \tilde{\delta}(X, Z)) \end{aligned}$$

□

### 3.3.1 A função potencial e a formulação primal-dual

Dentre todas as funções auxiliares definidas no início deste capítulo, a função potencial, definida na seção 3.2.3, talvez seja a que melhor proveito consiga tirar da formulação primal-dual. Grande parte das dificuldades associadas aos métodos estritamente primais que utilizam a função potencial reside no cálculo do parâmetro  $\rho$ , que deve, ao ser atualizado, manter-se ainda um limitante inferior do valor objetivo primal, ou seja, um objetivo dual factível. Caso se utilize a formulação primal-dual, a presença de um ponto dual factível durante todo o processo de otimização, permite que o cálculo de  $\rho$  seja feito de maneira adequada. Em Gonzaga [10], descrevem-se diversas variações do algoritmo seguidor de trajetória capazes de levar em conta esse fato em programação linear.

Uma outra classe de métodos, conhecida pelo nome de métodos redutores de potencial, explora a formulação primal-dual de uma maneira um pouco diferente: como o valor do objetivo ótimo  $\langle X^*, Z^* \rangle$  é conhecido e vale zero, a função

$$f_q(X, Z) = q \log \langle X, Z \rangle + B_{\mathcal{K}}(X) + B_{\mathcal{K}}(Z) \quad (3.59)$$

em que  $q$  é maior ou igual a  $p$ , que é a função potencial  $f_p(W)$  escrita para o maior limitante inferior do objetivo primal-dual, ou seja  $\rho = 0$ , tem como centro analítico o par de pontos solução dos problemas (2.26) e (2.29). A idéia, então, consiste em se tentar minimizar diretamente essa função, sem necessariamente ter que se manter próximo à trajetória central. No entanto, de uma certa maneira, grande parte da informação contida nessa função potencial primal-dual refere-se à trajetória central. Impondo-se  $q = p$ , esta função pode ser escrita como

$$f_p(X, Z) = p \log \langle X, Z \rangle - \log \det XZ \quad (3.60)$$

$$= p \log \frac{\text{tr}(XZ)}{(\det XZ)^{1/p}} \quad (3.61)$$

$$= p \log p + p \log \frac{\sum_{i=1}^p \lambda_i(XZ)/p}{(\prod_{i=1}^p \lambda_i(XZ))^{1/p}} \quad (3.62)$$

$$= p \log p + \psi(X, Z) \quad (3.63)$$

Assim, pode-se interpretar  $\psi(X, Z)$  — o logaritmo da razão entre a média aritmética e a média geométrica dos autovalores do produto  $XZ$  — como uma medida de centralização do par primal-dual porque no hiperplano definido por  $\langle X, Z \rangle = \eta$ , segundo o lema 3.11,

$$\psi(X, Z) = p \log \langle X, Z \rangle - \log \det XZ - p \log p \quad (3.64)$$

$$= -\log \det XZ - p \log(p/\eta)$$

$$= -\log \det XZ - \log \det \bar{X}(\eta)\bar{Z}(\eta) \geq 0$$

Isso implica que  $\psi(X, Z)$  é uma função que nunca assume valores negativos e, além do mais, apenas se anula na trajetória central primal-dual. Essa propriedade métrica associada a  $\psi(X, Z)$  motiva a definição da função

$$f_\mu(X, Z) = \mu \log \langle X, Z \rangle + \psi(X, Z) \quad (3.65)$$

Note-se que  $f_q(X, Z)$  e  $f_\mu(X, Z)$ , para  $q = \mu + p$ , apresentam como diferença apenas o termo constante ( $p \log p$ ) que é irrelevante do ponto de vista de otimização. Em relação à função  $f_\mu(X, Z)$ , formula-se o seguinte lema.

**Lema 3.14** Dado o par central primal-dual  $(\bar{X}(\eta_0), \bar{Z}(\eta_0))$  associado ao gap de dualidade  $\eta_0$ , se for possível gerar iterativamente uma seqüência de pontos tal que

$$f_\mu(X_{k+1}, Z_{k+1}) \leq f_\mu(X_k, Z_k) - \theta \quad (3.66)$$

em que  $\theta = O(1)$  é uma constante, então, obtém-se um par de pontos  $(X_k, Z_k)$  tal que  $\langle X_k, Z_k \rangle \leq \epsilon \eta_0$  em, no máximo,  $O(\mu \log(1/\epsilon))$  iterações.

**Prova:** Impondo-se  $X_0 = \bar{X}(\eta_0)$  e  $Z_0 = \bar{Z}(\eta_0)$ , é possível escrever, para toda seqüência que satisfaça o enunciado do lema acima, que

$$f_\mu(X_k, Z_k) \leq f_\mu(X_0, Z_0) - k\theta$$

Por sua vez, observando-se que  $\psi(X_k, Z_k) \geq 0$ , tem-se

$$\mu \log \langle X_k, Z_k \rangle \leq f_\mu(X_0, Z_0) - k\theta = \mu \log \langle X_0, Z_0 \rangle + \psi(X_0, Z_0) - k\theta$$

e, portanto

$$\frac{\langle X_k, Z_k \rangle}{\langle X_0, Z_0 \rangle} \leq \exp \frac{\psi(X_0, Z_0) - k\theta}{\mu}$$

cujos termos da direita serão menores do que  $\epsilon$  para

$$\bar{k} \geq (\mu \log(1/\epsilon) + \psi(X_0, Z_0))/\theta$$

e o resultado do lema é obtido observando-se que  $\psi(X_0, Z_0) = 0$  □

Esse lema pode ser ligeiramente modificado de modo a permitir pontos iniciais “razoavelmente” centrados. Na prova acima, percebe-se que o limitante do número de iterações  $\bar{k}$  permanece inalterado caso a ordem de  $\psi(X_0, Z_0)$  seja menor do que  $\mu \log(1/\epsilon)$ .

Em todo caso, convergência em  $O(\sqrt{p} \log(1/\epsilon))$  iterações pode ser obtida impondo-se  $\mu = \nu \sqrt{p}$ ,  $\nu = O(1)$ . Basta, então, saber como gerar uma seqüência que possua a propriedade descrita no enunciado do lema. De maneira análoga à formulação exclusivamente primal, o fato de a função potencial primal-dual ser a soma de uma função convexa e uma função côncava permite definir a função

$$\begin{aligned} fl_\mu(X, Z) &= (p + \mu) \log \langle \hat{X}, \hat{Z} \rangle - p \log p + \\ &+ \langle \nabla V(\hat{X}, \hat{Z}), (X - \hat{X}) \oplus (Z - \hat{Z}) \rangle + B_\kappa(X) + B_\kappa(Z) \end{aligned} \quad (3.67)$$

em que  $\nabla V(\hat{X}, \hat{Z})$  representa o valor do gradiente do termo côncavo  $V(X, Z) = (p + \mu) \log \langle X, Z \rangle$  avaliado no par de pontos  $(\hat{X}, \hat{Z})$ . Verifica-se que  $fl_\mu(X, Z)$  é equivalente à soma de um termo constante a uma função apenas cujo parâmetro  $\beta$  vale  $(p + \mu)/\langle \hat{X}, \hat{Z} \rangle$ , e o seguinte lema pode ser formulado.

**Lema 3.15** Dados o parâmetro  $\beta = (p + \mu)/\langle \hat{X}, \hat{Z} \rangle$ , com  $\mu = \nu \sqrt{p}$ ,  $\nu > 1$ , e a correspondente função apenas definida<sup>4</sup> por  $fl_\mu(X, Z)$  em torno do par de pontos factível  $(\hat{X}, \hat{Z})$ , o valor de  $\delta(W) = \delta(X, Z)$  é sempre maior ou igual a  $\epsilon = 0.35$ .

<sup>4</sup>A menos de um termo constante.

**Prova:** Sendo  $f_\beta(X, Z)$  a função apenada associada a  $fl_\mu(X, Z)$  construída em torno do par de pontos  $(\hat{X}, \hat{Z})$ , o componente em  $X$  da direção de Newton é dado por (omitindo-se os acentos circunflexos a fim de simplificar a notação)

$$X^{-\frac{1}{2}}H_XX^{-\frac{1}{2}} = -P_{L_{X^{-1}}} \left( X^{\frac{1}{2}}\nabla f_\beta(X, Z)X^{\frac{1}{2}} \right)$$

e, portanto, para uma matriz  $Z_X \in (\mathcal{L}^\perp + C)$ , pode-se escrever

$$\begin{aligned} -\beta X^{\frac{1}{2}}(Z_X - C)X^{\frac{1}{2}} &= X^{\frac{1}{2}}\nabla f_\beta(X, Z)_X X^{\frac{1}{2}} + X^{-\frac{1}{2}}H_XX^{-\frac{1}{2}} \\ &= X^{\frac{1}{2}}(\beta C - X^{-1})X^{\frac{1}{2}} + X^{-\frac{1}{2}}H_XX^{-\frac{1}{2}} \end{aligned}$$

ou, de maneira equivalente,

$$Z_X = \frac{1}{\beta}X^{-\frac{1}{2}} \left( \mathbf{I} - X^{-\frac{1}{2}}H_XX^{-\frac{1}{2}} \right) X^{-\frac{1}{2}}$$

Aplicando-se o mesmo raciocínio à direção  $H_Z$  obtém-se  $X_Z \in (\mathcal{L} + D)$  na forma

$$X_Z = \frac{1}{\beta}Z^{-\frac{1}{2}} \left( \mathbf{I} - Z^{-\frac{1}{2}}H_ZZ^{-\frac{1}{2}} \right) Z^{-\frac{1}{2}}$$

Supondo-se que o valor de  $\delta(W) = \sqrt{\delta^2(\hat{X}) + \delta^2(\hat{Z})} = \varepsilon < 1$ , então

$$\delta(X) = \|X^{-\frac{1}{2}}H_XX^{-\frac{1}{2}}\| = \|\beta X^{\frac{1}{2}}Z_XX^{\frac{1}{2}} - \mathbf{I}\| = \tilde{\delta}(X, Z_X) \leq \varepsilon$$

e

$$\delta(Z) = \|Z^{-\frac{1}{2}}H_ZZ^{-\frac{1}{2}}\| = \|\beta Z^{\frac{1}{2}}X_ZZ^{\frac{1}{2}} - \mathbf{I}\| = \tilde{\delta}(X_Z, Z) \leq \varepsilon$$

Assim, segundo o lema 2.5, os pontos  $X_Z$  e  $Z_X$  são factíveis e, conforme o segundo ponto do lema 3.13,

$$f_\beta(X, Z_X) \leq p + p \log \beta - \varepsilon - \log(1 - \varepsilon)$$

e

$$f_\beta(X_Z, Z) \leq p + p \log \beta - \varepsilon - \log(1 - \varepsilon)$$

Utilizando-se, então, o primeiro ponto do lema 3.13, pois  $X_Z$  e  $Z_X$  são factíveis, tem-se que

$$f_\beta(X_Z, Z_X) \geq p + p \log \beta$$

logo, combinando-se essas três últimas expressões,

$$\begin{aligned} \beta \langle X, Z \rangle - \log \det(XZ) - \beta \langle X - X_Z, Z - Z_X \rangle &= f_\beta(X, Z_X) + f_\beta(X_Z, Z) - f_\beta(X_Z, Z_X) \\ &\leq p + p \log \beta - 2(\varepsilon + \log(1 - \varepsilon)) \end{aligned}$$

Observando-se que

$$\begin{aligned} Z - Z_X &= \frac{1}{\beta} \left( \beta Z - X^{-1} + X^{-1}H_XX^{-1} \right) \\ &= \frac{1}{\beta} \left( \nabla f_\beta(X, Z)_X + X^{-1}H_XX^{-1} \right) \\ &= \frac{1}{\beta} X^{-\frac{1}{2}} \left( X^{\frac{1}{2}}\nabla f_\beta(X, Z)_X X^{\frac{1}{2}} + X^{-\frac{1}{2}}H_XX^{-\frac{1}{2}} \right) X^{-\frac{1}{2}} \end{aligned}$$

pode-se concluir, pela definição de  $H_X$ , que  $Z - Z_X \in \mathcal{L}^\perp$ . O mesmo raciocínio aplicado a  $X - X_Z$ , implica  $X - X_Z \in \mathcal{L}$  e, portanto,  $\langle X - X_Z, Z - Z_X \rangle = 0$ . Assim, da desigualdade anterior obtém-se

$$p \log \langle X, Z \rangle - \log \det(XZ) \leq p \log(p + \nu\sqrt{p}) - \nu\sqrt{p} - 2(\varepsilon + \log(1 - \varepsilon))$$

e, finalmente,

$$\begin{aligned} \psi(X, Z) &= p \log \langle X, Z \rangle - \log \det(XZ) - p \log p \\ &\leq p \log(p + \nu\sqrt{p}) - p \log p - \nu\sqrt{p} - 2(\varepsilon + \log(1 - \varepsilon)) \\ &\leq p(\log(1 + \nu/\sqrt{p}) - \nu/\sqrt{p}) - 2(\varepsilon + \log(1 - \varepsilon)) \\ &\leq p(\log(1 + 1/\sqrt{p}) - 1/\sqrt{p}) - 2(\varepsilon + \log(1 - \varepsilon)) \\ &\leq \frac{1}{3\sqrt{p}} - \frac{1}{2} - 2(\varepsilon + \log(1 - \varepsilon)) \\ &\leq -\frac{1}{6} - 2(\varepsilon + \log(1 - \varepsilon)) \end{aligned}$$

pois  $\nu > 1$  e  $\log(1 + a) \leq (a - a^2/2 + a^3/3)$  para  $a > -1$ . Obtém-se, assim, por meio dessa desigualdade, o limitante apresentado, pois  $\psi(X, Z)$  deve se manter sempre positiva, logo

$$\varepsilon + \log(1 - \varepsilon) \leq -1/12$$

o que implica  $\varepsilon \geq 0.35$ , e encerra esta prova.  $\square$

Esse lema, associado às seguintes propriedades

$$fl_\mu(\hat{X}, \hat{Z}) \geq f_\mu(\hat{X}, \hat{Z}) \text{ e } fl_\mu(\hat{X}, \hat{Z}) = f_\mu(\hat{X}, \hat{Z}) \quad (3.68)$$

fornece a chave para a geração da desejada seqüência: num par de pontos  $(X_k, Z_k)$  obtém-se a função  $fl_\mu(X, Z)$ ; aplica-se, então, o passo fornecido pelo teorema 3.2 na direção de Newton da correspondente função apenas a fim de obter  $(X_{k+1}, Z_{k+1})$ ; reavalia-se  $fl_\mu(X, Z)$  no par de pontos  $(X_{k+1}, Z_{k+1})$  e se repete o processo até se alcançar a precisão desejada. Procedendo-se dessa maneira, o item *iv*) do teorema 3.2 e a inequação acima asseguram que

$$f_\mu(X_{k+1}, Z_{k+1}) \leq f_\mu(X_k, Z_k) - \theta \quad (3.69)$$

e o lema 3.15 fornece o limitante inferior para  $\theta = \delta(W) - \log(1 + \delta(W)) \geq 0.05$ .

Mesmo para este “esboço” de algoritmo, caso  $\mu = \nu\sqrt{p}$  e a constante  $\nu$  seja maior do que um, é possível garantir, conforme o lema 3.14, que o número de iterações é limitado por  $O(\sqrt{p} \log(1/\varepsilon))$ . Na prática, a taxa de convergência pode ser enormemente acelerada através de procedimentos de busca linear que visam minimizar a função  $fl_\mu(X_k + \gamma H_{X_k}, Z_k + \gamma H_{Z_k})$ , por exemplo, por meio do teorema 3.4. Uma alternativa consiste em se tentar minimizar diretamente a função não convexa  $f_\mu(X_k + \gamma H_{X_k}, Z_k + \gamma H_{Z_k})$ . Alguns algoritmos chegam até mesmo a implementar a chamada busca plana, isto é, resolvem a cada iteração o problema

$$\min_{\gamma_X, \gamma_Z \in \Phi_\gamma} f_\mu(X + \gamma_X H_X, Z + \gamma_Z H_Z) \quad (3.70)$$

em que  $\Phi_\gamma$  é um conjunto que contém apenas valores de  $\gamma_X$  e  $\gamma_Z$  que asseguram a factibilidade dos pontos  $X + \gamma_X H_X$  e  $Z + \gamma_Z H_Z$ . Este procedimento visa aproveitar o relativo<sup>5</sup> desacoplamento entre as direções de decrescimento primal e dual, e voltará a ser discutido no próximo capítulo de maneira bem mais detalhada. É importante frisar que esse tipo de procedimento é que garante o bom desempenho desses métodos, uma vez que o limitante teórico obtido para  $\theta$  é bem pequeno.

Assim como discutido para a função apenada primal-dual, pode-se fazer uso das diversas transformações simétricas que reduzem a um o número de projeções calculadas por iteração. Felizmente, mesmo para estes casos, ainda é possível obter resultados semelhantes ao do lema 3.15. Veja, por exemplo, a seção 5 em Vandenberghe & Boyd [24].

Fechando esta seção, apresenta-se um dos possíveis métodos que podem ser construídos com base nas discussões anteriores.

**Algoritmo 3** Dados o par de pontos factível  $(X_0, Z_0)$ ,  $\mu = \nu\sqrt{p}$  e  $\nu > 1$

1. Calcular  $\beta_k = (p + \mu) / \langle X_k, Z_k \rangle$  e as direções

$$(H_X)_k = -X_k^{\frac{1}{2}} P_{\tilde{\mathcal{L}}_{X_k^{-1}}} \left( \beta_k X_k^{\frac{1}{2}} Z_k X_k^{\frac{1}{2}} - \mathbf{I} \right) X_k^{\frac{1}{2}} \quad (3.71)$$

e

$$(H_Z)_k = -Z_k^{\frac{1}{2}} P_{\tilde{\mathcal{L}}_{(Z_k^{-1})}} \left( \beta_k Z_k^{\frac{1}{2}} X_k Z_k^{\frac{1}{2}} - \mathbf{I} \right) Z_k^{\frac{1}{2}} \quad (3.72)$$

2. Efetuar a busca plana

$$\min_{\alpha_X, \alpha_Z \in \Phi_\alpha} f_\mu(X_k + \alpha_X (H_X)_k, Z_k + \alpha_Z (H_Z)_k) \quad (3.73)$$

3. Atualizar  $X_{k+1} = X_k + \alpha_X (H_X)_k$  e  $Z_{k+1} = Z_k + \alpha_Z (H_Z)_k$
4. Parar caso  $\langle X_{k+1}, Z_{k+1} \rangle < \epsilon$ . Caso contrário, incrementar  $k$  e voltar ao primeiro passo.

□

### 3.4 Algoritmo projetivo

O algoritmo discutido nesta seção trabalha os conceitos desenvolvidos ao longo do capítulo 2 de maneira um pouco diferente. Fundamentalmente, o problema (2.1), cujo conjunto de restrições consiste em uma combinação afim de matrizes, é transformado num problema equivalente cujas restrições são uma combinação linear de matrizes. Neste sentido, introduzem-se a variável auxiliar  $\tau$ , o vetor  $x_e = [x' \ \tau]'$  e o subespaço linear  $\mathcal{L}_e$  associado à LMI

$$F_e(x_e) = \begin{bmatrix} F_0\tau + \sum_{i=1}^n F_i x_i & \mathbf{0} \\ \mathbf{0} & \tau \end{bmatrix} \quad (3.74)$$

É claro que, do ponto de vista da factibilidade estrita, se existe  $x$  de tal forma que  $F(x) > 0$ , então  $F_e([x' \ 1]') > 0$  e, de modo contrário, se existe  $x_e = [x' \ \tau]'$  de tal forma que  $F_e(x_e) > 0$ ,

<sup>5</sup>Completo apenas para a função apenada e não mais para a função potencial.

então  $F(x/\tau) > 0$ , pois  $\tau > 0$ . Assim, o problema (2.1) torna-se equivalente ao problema de otimização da função não linear  $c'x/\tau$  sujeito a  $F_e(x_e) > 0$ , que pode ser escrito na forma

$$\begin{aligned} &\text{minimizar} && c'_e x_e / g'_e x_e \\ &\text{s.a.} && F_e(x_e) \geq 0 \end{aligned} \quad (3.75)$$

em que  $c_e = [c' \ 0]'$  e  $g_e = [0' \ 1]$ . Novamente, como na seção 2.4, esse conjunto de restrições pode ser reescrito em termos de variáveis matriciais pertencentes ao subespaço linear  $\mathcal{L}_e$  de dimensão  $p_e = p+1$ , de tal forma que os conjuntos  $\{x_e \mid F_e(x_e) \geq 0\}$  e  $\{X_e \mid X_e \in \{\mathcal{K}_e \cap \mathcal{L}_e\}\}$ , no qual  $\mathcal{K}_e$  denota o cone das matrizes reais e simétricas definidas positivas de dimensão  $p_e$ , sejam equivalentes. Para isso, o subespaço linear  $\mathcal{L}_e$  deve ter por base as matrizes

$$(F_e)_{i=1,\dots,p} = \begin{bmatrix} F_i & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \text{ e } (F_e)_{p+1} = \begin{bmatrix} F_0 & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \quad (3.76)$$

Quanto à função objetivo *linear fracional* ( $c'_e x_e / g'_e x_e$ ), note-se que  $g'_e x_e > 0$  sempre que  $F_e(x_e) > 0$ . Maiores detalhes encontram-se na seção 3.4.2.

Deste ponto em diante, enquanto se estiver descrevendo o algoritmo projetivo, apenas o problema (3.75) está sob foco e, a fim de simplificar a notação, não mais se faz referência explícita ao “e” subscrito, utilizado acima para introdução das variáveis e constantes auxiliares. Ao contrário dos demais métodos estudados, o algoritmo projetivo pode ser mais bem entendido analisando-se primeiramente o problema de factibilidade.

### 3.4.1 O problema de factibilidade

O objetivo do problema de factibilidade estrita é encontrar  $X \in \{\mathcal{K}^0 \cap \mathcal{L}\}$ . Para tal, o método projetivo se utiliza do seguinte lema.

**Lema 3.16** *O conjunto de factibilidade  $\{\mathcal{K}^0 \cap \mathcal{L}\}$  não é vazio se, e somente se,  $\{\mathcal{K} \cap \mathcal{L}^\perp\} = \{\mathbf{0}\}$ .*

**Prova:** Para todo  $X \in \mathcal{L}$  e  $E \in \mathcal{L}^\perp$ ,  $\langle X, E \rangle = 0$ , donde não é possível que  $X \in \mathcal{K}^0$  e  $E \in \mathcal{K}^0$  simultaneamente, pois não existem  $X > 0$  e  $E > 0$  de tal forma que  $\text{tr}(XE) = 0$ . Além do mais, caso  $X \in \mathcal{K}^0$ , existe uma única matriz  $E \in \mathcal{L}^\perp$ , a matriz nula, de tal forma que  $\text{tr}(XE) = 0$ .  $\square$

Isso significa que, caso o problema (3.75) seja estritamente factível, a interseção do cone  $\mathcal{K}$  com toda e qualquer variedade linear  $(\mathcal{L}^\perp + E)$ ,  $E > 0$ , consiste em um conjunto fechado e limitado, cujo fechamento é composto por pontos pertencentes à fronteira de  $\mathcal{K}$ . Caso contrário o conjunto  $\{\mathcal{K} \cap (\mathcal{L}^\perp + E)\}$  é aberto.

Observando-se que o centro analítico  $\bar{Y}$  desse conjunto, definido como a solução do problema

$$\begin{aligned} &\text{minimizar} && B_{\mathcal{K}}(Y) \\ &\text{s.a.} && Y \in \{\mathcal{K}^0 \cap (\mathcal{L}^\perp + E)\} \end{aligned} \quad (3.77)$$

apresenta a seguinte propriedade<sup>6</sup>

$$-P_{\mathcal{L}^\perp} \nabla B_{\mathcal{K}}(\bar{Y}) = P_{\mathcal{L}^\perp} \bar{Y}^{-1} = \mathbf{0} \quad (3.78)$$

<sup>6</sup>Veja seção 2.10.

o ponto  $\bar{Y}^{-1} \in \{\mathcal{K}^0 \cap \mathcal{L}\}$  é factível em relação ao problema (3.75). Esse ponto pode ser determinado conhecendo-se  $\bar{Y} \in \{\mathcal{K}^0 \cap (\mathcal{L}^\perp + E)\}$ , solução do problema (3.77), que, por sua vez, pode ser obtido através do algoritmo 1. Assim, para  $E = Y_0 > 0$ , uma iteração do algoritmo 1 em um ponto  $Y_k \in \{\mathcal{K}^0 \cap (\mathcal{L}^\perp + E)\}$  consiste em calcular a direção

$$H_k = Y_k^{\frac{1}{2}} P_{\mathcal{L}^\perp(Y_k^{-1})} (\mathbf{I}) Y_k^{\frac{1}{2}} \quad (3.79)$$

$$= Y_k^{\frac{1}{2}} (\mathbf{I} - P_{\tilde{\mathcal{L}}_{Y_k}} (\mathbf{I})) Y_k^{\frac{1}{2}} \quad (3.80)$$

e efetuar

$$Y_{k+1} = Y_k + \gamma_k H_k \quad (3.81)$$

$$= Y_k + \gamma_k \left( Y_k - Y_k^{\frac{1}{2}} P_{\tilde{\mathcal{L}}_{Y_k}} (\mathbf{I}) Y_k^{\frac{1}{2}} \right) \quad (3.82)$$

$$= Y_k + \gamma_k Y_k \left( Y_k^{-1} - Y_k^{-\frac{1}{2}} P_{\tilde{\mathcal{L}}_{Y_k}} (\mathbf{I}) Y_k^{-\frac{1}{2}} \right) Y_k \quad (3.83)$$

Definindo-se  $X_k = Y_k^{-1}$  e  $X_k^+ = X_k^{\frac{1}{2}} P_{\tilde{\mathcal{L}}(X_k^{-1})} (\mathbf{I}) X_k^{\frac{1}{2}}$ , a expressões anterior pode ser escrita de maneira mais compacta como

$$X_{k+1}^{-1} = X_k^{-1} + \gamma_k X_k^{-1} (X_k - X_k^+) X_k^{-1} \quad (3.84)$$

É interessante observar que a direção  $X_k^{-1} (X_k - X_k^+) X_k^{-1}$  é sempre ortogonal ao subespaço  $\mathcal{L}$  e que a seqüência  $X_k^{-1}$  tende ao infinito caso o problema (3.75) seja infactível, pois, nesse caso, o centro analítico  $\bar{Y}$  se encontra no infinito. Caso  $\mathcal{K}^0 \cap \mathcal{L} \neq \emptyset$ , a função  $B_{\mathcal{K}}(Y)$  é limitada inferiormente e essa seqüência tende a  $\bar{Y} < \infty$ . Conseqüentemente, nunca chega a interceptar o cone  $\mathcal{K}$ , como interpretado em Nemirovskii & Gahinet [16, 15]. Cabe, porém, ressaltar que, embora neste artigo a interpretação geométrica do mecanismo de convergência seja imprecisa, os procedimentos algébricos encontram-se absolutamente corretos. A razão desse engano, crê o autor deste trabalho, provém da natureza dos argumentos sobre os quais Nesterov & Nemirovskii<sup>7</sup> constroem o algoritmo projetivo, quais sejam, a função potencial primal e a *transformação de Legendre* da função de barreira primal. Note-se que em nenhum dos dois artigos, [15, 16], existem referências explícitas ao verdadeiro substrato utilizado na construção das atualizações, mas apenas provas de convergência.

Ora, mas se o objetivo deste procedimento consiste apenas em encontrar um ponto factível, não chega a ser necessário determinar  $\bar{Y}$  exatamente: basta verificar o momento em que a projeção  $X_k^+$  se torna definida positiva, fato que pode ocorrer até mesmo bem distante do ponto  $\bar{Y}$ . De qualquer modo, esse algoritmo garante que, mais cedo ou mais tarde, sempre que  $\bar{Y} < \infty$ , isso efetivamente irá ocorrer. A fim de se obter uma estimativa do número de iterações necessárias para que  $X_k^+$  torne-se positiva, utiliza-se o lema 2.5. Observando-se que o medidor de centralização para o problema (3.77) é dado por

$$\delta(X_k^{-1}) = \delta(Y_k) = \|\mathbf{I} - X_k^{-\frac{1}{2}} X_k^+ X_k^{-\frac{1}{2}}\| \quad (3.85)$$

<sup>7</sup>Veja seção 4.4 em Nesterov & Nemirovskii [14].

e coincide com o raio do elipsóide centrado em  $X_k$  que passa pelo ponto  $X_k^+$ , basta iterar apenas até que  $\delta(X_k^{-1})$  seja menor do que um, ou seja, até que a matriz  $X_k^+$  esteja contida no elipsóide de Dikin centrado em  $X_k$ , sendo, portanto, definida positiva. Note-se que essa consideração é relevante apenas do ponto de vista da análise de complexidade, uma vez que  $X_k^+$  pode se tornar positiva antes mesmo que o elipsóide de Dikin centrado em  $X_k$  intercepte  $\mathcal{L}$ . Novamente em Nemirovski & Gahinet [16], este mecanismo é identificado como o responsável por aumentar “significativamente” as “chances” do elipsóide de Dikin interceptar  $\mathcal{L}$ . Segundo o teorema 3.2 e pelo exposto acima,  $\delta(X_k^{-1})$  permanece maior do que um até que  $X_k^+$  se encontre no interior do elipsóide de Dikin centrado em  $X_k$ , logo

$$\log \det X_{k+1}^{-1} - \log \det X_k^{-1} \geq (1 - \log 2) \geq 0.306 = \theta \quad (3.86)$$

ou, escrito de outra maneira,

$$\det X_k^{-1} \geq e^{k\theta} \det X_0^{-1} \quad (3.87)$$

o que fornece um limitante para o número de iterações

$$\bar{k} \leq \frac{\log \det \bar{Y} - \log \det X_0^{-1}}{0.306} = 3.26 \left( \log \det \bar{Y} - \log \det X_0^{-1} \right) \quad (3.88)$$

Um outro limitante pode ser obtido observando-se que a seguinte desigualdade<sup>8</sup>, válida para todo  $X > 0$ ,

$$\log(\operatorname{tr}(X)/p) \geq (\log \det X)/p \quad (3.89)$$

permite escrever

$$\langle X_k^{-1}, \bar{Y}^{-1} \rangle / p \geq \left[ \det(X_0^{-1} \bar{Y}^{-1}) \right]^{1/p} e^{k\theta/p} \quad (3.90)$$

Como para todo valor de  $k$  os pontos  $X_k^{-1} \in (\mathcal{L}^\perp + E)$  e  $\bar{Y}^{-1} \in \mathcal{L}$ , logo

$$\langle X_k^{-1}, \bar{Y}^{-1} \rangle = \langle X_0^{-1}, \bar{Y}^{-1} \rangle \quad (3.91)$$

então,

$$\begin{aligned} e^{k\theta/p} &\leq \frac{\operatorname{tr}(X_0^{-1} \bar{Y}^{-1}) / p}{\left[ \det(X_0^{-1} \bar{Y}^{-1}) \right]^{1/p}} \\ &\leq \frac{\max_i \lambda_i \left( X_0^{-\frac{1}{2}} \bar{Y}^{-1} X_0^{-\frac{1}{2}} \right)}{\min_i \lambda_i \left( X_0^{-\frac{1}{2}} \bar{Y}^{-1} X_0^{-\frac{1}{2}} \right)} \\ &= \operatorname{cond} \left( X_0^{-\frac{1}{2}} \bar{Y}^{-1} X_0^{-\frac{1}{2}} \right) \end{aligned} \quad (3.92)$$

Nessa desigualdade, o termo final é o chamado *condition number* da matriz  $X_0^{-\frac{1}{2}} \bar{Y}^{-1} X_0^{-\frac{1}{2}}$ . Sabendo-se que é possível impor, sem perda de generalidade,  $X_0^{-1} = \alpha \mathbf{I}$ , esse limitante indica que o número de iterações dependerá basicamente de um bom condicionamento dos autovalores de  $\bar{Y}^{-1}$ , conseqüentemente, da “distância” entre o subespaço  $\mathcal{L}$  e o cone  $\mathcal{K}$ .

<sup>8</sup> Análoga à desigualdade pertinente à função  $\psi(X, Z)$ , definida na seção 3.3.

Nesse sentido, matrizes  $F_i$  quase linearmente dependentes são capazes de aumentar demasiadamente o número de iterações requeridas para o cálculo de um ponto factível. Finalmente, cabe ainda ressaltar que é possível aplicar diretamente o teorema 3.2 na determinação do passo ótimo  $\gamma_k$ , pois o problema (3.77) já se encontra na forma de uma função apenas cujo termo  $\langle C, X \rangle = 0$ .

### 3.4.2 O problema de otimização

O problema de otimização pode ser resolvido mediante uma pequena modificação do esquema discutido na seção anterior. Conforme a seção 2.4, é possível definir matrizes  $C$  e  $G$  tais que produto internos matriciais se igualem a produtos internos vetoriais, isto é, para todo  $X \in \mathcal{L}$

$$\langle C, X \rangle = c'x \quad \text{e} \quad \langle G, X \rangle = g'x \quad (3.93)$$

Observando-se que o valor do objetivo ótimo

$$f^* = f(X^*) = \langle C, X^* \rangle / \langle G, X^* \rangle \quad (3.94)$$

determina um hiperplano em  $\mathbb{R}^{p \times p}$ , pode-se utilizar o mesmo algoritmo da seção anterior para determinar  $X^*$ , bastando que a projeção  $X_k^+$  da matriz  $X_k$  sobre o subespaço  $\mathcal{L}$  seja substituída pela projeção  $X_k^+(f^*)$  da matriz  $X_k$  sobre o subespaço

$$\mathcal{L}(f^*) = \mathcal{L} \cap \{X \mid \langle C - f^* G, X \rangle = 0\} \quad (3.95)$$

O problema é que quase nunca se conhece o valor do objetivo ótimo  $f^*$  *a priori* e deve-se recorrer ao cálculo de limitantes superiores desse valor. Felizmente, a estrutura do algoritmo projetivo permite que isto seja feito de maneira bastante elegante e eficiente. Procedese da seguinte maneira: inicialmente, utiliza-se o algoritmo anterior sem modificações até que se consiga obter  $X_k^+ > 0$ . Determinam-se então, a cada iteração, as matrizes<sup>9</sup>  $C_k \in \mathcal{L}$  e  $G_k \in \mathcal{L}$  de tal forma que, para todo  $X \in \mathcal{L}$ ,

$$\langle C_k, X \rangle_{X_k^{-1}} = c'x \quad \text{e} \quad \langle G_k, X \rangle_{X_k^{-1}} = g'x \quad (3.96)$$

Desse modo, obtido um limitante superior  $f$  para o valor do objetivo ótimo, a projeção da matriz  $X_k$  sobre  $\mathcal{L}(f)$  pode ser facilmente determinada como

$$X_k^+(f) = X_k^+ - \frac{\langle C_k - f G_k, X_k^+ \rangle_{X_k^{-1}}}{\|C_k - f G_k\|_{X_k^{-1}}^2} (C_k - f G_k) \quad (3.97)$$

pois, como  $C_k$  e  $G_k$  pertencem a  $\mathcal{L}$  por construção, a projeção de  $X_k^+ \in \mathcal{L}$  sobre o hiperplano  $\{(C_k - f G_k, X) = 0\}$  pode ser obtida por meio do produto interno de  $X_k^+$  e da matriz  $(C_k - f G_k)$ , ortogonal ao hiperplano.

Entretanto, essa expressão para a matriz projetada  $X_k^+(f)$  torna possível determinar o menor limitante superior de  $f^*$  de tal forma que a matriz  $X_k^+(f)$  seja ainda positiva, ou seja, o menor objetivo associado a um ponto factível que possa ser obtido como projeção

<sup>9</sup>O cálculo destas matrizes é bastante simples uma vez que podem ser obtidas mediante a aplicação do mesmo operador  $P_{\mathcal{L}(X_k^{-1})}$  requerido pela avaliação da projeção  $X_k^+$ .

de  $X_k$  sobre a família de subespaços  $\mathcal{L}(f)$ . Esse valor de  $f$  será denotado por  $f_k^*$ . Devido à estrutura da família de hiperplanos  $\mathcal{L}(f)$ , a matriz  $X_k^+(f_k^*)$  está sempre localizada na fronteira do cone  $\mathcal{K}$ , e um raciocínio similar ao empregado na prova do lema 2.6 fornece o valor de  $f_k^*$ . Esse é o assunto do próximo lema.

**Lema 3.17** *Dados a projeção  $X_k^+ > 0$  de  $X_k$  sobre  $\mathcal{L}$  e  $f_k = \langle C_k, X_k^+ \rangle / \langle G_k, X_k^+ \rangle$ , o menor valor possível de  $f$  de tal forma que a projeção  $X_k^+(f)$  de  $X_k$  sobre o subespaço  $\mathcal{L}(f)$  seja ainda semi-definida positiva é dado por*

$$f_k^* = f_k - \frac{1}{\max_i(\mu_i)} \quad (3.98)$$

em que  $\mu_i$  denota o  $i$ ésimo autovalor generalizado real do feixe de matrizes

$$\begin{bmatrix} \Omega_3 & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix} + \mu \begin{bmatrix} \Omega_2 & \Omega_1^{\frac{1}{2}} \\ \Omega_1^{\frac{1}{2}} & \mathbf{0} \end{bmatrix} \quad (3.99)$$

e

$$\begin{aligned} \Omega_1 &= \|C_k - f_k G_k\|_{X_k^{-1} X_k^+}^2 \\ \Omega_2 &= 2 \langle C_k - f_k G_k, G_k \rangle_{X_k^{-1} X_k^+} - \langle G_k, X_k^+ \rangle_{X_k^{-1}} \langle C_k - f_k G_k \rangle_{X_k^{-1}} \\ \Omega_3 &= \langle G_k, G_k \rangle_{X_k^{-1} X_k^+} - \langle G_k, X_k^+ \rangle_{X_k^{-1}} \langle G_k \rangle_{X_k^{-1}} \end{aligned}$$

**Prova:** Definindo-se  $\mu := 1/(f_k - f)$  e efetuando-se algumas manipulações algébricas, o determinante de  $X_k^+(f)$  pode ser escrito na forma

$$\det X_k^+(\mu) = \frac{\det(\mu^2 \Omega_1 + \mu \Omega_2 + \Omega_3)}{\|\mu C_k + (1 - f_k \mu) G_k\|_{X_k^{-1}}^{2p}}$$

Então, procedendo-se como na prova do lema 2.6, a fim de encontrar a matriz  $X_k^+$  na fronteira do cone  $\mathcal{K}$ , deve-se tentar anular esse determinante. Observando-se que  $\Omega_1 > 0$  é possível resolver essa equação, para valores não nulos do denominador, expandindo-se o numerado como

$$\det(\mu^2 \Omega_1 + \mu \Omega_2 + \Omega_3) = - \det \left( \begin{bmatrix} \Omega_3 + \mu \Omega_2 & \mu \Omega_1^{\frac{1}{2}} \\ \mu \Omega_1^{\frac{1}{2}} & -\mathbf{I} \end{bmatrix} \right) = 0$$

o que fornece o feixe de matrizes apresentado no enunciado do lema. Como  $f_k^*$  deve ser o primeiro valor menor do que  $f_k$  que anula esse determinante, deve-se escolher o maior valor real positivo possível para  $\mu$ , ou seja, o maior autovalor generalizado real positivo.  $\square$

Esclarecido o mecanismo de cálculo de  $X_k^+(f_k^*)$ , formalizar-se-á o algoritmo projetivo.

**Algoritmo 4** *Dado  $X_0 \in \mathcal{K}^0$*

1. Calcular a projeção

$$X_k^+ = X_k^{\frac{1}{2}} P_{\tilde{\mathcal{L}}(X_k^{-1})}(\mathbf{I}) X_k^{\frac{1}{2}} \quad (3.100)$$

2. Caso  $X_k^+ > 0$  ir para o passo 3, caso contrário impor

$$Z_k = X_k - X_k^+ \quad (3.101)$$

e ir para o passo 4.

3. A partir de  $f_k = \langle C_k, X_k^+ \rangle / \langle G_k, X_k^+ \rangle$ , encontrar o valor  $f_k^* < f_k$  de tal forma que a projeção  $X_k^+(f_k^*) \geq 0$  esteja na fronteira do cone  $\mathcal{K}$  e impor

$$Z_k = X_k - X_k^+(f_k^*) \quad (3.102)$$

4. Atualizar  $X_k$

$$X_{k+1}^{-1} = X_k^{-1} + \gamma_k X_k^{-1} Z_k X_k^{-1} \quad (3.103)$$

seleccionando-se o valor do passo  $\gamma_k$  através do teorema 3.4.

5. Caso o critério de parada não esteja satisfeito voltar ao passo 1.

□

Como um primeiro comentário a respeito desse algoritmo, o passo 2 implementa as estratégias apresentadas na seção anterior de busca de um ponto factível, ou seja, ele é capaz de resolver o problema de otimização 3.75 a partir de um ponto  $X_0 \in \mathcal{K}$  não factível.

Neste ponto, deve estar claro que esse algoritmo é conceitualmente equivalente ao algoritmo que resolvia apenas o problema de factibilidade. Ao se calcular a projeção  $X_k^+(f_k^*)$ , ao invés de  $X_k^+$ , pode-se interpretar a atualização de  $X_k^{-1}$  como proveniente do problema (3.77) em que o a variedade linear  $(\mathcal{L}^\perp + X_0^{-1})$  se vê substituída por  $(\mathcal{L}^\perp(f_k^*) + X_k^{-1})$  a cada iteração. É interessante notar que o medidor de centralização deste problema é dado por

$$\delta(X_k^{-1}) = \|\mathbf{I} - X_k^{-\frac{1}{2}} X_k^+(f_k^*) X_k^{-\frac{1}{2}}\| \quad (3.104)$$

e que  $\delta(X_k^{-1}) \geq 1$ , já que  $X_k^+(f_k^*)$  é um ponto na fronteira de  $\mathcal{K}$ . Isso significa, portanto, que as propriedades de convergência obtidas para o problema de factibilidade não são alteradas no caso do problema de otimização. Esta estratégia de cálculo do valor  $f_k^*$  apresenta um desempenho superior à sugerida em Nemirovskii & Gahinet [15]. Nesse artigo,  $f_k^*$  é obtido mediante a seleção da menor raiz da equação quadrática em  $f$

$$\|X_k - X_k^+\|_{X_k^{-1}}^2 + \frac{\langle C_k - f G_k, X_k^+ \rangle_{X_k^{-1}}^2}{\|C_k - f G_k\|_{X_k^{-1}}^2} = 0.99 \quad (3.105)$$

que consiste em escolher  $f_k^*$  na fronteira do elipsóide de Dikin centrado em  $X_k$ , ou seja,  $\delta(X_k^{-1}) = 0.99$ . Esse procedimento falha, por exemplo, quando  $\|X_k - X_k^+\|$  já é maior do que 0.99 obrigando a execução de mais um passo do tipo factibilidade, durante o qual não há decréscimo de  $f$ . O próximo lema resume fatos relativos à complexidade do algoritmo projetivo como um todo.

**Lema 3.18** *Os seguintes fatos relativos ao algoritmo 4 são verdadeiros*

- i) *Dado um ponto  $X_0^f > 0$  não factível, o algoritmo 4 encontra um ponto factível num número de iterações limitado por  $O(p \log \chi)$ , em que  $\chi = \text{cond} \left( (X_0^f)^{-\frac{1}{2}} \bar{Y}^{-1} (X_0^f)^{\frac{1}{2}} \right)$  e  $\bar{Y}^{-1}$  é o centro analítico associado ao conjunto  $\left\{ \mathcal{K}^0 \cap \left( \mathcal{L}^\perp + (X_0^f)^{-1} \right) \right\}$ .*
- ii) *Dados  $\epsilon \in (0, 1)$  e o ponto factível  $X_0^o > 0$ , obtido ou não pelo algoritmo 4, esse algoritmo encontra um ponto  $X_\epsilon$  de tal forma que  $f(X_\epsilon) \leq f^* + \alpha\epsilon$ ,  $\alpha = O(1)$ , em no máximo  $O(p \log(1/\epsilon))$  iterações.*

**Prova:** A prova do primeiro ponto provém diretamente da discussão da seção anterior pois, segundo a inequação (3.92),

$$\bar{k}_f \leq \frac{p}{\theta} \log \text{cond} \left( (X_0^f)^{-\frac{1}{2}} \bar{Y}^{-1} (X_0^f)^{-\frac{1}{2}} \right) = O \left( p \log \text{cond} \left( (X_0^f)^{-\frac{1}{2}} \bar{Y}^{-1} (X_0^f)^{\frac{1}{2}} \right) \right)$$

Quanto ao cálculo do limitante do número de iterações da fase de otimização, como o cálculo de  $f_k^*$  é feito de modo a manter válida a relação  $\delta(X_k^{-1}) \geq 1$ , pode-se novamente lançar mão do item iv) do teorema 3.2 e obter<sup>10</sup>

$$\det X_k^{-1} \geq e^{k\theta} \det(X_0^o)^{-1}$$

Definindo-se a matriz  $X_\epsilon \in \mathcal{L}$  mediante

$$X_\epsilon = (1 - \epsilon)X^* + \epsilon X_0^o$$

ou seja, uma combinação convexa do ponto ótimo  $X^*$  e do ponto factível  $X_0^o$ , torna-se possível determinar uma constante  $\alpha > 0$ , que dependa apenas de  $X^*$  e  $X_0^o$ , de tal forma que  $f(X_\epsilon) \leq f^* + \alpha\epsilon$ . Sabendo-se que o algoritmo 4 prossegue enquanto  $f_k^* \geq f(X_\epsilon)$  e observando-se que  $\langle X_k - X_k^+, X_\epsilon \rangle_{X_k^{-1}} = 0$ , a atualização de  $X_k^{-1}$  no problema de otimização implica

$$\begin{aligned} \langle X_{k+1}^{-1} - X_k^{-1}, X_\epsilon \rangle &= \frac{1}{\gamma_k} \langle X_k^{-1} (X_k - X_k^+(f_k^*)) X_k^{-1}, X_\epsilon \rangle \\ &= \frac{1}{\gamma_k} \langle X_k - X_k^+(f_k^*), X_\epsilon \rangle_{X_k^{-1}} \\ &= \frac{1}{\gamma_k} \langle X_k^+ - X_k^+(f_k^*), X_\epsilon \rangle_{X_k^{-1}} \\ &= \frac{1}{\gamma_k} \frac{\langle C_k - f_k^* G_k, X_k^+ \rangle_{X_k^{-1}}}{\|C_k - f_k^* G_k\|_{X_k^{-1}}^2} \langle C_k - f_k^* G_k, X_\epsilon \rangle_{X_k^{-1}} \\ &= \frac{(f_k - f_k^*) \langle G_k, X_k^+ \rangle_{X_k^{-1}} (f(X_\epsilon) - f_k^*) \langle G_k, X_\epsilon \rangle_{X_k^{-1}}}{\gamma_k \|C_k - f_k^* G_k\|_{X_k^{-1}}^2} \\ &\leq 0 \end{aligned}$$

<sup>10</sup>Para o mesmo valor de  $\theta$  do problema de factibilidade!

pois  $\gamma_k > 0$ ,  $f_k^* \leq f_k$ ,  $\langle G_k, X_k^+ \rangle = g'x_k^+ > 0$  e  $\langle G_k, X_\epsilon \rangle = g'x_\epsilon > 0$ . Isso quer dizer que

$$\langle (X_0^o)^{-1}, X_\epsilon \rangle \geq \langle X_k^{-1}, X_\epsilon \rangle \geq p \left[ \det \left( (X_0^o)^{-1} X_\epsilon \right) \right]^{1/p} e^{k\theta/p}$$

portanto, que se pode definir o limitante

$$\begin{aligned} \bar{k}_o &\leq \frac{p}{\theta} \log \text{cond} \left( (X_0^o)^{-\frac{1}{2}} X_\epsilon (X_0^o)^{-\frac{1}{2}} \right) \\ &\leq \frac{p}{\theta} \log \text{cond} \left( (1 - \epsilon)(X_0^o)^{-\frac{1}{2}} X^*(X_0^o)^{-\frac{1}{2}} + \epsilon \mathbf{I} \right) \\ &\leq \frac{p}{\theta} \log \frac{(1 - \epsilon) \max_i \lambda_i \left( (X_0^o)^{-\frac{1}{2}} X^*(X_0^o)^{-\frac{1}{2}} \right) + \epsilon}{(1 - \epsilon) \min_i \lambda_i \left( (X_0^o)^{-\frac{1}{2}} X^*(X_0^o)^{-\frac{1}{2}} \right) + \epsilon} \\ &\leq \frac{p}{\theta} \log \left[ \left( \frac{1}{\epsilon} - 1 \right) \max_i \lambda_i \left( (X_0^o)^{-\frac{1}{2}} X^*(X_0^o)^{-\frac{1}{2}} \right) + 1 \right] \\ &= O(p \log(1/\epsilon)) \end{aligned}$$

□

Apesar da análise desse lema indicar uma convergência  $O(\sqrt{p})$  maior do que as obtidas para os métodos seguidores de trajetória e primal-dual, na prática, o desempenho do algoritmo projetivo é muito bom graças ao uso de estratégias baseadas no resultado do teorema 3.4, fato que a análise teórica de complexidade não consegue considerar.

Finalmente, o algoritmo projetivo pode ser utilizado também na resolução de outros tipos de problema que não estejam na forma (2.1). Observa-se que o método, em si, não impõe restrição alguma ao vetor objetivo  $g$ , de tal forma que o problema *quase convexo* no formato

$$\begin{aligned} &\text{minimizar} && (c'x + a)/(g'x + b) \\ &\text{s.a.} && \begin{cases} F(x) \geq 0 \\ (g'x + b) > 0 \end{cases} \end{aligned} \quad (3.106)$$

em que  $a$  e  $b$  são escalares, chamado *problema linear-fracional generalizado*, pode ser resolvido sem modificação alguma do algoritmo. De fato, até mesmo o problema também *quase convexo*

$$\begin{aligned} &\text{minimizar} && f \\ &\text{s.a.} && \begin{cases} fG(x) \geq C(x) \\ F(x) \geq 0 \\ G(x) \geq 0 \end{cases} \end{aligned} \quad (3.107)$$

em que  $C(x)$  e  $G(x)$  são combinações afins de matrizes, ou seja, a determinação do menor autovalor generalizado do feixe de matrizes  $(C(x) - fG(x))$  sobre o domínio  $\{F(x) \geq 0, G(x) \geq 0\}$ , de especial interesse para a área de controle, pode ser resolvido pelo algoritmo projetivo. Como discutido no início desta seção, as combinações afins podem ser transformadas em combinações lineares mediante a introdução da variável auxiliar  $\tau$ . Uma vez feito isso, a única modificação necessária à resolução desse problema se refere ao cálculo do valor de  $f_k^*$  e da projeção  $X_k^+(f_k^*)$ . Na seção 4.4 em Nesterov & Nemirovskii [14] sugere-se que isso seja feito por meio de um procedimento *dicotômico*<sup>11</sup>. É interessante notar que essa

<sup>11</sup>A não ser para o caso linear, a introdução desse tipo de procedimento não permite obter limitantes do número de iterações requeridas para o cálculo de  $f_k^*$  e Nesterov & Nemirovskii se referem a essa generalização como um algoritmo *conceitualmente* polinomial. Veja *remark* 4.4.2 em [14].

complicação adicional é necessária devido à desigualdade *não linear*  $fG(x) \geq C(x)$ , à qual não é mais possível associar um subespaço vetorial. Estuda-se a possibilidade de tratar este problema mediante a introdução de uma variável de *folga* matricial  $W$ , que permite obter o problema equivalente

$$\begin{array}{ll} \text{minimizar} & f \\ \text{s.a} & \left\{ \begin{array}{l} fG(x) + W = C(x) \\ F(x) \geq 0 \\ G(x) \geq 0 \\ W \geq 0 \end{array} \right. \end{array} \quad (3.108)$$

de tal forma que se possa recorrer ao conceito de projeção no subespaço linear  $fG(x) + W - C(x) = 0$ .

### 3.5 O problema de factibilidade

Todos os algoritmos abordados neste trabalho, com exceção do algoritmo projetivo, pressupõem a existência de um ponto inicial estritamente factível, embora nada se tenha dito a esse respeito, até este momento. Para os algoritmos seguidores de trajetória, exige-se, adicionalmente, que esse ponto inicial seja quase central. Essa última exigência não é muito difícil de ser atendida pois, mediante o emprego do algoritmo 1 na resolução do problema (2.74), é possível obter um ponto quase central a partir de um ponto inicial factível qualquer. Note-se que o número de iterações requerido por esta etapa inicial de centralização, até que se atinja a região em que  $\delta(X) < 1$ , pode ser obtido por meio de uma análise semelhante à efetuada para o algoritmo de factibilidade projetivo, e resulta no limitante<sup>12</sup>

$$\bar{k} = 3.26 (f_{\beta_0}(x_0) - f_{\beta_0}(\bar{x}(\beta_0))) \quad (3.109)$$

que está diretamente relacionado com a localização do ponto inicial  $x_0$ , o que significa que se deve esperar um maior número de iterações para pontos iniciais próximos à fronteira do cone.

Quanto ao ponto inicial factível, existem diversas maneiras de determiná-lo, a maioria delas envolvendo a resolução de um problema de otimização auxiliar e a introdução de variáveis de folga. Por exemplo, pode-se encontrar um ponto inicial primal por meio do problema

$$\begin{array}{ll} \text{minimizar} & \tau \\ \text{s.a} & F(x) + \tau \mathbf{I} \geq 0 \end{array} \quad (3.110)$$

para o qual se conhece o ponto factível  $\tau = (\epsilon - \min_i \lambda_i(F(x_0)))$ ,  $\epsilon > 0$ , com  $x_0$  escolhido arbitrariamente. Esse problema pode ser resolvido por qualquer um dos métodos já estudados e um ponto factível é obtido no momento em que  $\tau$  se torna negativo. Caso o mínimo desse problema apresente um valor positivo para  $\tau$ , isso significa que o conjunto  $F(x) > 0$  não possui nenhum ponto interior, ou seja, o problema (2.1) é infactível.

No caso dos métodos primais-duais, pontos factíveis podem ser obtidos, por exemplo, mediante a introdução do chamado *M grande* (do inglês, *big M*). Inicialmente, suponha-se

<sup>12</sup>Em termos da função apenas.

ter à mão um ponto inicial primal factível  $x_0$ . A introdução de uma restrição extra ao problema (2.1) na forma

$$\begin{aligned} & \text{minimizar} && c'x \\ & \text{s.a.} && \begin{cases} F(x) \geq 0 \\ \text{tr}(F(x)) \leq M \end{cases} \end{aligned} \quad (3.111)$$

que não altera o seu resultado desde que o valor da constante  $M$  seja “suficientemente” grande, gera o seguinte problema dual

$$\begin{aligned} & \text{minimizar} && \text{tr}((Z - \tau\mathbf{I})F_0) + M\tau \\ & \text{s.a.} && \begin{cases} \text{tr}((Z - \tau\mathbf{I})F_i) = c_i, \text{ para } i = 1, \dots, n \\ Z \geq 0 \\ \tau \geq 0 \end{cases} \end{aligned} \quad (3.112)$$

para o qual um ponto factível pode ser calculado da seguinte maneira: determina-se qualquer uma das soluções do sistema linear subdeterminado

$$\text{tr}(F_i C) = c_i, \quad i = \{1, \dots, n\} \quad (3.113)$$

em que  $C = C'$ . Um ponto inicial factível dual pode ser obtido observando-se que  $\tau_0 = (-\min\{\lambda_i(C), 0\} + \epsilon)$ ,  $\epsilon > 0$  e  $Z_0 = (C + \tau_0\mathbf{I})$  implicam  $Z_0 \geq 0$ ,  $\tau_0 \geq 0$  e

$$\text{tr}((Z_0 - \tau_0\mathbf{I})F_i) = \text{tr}(F_i C) = c_i, \quad i = \{1, \dots, n\} \quad (3.114)$$

Sem dúvida, um ponto sensível associado a essa formulação consiste na escolha do valor “suficientemente” grande da constante  $M$ . Assim, a solução desse par de problemas somente coincidirá com a solução dos problemas (2.1) e (2.8) caso  $\tau = 0$ , o que ocorre, segundo a teoria dos multiplicadores de Lagrange, apenas se a restrição  $\text{tr}(F(x)) < M$  estiver inativa, ou seja, se  $M \geq F(x^*)$ . Note-se, entretanto, que a presença do ponto inicial primal factível sempre permite calcular um valor adequado para  $M$ .

A versão dual desse par de problemas auxiliares consiste em

$$\begin{aligned} & \text{minimizar} && c'x + M\tau \\ & \text{s.a.} && \begin{cases} F(x) + \tau\mathbf{I} \geq 0 \\ \tau \geq 0 \end{cases} \end{aligned} \quad (3.115)$$

e

$$\begin{aligned} & \text{minimizar} && \text{tr}(ZF_0) \\ & \text{s.a.} && \begin{cases} \text{tr}(ZF_i) = c_i, \text{ para } i = 1, \dots, n \\ \text{tr}(Z) + \rho = M \\ Z \geq 0 \\ \rho \geq 0 \end{cases} \end{aligned} \quad (3.116)$$

que pode ser escrito sem a variável de folga  $\rho$  como

$$\begin{aligned} & \text{minimizar} && \text{tr}(ZF_0) \\ & \text{s.a.} && \begin{cases} \text{tr}(ZF_i) = c_i, \text{ para } i = 1, \dots, n \\ \text{tr}(Z) \leq M \\ Z \geq 0 \end{cases} \end{aligned} \quad (3.117)$$

para os quais o ponto primal  $x_0$ ,  $\tau_0 = (-\min\{\lambda_i(F(x_0)), 0\} + \epsilon)$ ,  $\epsilon > 0$ , com  $x_0$  escolhido arbitrariamente, é factível. Essa formulação é útil apenas quando se possui um ponto inicial factível dual.

Finalmente, caso pontos iniciais factíveis primais e duais não estejam disponíveis, ainda assim é possível utilizar um procedimento semelhante aos anteriores, mediante a introdução de duas constantes nos problemas

$$\begin{aligned} & \text{minimizar} && c'x + M_d\tau_p \\ & \text{s.a.} && \begin{cases} F(x) + \tau_p\mathbf{I} \geq 0 \\ \text{tr}(F(x)) \leq M_p \\ \tau_p \geq 0 \end{cases} \end{aligned} \quad (3.118)$$

e

$$\begin{aligned} & \text{minimizar} && \text{tr}(ZF_0) + M_p\tau_d \\ & \text{s.a.} && \begin{cases} \text{tr}((Z - \tau_d\mathbf{I})F_i) = c_i, \text{ para } i = 1, \dots, n \\ \text{tr}(Z) \leq M_d \\ Z \geq 0 \\ \tau_d \geq 0 \end{cases} \end{aligned} \quad (3.119)$$

Para esses problemas, caso não seja possível obter limitantes *a priori* para os valores das constantes que garantam a inatividade das restrições auxiliares recém introduzidas, deve-se submeter o par solução à seguinte análise:

- Se  $\tau_p = 0$ ,  $\tau_d = 0$ ,  $\text{tr}(F(x)) < M_p$  e  $\text{tr}(Z) < M_d$ , as soluções encontradas coincidem com as soluções de (2.1) e (2.8).
- Se  $\tau_p > 0$  e  $\text{tr}(Z) = M_d$ , deve-se aumentar o valor de  $M_d$  e obter novas soluções. Proceder-se da mesma maneira em relação a  $\tau_d > 0$  e  $\text{tr}(F(x)) = M_p$ .

Caso os problemas originais (2.1) e (2.8) não sejam estritamente factíveis, podem ocorrer situações em que  $\tau_p = 0$  e  $\text{tr}(Z) = M_d$ , ou  $\tau_d = 0$  e  $\text{tr}(F(x)) = M_p$ .

Finalizando, o problema de factibilidade admite ainda outras formulações (veja, por exemplo, a seção 3.2 em Nesterov & Nemirovskii [14]), as quais fornecem interessante material para discussão. O tratamento detalhado desses problemas foge, entretanto, ao escopo deste trabalho.

## Capítulo 4

# Algoritmos: aspectos práticos

Neste capítulo são abordados diversos aspectos referentes à implementação dos algoritmos discutidos no capítulo anterior, visando adequá-los às estruturas normalmente presentes em problemas de controle. Diversos procedimentos pertinentes a esses algoritmos são analisados levando-se em conta que o conjunto de restrições consiste em  $N$  blocos de LMI de dimensão  $m \times m$  empilhados diagonalmente, o que resulta numa LMI de dimensão  $p = Nm$ . Uma justificativa do porquê desta estrutura assaz particular pode ser encontrada na seção 5.2 do capítulo 5.

Observando-se a estrutura do algoritmo 1, núcleo básico de todos os métodos estudados neste trabalho, nota-se que o custo computacional por iteração está basicamente dividido entre os itens

- cálculo do operador projeção  $P_{\tilde{\mathcal{L}}}$  e
- avaliação do passo  $\gamma_k$ .

Além disso, no caso específico do algoritmo projetivo, deve-se analisar também o cálculo de  $f_k^*$ . São esses os três temas deste capítulo. Apresentam-se, ainda, os resultados de experimentos numéricos envolvendo algumas dessas questões.

### 4.1 O cálculo das projeções

Segundo as seções 2.5 e 2.6, a operação mais custosa envolvida na avaliação do operador projeção consiste na construção e inversão da matriz simétrica definida positiva  $\tilde{M}$ . É interessante perceber que a matriz  $\tilde{M}$  tem dimensão  $n \times n$ , e sua inversão depende, portanto, somente do número de variáveis do problema (2.1). Por sua vez, o processo de construção de  $\tilde{M}$  exige a avaliação de componentes na forma

$$\tilde{M}_{i,j} = \tilde{M}_{j,i} = \left\langle X_k^{-\frac{1}{2}} F_j X_k^{-\frac{1}{2}}, X_k^{-\frac{1}{2}} F_i X_k^{-\frac{1}{2}} \right\rangle \quad (4.1)$$

que dependem somente da dimensão  $p = Nm$  do conjunto de restrições LMI. Como será visto posteriormente, a avaliação individual de cada componente da matriz  $\tilde{M}$  consiste em uma operação vantajosa do ponto de vista computacional, pois permite explorar de maneira muitíssimo eficiente várias estruturas particulares das matrizes básicas  $F_i$ 's, principalmente

quando as variáveis do problema são matriciais. Entretanto, devido à presença da função de barreira, a matriz  $\tilde{M}$  torna-se mal condicionada próximo à fronteira do cone  $\mathcal{K}$ , o que acarreta problemas numéricos no processo de inversão. Isso pode vir a ser um problema quando se exige da solução um grau muito grande de precisão<sup>1</sup>, o que, muitas vezes, implica proximidade excessiva da fronteira do cone  $\mathcal{K}$ . Nessas situações, pode-se tentar contornar esses problemas por meio da chamada fatoração QR, que apresenta um comportamento numérico mais estável, mas que não permite explorar satisfatoriamente a estrutura das matrizes básicas. Para efetuar a fatoração QR, segundo a discussão da seção 2.6, aplica-se a transformação simétrica definida por  $X_k^{-1}$  às matrizes básicas e constrói-se a matriz  $\mathcal{A} \in \mathfrak{R}^{(Nm(m+1)/2) \times n}$

$$\mathcal{A} = \left[ \text{vetor}(\tilde{F}_1) \cdots \text{vetor}(\tilde{F}_n) \right] \quad (4.2)$$

em que  $\text{vetor}(\cdot)$  transforma a matriz simétrica  $(\cdot)$  constituída de  $N$  blocos de dimensão  $m$  em um vetor<sup>2</sup>. Determinam-se, então, as matrizes  $Q$ , ortonormal, e  $R$ , diagonal superior, na forma

$$Q\mathcal{A} = \begin{bmatrix} R \\ \mathbf{0} \end{bmatrix} \quad (4.3)$$

de tal forma que a projeção de uma matriz<sup>3</sup>  $B$  sobre  $\tilde{\mathcal{L}}$  seja dada por

$$P_{\tilde{\mathcal{L}}}(B) = R^{-1}B_1 \quad (4.4)$$

em que

$$B = [B'_1 \ B'_2]' = Q \text{vetor} \left( X_k^{-\frac{1}{2}} B X_k^{-\frac{1}{2}} \right) \quad (4.5)$$

Esse procedimento torna-se mais eficiente caso as transformações simétricas sejam avaliadas para os fatores de Cholesky da matriz  $X_k^{-1}$  e não para sua raiz quadrada, isto é, caso se obtenha o operador  $P_{\tilde{\mathcal{L}}_{X_k^{-1}}} = P_{\tilde{\mathcal{L}}_{(X_k^{-1})}}$ . De qualquer maneira, os custos associados a cada etapa desse procedimento são

- obtenção de  $\mathcal{A}$ :  $O(Nm^3n)$
- obtenção de  $Q$  e  $R$ :  $O(Nm^2n^2/2)$

Esses custos devem ser comparados aos custos associados à obtenção do operador projeção por meio da construção da matriz  $M$ . Para tal, calculando-se previamente as  $n$  matrizes  $\tilde{F}_i$  em  $O(Nm^3)$  operações, é necessário obter  $n(n+1)/2$  componentes a um custo de  $O(Nm(m+1)/2)$  operações por componente. Deve-se, em seguida, calcular a inversa de  $M$ , que pode ser obtida via fatoração de Cholesky em  $O(n^3/6)$  operações. Assim, os custos de cada etapa são

- obtenção das matrizes  $\tilde{F}_i$ :  $O(Nm^3n)$
- obtenção de  $M$ :  $O(Nm^2n^2/4)$

<sup>1</sup>Nemirovskii & Gahinet reportam em [16] que isso ocorre, em geral, quando se exige uma precisão maior do que 1-5% em relação ao valor objetivo ótimo.

<sup>2</sup>Cujos elementos fora da diagonal são devidamente ponderados a fim de que se mantenha a equivalência entre os produtos internos vetorial e matricial.

<sup>3</sup>Geralmente o gradiente de uma das funções auxiliares.

- obtenção de  $R'R = M$ :  $O(n^3/6)$

A vantagem da segunda abordagem sobre a primeira é revelada ao se analisar estruturas particulares para as LMI. Como exemplo, apresenta-se uma LMI frequentemente encontrada em problemas de controle<sup>4</sup>, a inequação de Lyapunov generalizada, dada por

$$AXB' + BXA' \leq 0 \quad (4.6)$$

em que as matrizes  $A$  e  $B$  e a variável matricial quadrada e simétrica  $X$  têm dimensão  $q \times q$ , logo,  $n = q(q+1)/2$  e  $m = q$ . Calculam-se os custos referentes à LMI resultante do empilhamento na diagonal de  $N$  blocos de inequações de Lyapunov, o que fornece  $p = Nm = Nq$ . Nesse caso, substituindo-se as dimensões das matrizes  $F_i$  associadas a essa LMI nos cálculos anteriores, obtêm-se os seguintes custos

- fatoração QR:  $O(Nq^5/2) + O(Nq^6/8)$
- montagem e Cholesky:  $O(Nq^5/2) + O(Nq^6/16) + O(q^6/48)$

Entretanto, a natureza matricial da variável  $X$  permite adotar uma representação *estruturada* para as matrizes  $F_i$ , o que consiste, basicamente, em se utilizar bases matriciais para as variáveis do problema. Por exemplo, uma representação estruturada da inequação de Lyapunov é dada por

$$AXB' + BXA' = \sum_{i=1}^{q(q+1)/2} (AX_i B' + BX_i A') x_i \quad (4.7)$$

em que  $X_i$  são as bases matriciais canônicas da variável  $X$ . Nessa representação, o cálculo de um componente da matriz  $\tilde{M}$  consiste em efetuar

$$\tilde{M}_{i,j} = \left\langle X_k^{-\frac{1}{2}} (AX_i B' + BX_i A') X_k^{-\frac{1}{2}}, X_k^{-\frac{1}{2}} (AX_j B' + BX_j A') X_k^{-\frac{1}{2}} \right\rangle \quad (4.8)$$

$$= 2 \left( \text{tr} \left( X_i (B' X_k^{-1} A) X_j (B' X_k^{-1} A) \right) + \text{tr} \left( X_i (B' X_k^{-1} B) X_j (A' X_k^{-1} A) \right) \right) \quad (4.9)$$

Por sua vez, esse cálculo pode ser efetuado de modo extremamente eficiente representando-se cada base canônica matricial  $X_i$  como o produto de bases canônicas vetoriais em  $\mathfrak{R}^q$ , isto é

$$X_i = \sum_{k=1}^{\alpha_i} s_{ik} t'_{ik} \quad (4.10)$$

Nesse caso, em que  $X$  é simétrica,  $\alpha_i$  é igual a dois para as bases dos elementos simétricos fora da diagonal e igual a um para as bases da diagonal principal. Isso tudo permite calcular cada um dos  $n(n+1)/2$  componentes de  $\tilde{M}$  como

$$\tilde{M}_{i,j} = 2 \sum_{k=1}^{\alpha_i} \sum_{l=1}^{\alpha_j} \left( t'_{ik} (B' X_k^{-1} A) s_{jl} t'_{jl} (B' X_k^{-1} A) s_{ik} + t'_{ik} (B' X_k^{-1} B) s_{ij} t'_{ij} (A' X_k^{-1} A) s_{ij} \right) \quad (4.11)$$

Note-se a tremenda economia obtida, uma vez que o cálculo do produto  $t'_{ik} Y s_{jl}$  reduz-se à seleção de um único componente da matriz  $Y$ , que pode ser previamente calculada e utilizada na obtenção de todos os componentes de  $\tilde{M}$ . Assim, os custos para a montagem estruturada no caso da inequação de Lyapunov são

<sup>4</sup>Veja capítulo 5.

- Cálculo das  $N$  matrizes  $(B'X_k^{-1}A)$ ,  $(B'X_k^{-1}B)$  e  $(A'X_k^{-1}A)$ :  $O(Nq^3)$
- Obtenção de  $\tilde{M}$ :  $O(Nq^4)$

o que fornece um custo total para a montagem do operador projeção, incluindo-se a decomposição de Cholesky, de  $O(Nq^3)+O(Nq^4)+O(q^6/48)$ . Observe-se o ganho obtido em relação à fatoração QR, principalmente no que se refere ao aumento do número de blocos diagonais  $N$ , pois o processo de montagem estruturada consegue diminuir em duas ordens de grandeza o custo das operações dependentes de  $N$ . Cabe, ainda, ser lembrado que o custo da montagem estruturada no caso de variáveis escalares, recai nos valores calculados inicialmente, os quais são da mesma ordem de grandeza que os custos da fatoração QR; percebe-se que o cálculo das bases matriciais  $X_i$  consegue levar em conta estruturas escalares como um caso particular. Infelizmente, como comentado anteriormente, caso seja necessário obter soluções com alto grau de precisão, deve-se recorrer à fatoração QR para obtenção das projeções das últimas iterações, pois o cálculo das matrizes Q e R é, em geral, um procedimento mais estável numericamente do que a fatoração de Cholesky da matriz  $M$ .

Como desvantagem, a representação estruturada exige um maior esforço de programação, pois é necessário prover uma maneira estruturada de cálculo dos componentes da matriz  $\tilde{M}$  para as mais diversas estruturas possíveis. Nesse sentido, uma boa interface é capaz de obter os dados do usuário em uma representação já estruturada.

Finalmente, conforme a seção 3.1.2, o operador projeção pode ser sempre interpretado como a direção de Newton de uma função apenas e, portanto, torna-se possível calculá-lo por meio de métodos *indiretos* como, por exemplo, o método dos gradientes conjugados. O caráter iterativo desses métodos abre espaço para a introdução de aproximações no cálculo das projeções, que também pode contribuir para a diminuição da complexidade computacional por iteração. Um exemplo desse tipo de abordagem aplicado a desigualdades de Lyapunov pode ser encontrado em Vandenberghe & Boyd [23].

## 4.2 O cálculo do passo ótimo

Todos os algoritmos abordados no capítulo anterior podem tirar proveito do resultado do teorema 3.4, especialmente os algoritmos projetivo e primal-dual, cujas trajetórias não devem se manter, necessariamente, próximas à trajetória central. Para esses métodos, o uso do passo ótimo chega a ser a principal justificativa de bons desempenhos. *Alargar* o passo significa reduzir o número de iterações do algoritmo 1, o que só constitui vantagem caso o custo numérico associado ao cálculo do passo ótimo seja inferior ao requerido pelo cálculo das projeções, assunto da seção anterior. Entretanto, à primeira vista, essa afirmação não soa verdadeira uma vez que a avaliação de

$$\mu^* = \frac{1}{\gamma^*} = \lambda_M(\Gamma) \quad (4.12)$$

exige que se determine o maior autovalor da matriz  $\Gamma$ , de dimensão  $Nm$  e sem estrutura bloco diagonal. Isso demanda  $O(N^3m^3)$  operações aritméticas, o que representa um custo superior ao custo do cálculo da projeção no que diz respeito ao fator  $N$ . Como  $N$  é, em geral, bastante grande, necessita-se de um modo mais eficiente de cálculo desse autovalor.

De fato, para se calcular  $\gamma^*$  basta apenas determinar o maior autovalor de  $\Gamma$ , ou seja, resolver o problema

$$\begin{aligned} \mu^* &= \min \mu \\ \text{s.a. } &\Gamma = \Lambda\Lambda'/\delta^2(X) - \text{diag}(\Lambda) \leq \mu\mathbf{I} \end{aligned} \quad (4.13)$$

Ora, o complemento de Schur<sup>5</sup> permite concluir que esse problema é convexo no intervalo  $\mu\mathbf{I} > -\text{diag}(\Lambda)$  e equivale a

$$\begin{aligned} \text{minimizar } &\mu \\ \text{s.a. } &\begin{bmatrix} \mu\mathbf{I} + \text{diag}(\Lambda) & \Lambda/\delta(X) \\ \Lambda'/\delta(X) & 1 \end{bmatrix} \geq 0 \end{aligned} \quad (4.14)$$

ou, escrito de modo diferente, ao problema

$$\begin{aligned} \text{minimizar } &\mu \\ \text{s.a. } &\begin{cases} f(\mu) = \Lambda'[\mu\mathbf{I} + \text{diag}(\Lambda)]^{-1}\Lambda/\delta^2(X) - 1 \leq 0 \\ \mu > -\lambda_m(\Psi(H)) \end{cases} \end{aligned} \quad (4.15)$$

Observando-se que

$$\lim_{\mu \rightarrow -\lambda_m(\Psi(H))^+} f(\mu) = +\infty \text{ e } \lim_{\mu \rightarrow \infty} f(\mu) = -1 \quad (4.16)$$

e que  $f(\mu)$  é convexa no intervalo  $\{\mu \mid \mu > -\lambda_m(\Psi(H))\}$ ,  $\mu^*$  deve ser a única raiz do polinômio  $f(\mu)$  nesse intervalo, a qual pode ser encontrada, por exemplo, mediante o emprego do método de Newton. Nesse caso, obtém-se iterações na forma

$$\mu_{k+1} = \mu_k - \frac{f(\mu)}{f'(\mu)} \quad (4.17)$$

em que

$$f'(\mu) = -\Lambda'[\mu\mathbf{I} + \text{diag}(\Lambda)]^{-2}\Lambda/\delta^2(X) < 0 \quad (4.18)$$

Como  $f'(\mu)$  é negativa e a função  $f(\mu)$  é convexa no intervalo de interesse, caso se inicie este procedimento num ponto  $-\lambda_m(\Psi(H)) < \mu_0 < \mu^*$  tal que  $f(\mu)$  seja positiva, a equação (4.17) garante que todos os pontos obtidos serão maiores do que  $\mu_0$ , portanto, factíveis. Um ponto inicial com tal propriedade pode ser obtido com o auxílio do seguinte limitante inferior da função  $f(\mu)$

$$\frac{1}{\delta^2(X)} \frac{\lambda_m^2(\Psi(H))}{\mu + \lambda_m(\Psi(H))} - 1 \leq f(\mu) \quad (4.19)$$

que é obtido tomando-se apenas o termo relativo ao autovalor  $\lambda_m(\Psi(H))$  no produto  $\Lambda'[\mu\mathbf{I} + \text{diag}(\Lambda)]^{-1}\Lambda/\delta^2(X)$ . Note-se que isto é verdade pois todos os termos desprezados são positivos no intervalo em questão. Ora, se o limitante inferior de  $f(\mu)$  for positivo, este também o será, portanto,  $\mu_0$  deve pertencer ao intervalo

$$-\lambda_m(\Psi(H)) < \mu_0 \leq \frac{\lambda_m^2(\Psi(H))}{\delta^2(X)} - \lambda_m(\Psi(H)) \quad (4.20)$$

Note-se também que  $\lambda_m^2(\Psi(H))/\delta^2(X) < 1$  na direção de Newton e, portanto,  $\mu_0 < \mu^+ = 1 - \lambda_m(\Psi(H))$ , fornecido pelo corolário 3.1.4.

<sup>5</sup>Veja lema 5.1.

A vantagem obtida ao se calcular  $\gamma^*$  dessa maneira é que a maior parte do esforço computacional fica associado ao cálculo da matriz  $\Lambda$ , ou seja, ao cálculo dos autovalores de  $\Psi(H)$ , obtidos em  $O(Nm^3)$  operações aritméticas. Daí em diante, o método de Newton descrito acima requer apenas o cálculo de  $f(\mu)$  e de sua derivada primeira, que podem ser feitos em  $O(Nm)$  operações. Assim, com a introdução de um critério de parada adequado, pode-se dizer que o número de iterações deste procedimento fica limitado por  $O(1)$  iterações, o que implica complexidade total de  $O(Nm^3)$  operações, inferior, portanto, ao requerido pelo cálculo das projeções analisadas na seção anterior. Percebe-se ainda que este procedimento é completamente equivalente ao método de Newton aplicado à resolução do problema de busca unidimensional

$$\min_{\gamma} f_{\beta}(X + \gamma H) \quad (4.21)$$

o que significa que os resultados do teorema 3.2 continuam válidos e um critério de parada adequado do ponto de vista computacional pode ser construído baseado na avaliação do escalar  $\delta(\alpha_k)$ , definido como no teorema 3.2. De fato, embora se tenha conseguido um custo computacional adequado mediante o emprego do método de Newton como auxiliar na busca das raízes da condição de otimalidade (3.13), mediante o emprego do lema 2.7 pode-se escrever

$$f_{\beta}(X + \gamma H) = f_{\beta}(X) - \sum_{i=1}^p \log [1 + \gamma \lambda_i(\Psi(H))] + \gamma \langle \beta C, H \rangle \quad (4.22)$$

o que implica que o cálculo das derivadas primeira e segunda de  $f_{\beta}(X + \gamma H)$  pode também ser efetuado em  $O(Nm)$  operações, uma vez obtidos os autovalores de  $\Psi(H)$ . Na realidade, as iterações em ambos os problemas são idênticas.

Como todos esses procedimentos são derivados do teorema 3.4, pode-se aplicá-los diretamente a todos os métodos baseados na função apenada e, por extensão, aos métodos baseados na função central. Já para os algoritmos derivados da função potencial, são requeridas algumas pequenas modificações, que são introduzidas no restante desta seção para o método redutor de potencial primal-dual. Resultados semelhantes para os demais métodos potenciais podem ser obtidos a partir das expressões apresentadas com quase nenhum esforço algébrico.

Como discutido no final da seção 3.3, deve-se resolver o problema quase convexo<sup>6</sup>

$$\min_{\gamma} f_{\mu}(W + \gamma H_W) \quad (4.23)$$

Utiliza-se nesse caso a mesma estratégia que permitiu a obtenção da direção de decrescimento  $H_W$ : aplicar o método de Newton à função com o termo côncavo linearizado  $fl_{\mu}(W + \gamma H_W)$ , definida de maneira análoga a (3.67), reavaliando-a a cada iteração. A derivada primeira dessa função, que deve coincidir com a derivada primeira<sup>7</sup> da função  $f_{\mu}(W + \gamma H_W)$ , é dada por

$$\begin{aligned} Dfl_{\mu}(W, \gamma) &= \frac{df_{\mu}(W + \gamma H_W)}{d\gamma} \\ &= \frac{\alpha_W(p + \mu)}{\alpha + \gamma \alpha_W} - \sum_{i=1}^p \frac{\lambda_i(\Psi(H_X))}{1 + \gamma \lambda_i(\Psi(H_X))} - \sum_{i=1}^p \frac{\lambda_i(\Psi(H_Z))}{1 + \gamma \lambda_i(\Psi(H_Z))} \end{aligned} \quad (4.24)$$

<sup>6</sup>Pois a função côncava é monotônica em relação de  $\gamma$ .

<sup>7</sup>A notação em função de  $W$  visa apenas realçar que esta derivada depende da linearização do termo côncavo de  $f_{\mu}(W + \gamma H_W)$  no ponto  $(W + \gamma H_W)$ .

enquanto a derivada segunda vale

$$\begin{aligned} D^2 fl_\mu(\gamma) &= \frac{d^2 fl_\mu(W + \gamma H_W)}{d\gamma^2} \\ &= \sum_{i=1}^p \frac{\lambda_i^2(\Psi(H_X))}{(1 + \gamma \lambda_i(\Psi(H_X)))^2} + \sum_{i=1}^p \frac{\lambda_i^2(\Psi(H_Z))}{(1 + \gamma \lambda_i(\Psi(H_Z)))^2} \end{aligned} \quad (4.25)$$

em que

$$\alpha = \langle X, Z \rangle \text{ e } \alpha_W = \langle Z, H_X \rangle + \langle X, H_Z \rangle \quad (4.26)$$

Mais uma vez, caso os autovalores de  $\Psi(H_X)$  e  $\Psi(H_Z)$  estejam disponíveis, essas derivadas podem ser calculadas em  $O(Nm)$  operações aritméticas. As iterações deste procedimento consistem na obtenção da direção de Newton

$$H_k = -D fl_\mu(W, \gamma_k) / D^2 fl_\mu(\gamma_k) \quad (4.27)$$

seguida da atualização na forma

$$\gamma_{k+1} = \gamma_k + H_k \quad (4.28)$$

Caso  $H_X$  e  $H_Z$  sejam direções de Newton da função  $fl_\mu(X, Z)$ , o corolário (3.1.4) permite novamente concentrar a busca por  $\gamma^*$  no intervalo  $\left(0, \frac{1}{1 - \min\{\lambda_m(\Psi(H_X)), \lambda_m(\Psi(H_Z))\}}\right]$ . Mesmo se  $H_X$  ou  $H_Z$  não forem direções de Newton, uma pequena modificação do teorema (3.4) e do corolário (3.1.4) permite limitar a busca ao intervalo  $\left(0, \frac{1}{\sigma - \min\{\lambda_m(\Psi(H_X)), \lambda_m(\Psi(H_Z))\}}\right]$  para o escalar  $\sigma$  definido como a razão entre os módulos das derivadas direcionais segunda e primeira de  $fl_\mu(W)$  tomadas na direção  $H_W$ , isto é

$$\begin{aligned} \sigma &= \frac{D^2 fl_\mu(0)}{|D fl_\mu(W, 0)|} \\ &= \frac{\sum_i^p \lambda_i^2(\Psi(H_X)) + \sum_i^p \lambda_i^2(\Psi(H_Z))}{\sum_i^p \lambda_i(\Psi(H_X)) + \sum_i^p \lambda_i(\Psi(H_Z)) - \alpha_W(p + \mu) / \alpha} \end{aligned} \quad (4.29)$$

Na verdade, os resultados deste teorema e deste corolário não dependem da escolha da direção de Newton, para a qual  $\sigma = \Lambda' \Lambda / \delta^2(X) = 1$ .

No final da seção 3.3 discute-se a possível resolução do problema

$$\min_{\gamma_X, \gamma_Z} f_\mu(X + \gamma_X H_X, Z + \gamma_Z H_Z) \quad (4.30)$$

chamado de busca plana, que explora de maneira bastante proveitosa o relativo desacoplamento entre as direções  $H_X$  e  $H_Z$ . As derivadas da linearização do termo côncavo da função  $f_\mu(X + \gamma_X H_X, Z + \gamma_Z H_Z)$  valem

$$\begin{aligned} D fl_\mu(X, Z, \gamma_X) &= \frac{\partial f_\mu(X + \gamma_X H_X, Z + \gamma_Z H_Z)}{\partial \gamma_X} \\ &= \frac{\alpha_X(p + \mu)}{\alpha + \gamma_X \alpha_X + \gamma_Z \alpha_Z} - \sum_{i=1}^p \frac{\lambda_i(\Psi(H_X))}{1 + \gamma_X \lambda_i(\Psi(H_X))} \end{aligned} \quad (4.31)$$

$$\begin{aligned} D fl_\mu(X, Z, \gamma_Z) &= \frac{\partial f_\mu(X + \gamma_X H_X, Z + \gamma_Z H_Z)}{\partial \gamma_Z} \\ &= \frac{\alpha_Z(p + \mu)}{\alpha + \gamma_X \alpha_X + \gamma_Z \alpha_Z} - \sum_{i=1}^p \frac{\lambda_i(\Psi(H_Z))}{1 + \gamma_Z \lambda_i(\Psi(H_Z))} \end{aligned} \quad (4.32)$$

e

$$D^2 fl_\mu(\gamma_X) = \frac{\partial^2 fl_\mu(X + \gamma_X H_X, Z + \gamma_Z H_Z)}{\partial \gamma_X^2} \quad (4.33)$$

$$= \sum_{i=1}^p \frac{\lambda_i^2(\Psi(H_X))}{(1 + \gamma_X \lambda_i(\Psi(H_X)))^2}$$

$$D^2 fl_\mu(\gamma_Z) = \frac{\partial^2 fl_\mu(X + \gamma_X H_X, Z + \gamma_Z H_Z)}{\partial \gamma_Z^2} \quad (4.34)$$

$$= \sum_{i=1}^p \frac{\lambda_i^2(\Psi(H_Z))}{(1 + \gamma_Z \lambda_i(\Psi(H_Z)))^2}$$

em que

$$\alpha = \langle X, Z \rangle, \alpha_X = \langle Z, H_X \rangle \text{ e } \alpha_Z = \langle X, H_Z \rangle \quad (4.35)$$

e são, portanto, também passíveis de serem calculadas em  $O(Nm)$  operações. Mais uma vez, pode-se resolver esse problema de busca plana pelo método de Newton. As iterações consistem no cálculo da direção de Newton

$$H_k = (H_{\gamma_X}, H_{\gamma_Z})_k \quad (4.36)$$

em que

$$(H_{\gamma_X})_k = -Dfl_\mu(X, Z, (\gamma_X)_k) / D^2 fl_\mu((\gamma_X)_k) \quad (4.37)$$

$$(H_{\gamma_Z})_k = -\mathcal{D}fl_\mu(X, Z, (\gamma_Z)_k) / D^2 fl_\mu((\gamma_Z)_k) \quad (4.38)$$

e na avaliação do escalar

$$\delta^2((\gamma_X, \gamma_Z)_k) = D^2 fl_\mu(X, Z, (\gamma_X)_k) (Dfl_\mu(X, Z, (\gamma_X)_k))^2 + D^2 fl_\mu(X, Z, (\gamma_Z)_k) (Dfl_\mu(X, Z, (\gamma_Z)_k))^2 \quad (4.39)$$

seguidos de atualizações na forma

$$(\gamma_X, \gamma_Z)_{k+1} = (\gamma_X, \gamma_Z)_k + \rho_k H_k \quad (4.40)$$

em que  $\rho_k$  é o fator de amortecimento

$$\rho_k := \begin{cases} 1, & \text{se } \delta((\gamma_X, \gamma_Z)_k) < 1 \\ 1/(1 + \delta((\gamma_X, \gamma_Z)_k)), & \text{se } \delta((\gamma_X, \gamma_Z)_k) \geq 1 \end{cases} \quad (4.41)$$

análogo ao passo (3.9). Segundo o teorema 3.2, o emprego desse fator de amortecimento é capaz de garantir a factibilidade dos pontos obtidos por esse procedimento. Note-se também que o uso de direções  $H_X$  e  $H_Z$  que não sejam, ambas, direções de Newton, não implica modificações desse esquema.

### 4.3 O cálculo de $f_k^*$

Na seção 3.4 do capítulo anterior, que versa sobre o algoritmo projetivo, define-se o escalar  $f_k^*$  como

$$f_k^* = f_k - \frac{1}{\max_i \mu_i} \quad (4.42)$$

em que  $\mu_i$  denota o  $i$ ésimo autovalor generalizado real do feixe de matrizes

$$\begin{bmatrix} \Omega_3 & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix} + \mu \begin{bmatrix} \Omega_2 & \Omega_1^{\frac{1}{2}} \\ \Omega_1^{\frac{1}{2}} & \mathbf{0} \end{bmatrix} \quad (4.43)$$

Mais uma vez, depara-se com um cálculo que não respeita uma estrutura bloco diagonal, nesse caso devido à presença da raiz quadrada da matriz  $\Omega_1$  fora da diagonal. Isso faz com que este procedimento de cálculo de  $f_k^*$  tenha um custo de  $O(N^3 m^3)$  operações aritméticas, que é maior do que o custo associado ao cálculo das projeções. No entanto, observando-se atentamente a equação que deu origem a este problema de autovalores

$$\det(\mu^2 \Omega_1 + \mu \Omega_2 + \Omega_3) = 0 \quad (4.44)$$

nota-se que  $\Omega_1$ ,  $\Omega_2$  e  $\Omega_3$  apresentam estrutura bloco-diagonal o que significa que esse determinante pode ser calculado como o produto dos determinantes de cada bloco

$$\det(\mu^2 \Omega_1 + \mu \Omega_2 + \Omega_3) = \prod_i^N \det(\mu^2(\Omega_1)_i + \mu(\Omega_2)_i + (\Omega_3)_i) = 0 \quad (4.45)$$

O cálculo da maior raiz dessa equação requer apenas a avaliação dos  $N$  maiores autovalores de blocos de matrizes de dimensão  $2m$ , o que pode ser feito em  $O(Nm^3)$  operações aritméticas.

### 4.4 Alguns experimentos computacionais

A fim de ressaltar a importância do papel exercido pelos procedimentos discutidos neste capítulo sobre taxa de convergência dos algoritmos de pontos interiores, apresentam-se os resultados de uma série de testes computacionais envolvendo o algoritmo projetivo<sup>8</sup>, em quatro diferentes versões. A partir de um ponto inicial factível comum a todas as versões, efetuam-se iterações na forma

$$X_{k+1}^{-1} = X_k^{-1} + \gamma_k X_k^{-1} (X_k - X_k^+(f_k^*)) X_k^{-1} \quad (4.46)$$

Em cada uma das versões, denominadas pelas siglas definidas abaixo, os seguintes esquemas de cálculo do valor de  $\gamma_k$  foram utilizados

- **P1:**  $\gamma_k = 1/(1 + \max_i |\lambda_i(\Psi(H_k))|)$ , como sugerido em Nemirovskii & Gahinet [15],
- **P2:**  $\gamma_k = \gamma_k^+ = 1/(1 - \lambda_m(\Psi(H_k)))$ , como no corolário 3.1.4,

---

<sup>8</sup>Veja seção 3.4.

Dimensões			Número médio de Mflops			
$q$	$p$	$n$	<b>P1</b>	<b>P2</b>	<b>P3</b>	<b>P4</b>
2	2+3	6	0.12	0.10	0.07	0.08
3	3+4	10	0.35	0.30	0.22	0.22
4	4+5	15	0.96	0.84	0.63	0.52
5	5+6	21	2.35	2.09	1.57	1.19
6	6+7	28	4.83	4.33	3.31	2.10
7	7+8	36	9.68	8.79	6.82	3.91
8	8+9	45	17.60	15.95	12.50	7.06
9	9+10	55	32.84	30.31	23.80	12.06
10	10+11	66	56.33	53.17	42.80	20.25

Tabela 4.1: Desempenhos em função do número médio de Mflops

- **P3**:  $\gamma_k = \gamma_k^* = 1/\lambda_M(\Gamma_k)$ , como no teorema 3.4,
- **P4**: como em **P3**.

No algoritmo **P4** o valor de  $f_k^*$  foi calculado por meio da avaliação do máximo autovalor generalizado do feixe de matrizes (3.99) conforme discutido na seção anterior, enquanto nas demais versões utilizou-se o valor de  $f_k^*$  resultante da resolução da equação quadrática (3.105), como sugerido em Gahinet & Nemirovskii [15].

Como problema teste, selecionou-se o problema de controle por realimentação de estados  $\mathcal{H}_2$ , cuja descrição pode ser encontrada na seção 5.2 do próximo capítulo. Fixaram-se os valores das seguintes matrizes

$$B_1 = \mathbf{I}, C = \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix}, D = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} \quad (4.47)$$

de tal forma que a dimensão do vetor de controle fosse  $v = 1$  e  $C'D = \mathbf{0}$  e  $D'D = 1 > 0$ . Em seguida, obtiveram-se pares de matrizes  $(A, B_2)$  cujos componentes foram gerados aleatoriamente segundo uma distribuição Gaussiana com média zero e variância um, para a dimensão  $q$  variando entre dois e dez. Foram gerados vinte pares para cada dimensão, ou seja, cento e oitenta problemas ao todo. Os algoritmos foram inteiramente implementados e executados em MATLAB versão 4.2c, compartilhando um núcleo comum de construção e manipulação das LMI. Como o objetivo principal destes testes consiste em avaliar o papel representado pela introdução das estratégias de cálculo do passo ótimo (teorema 3.4) e de cálculo de  $f_k^*$  (lema 3.17), o cálculo estruturado das projeções discutido na seção 4.1 não foi implementado. Para obtenção de  $\gamma_k$  nos algoritmos **P3** e **P4** utilizou-se o método de Newton discutido na seção anterior. Os resultados, em número médio de Mflops para cada dimensão  $q$ , encontram-se na tabela 4.1. Nessa tabela, as dimensões  $p$  e  $n$  correspondem, respectivamente, à dimensão total da LMI e ao número de variáveis. Os resultados, em termos do número de iterações da fase de otimização, encontram-se na tabela 4.2.

Percebe-se que os cálculos adicionais requeridos pela determinação de  $\gamma_k^*$  e  $f_k^*$  em **P4** reduzem bruscamente o número de iterações necessárias à obtenção de soluções. Ao mesmo

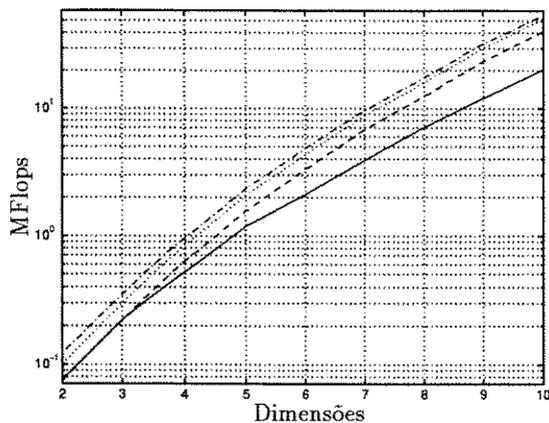


Figura 4.1: Desempenhos em função do número médio de Mflops

$q$	Número médio de iterações			
	P1	P2	P3	P4
2	23.5	19.3	13.8	7.5
3	23.5	19.9	14.5	8.2
4	26.3	23.0	17.0	8.4
5	29.2	25.9	19.2	9.6
6	29.5	26.4	20.0	9.0
7	31.3	28.4	21.8	9.4
8	31.8	28.8	22.4	9.9
9	34.9	32.2	25.0	10.3
10	36.6	34.5	27.6	10.9

Tabela 4.2: Desempenhos em função do número médio de iterações

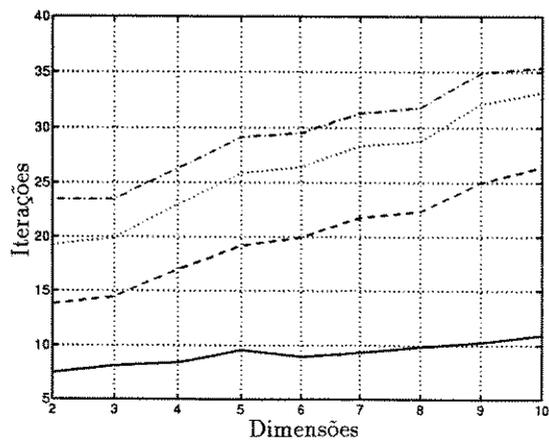


Figura 4.2: Desempenhos em função do número médio de iterações

tempo, pelos dados apresentados na tabela 4.1, confirmam-se as análises das seções anteriores que indicam que estes procedimentos apresentam um baixo custo comparados ao cálculo das projeções, o que se reflete na razão entre os ganhos percentuais em termos do número de flops e em termos do número de iterações, aproximadamente coincidentes. As figuras 4.1 e 4.2 ilustram graficamente este comportamento. Nestas figuras a correspondência entre o tipo das linha e o algoritmo é a seguinte: traço-ponto, **P1**; pontilhada, **P2**; tracejada, **P3**; linha cheia, **P4**.

## Capítulo 5

# Algumas aplicações

Neste último capítulo, na ânsia de convencer o leitor das vantagens da formulação de problemas no formato LMI (o que pretensiosamente justificaria as intrincadas análises apresentadas ao longo dos capítulos anteriores), apresenta-se de maneira rápida e sucinta uma série de conhecidos problemas convexos e suas respectivas formulações LMI. Algumas aplicações provenientes da área de controle são discutidas na segunda seção.

### 5.1 Abrangência da formulação LMI

Nesta seção, diversos problemas de otimização convexos são reescritos no formato LMI. Concentram-se esforços principalmente no tratamento de diferentes conjuntos de restrições, uma vez que todo problema de otimização convexo pode ser reformulado de modo a apresentar uma função objetivo linear. Inicia-se pelas restrições lineares.

Um conjunto de restrições lineares pode ser escrito na forma

$$\{x \mid Ax + b \geq 0\} \quad (5.1)$$

em que  $x \in \mathbb{R}^l$ ,  $A \in \mathbb{R}^{m \times l}$  e  $b \in \mathbb{R}^m$ . Esse conjunto equivale à LMI

$$F(x) = [\text{diag}(Ax + b)] \geq 0 \quad (5.2)$$

cujas matrizes básicas são

$$F_0 = \text{diag}(b) \text{ e } F_i = \text{diag}(A_i) \quad (5.3)$$

em que  $A_i$  denota a  $i$ ésima coluna de  $A$ .

A conversão de conjuntos de restrições mais complexos ao formato LMI dá-se, quase sempre, com o auxílio da *fórmula complementar de Schur*, ou simplesmente *complemento de Schur*, que é assunto do próximo lema.

**Lema 5.1** *Assumindo-se que  $S_1(x) > 0$ , o conjunto*

$$\{x \mid S_1(x) > 0, S_3(x) \geq S_2(x)'S_1(x)^{-1}S_2(x)\} \quad (5.4)$$

*em que  $S_{i=\{1,2 \text{ e } 3\}}$ , matrizes reais de dimensões adequadas, é completamente equivalente ao conjunto descrito pela desigualdade matricial*

$$S(x) = \begin{bmatrix} S_1(x) & S_2(x) \\ S_2(x)' & S_3(x) \end{bmatrix} \geq 0 \quad (5.5)$$

**Prova:** A matriz

$$V(x) = \begin{bmatrix} S_1(x) & \mathbf{0} \\ \mathbf{0} & S_3(x) - S_2(x)'S_1(x)^{-1}S_2(x) \end{bmatrix}$$

é semidefinida positiva se, e somente se,  $S_3(x) \geq S_2(x)'S_1(x)^{-1}S_2(x)$ . Assim, como  $V(x) = T(x)S(x)T(x)'$ , em que

$$T(x) = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -S_2(x)'S_1(x)^{-1} & \mathbf{I} \end{bmatrix}$$

é uma matriz não singular cujos autovalores são todos iguais a um,  $S(x) \geq 0$  se, e somente se,  $V(x) \geq 0$ .  $\square$

Desse modo, fazendo-se uso do lema anterior, é possível escrever uma restrição quadrática convexa

$$\{x \mid c'x + d \geq (Ax + b)'Q(Ax + b)\} \quad (5.6)$$

em que  $x$  e  $c \in \mathfrak{R}^l$ ,  $Q \in \mathfrak{R}^{m \times m}$ ,  $A \in \mathfrak{R}^{m \times l}$ ,  $b \in \mathfrak{R}^m$ ,  $d \in \mathfrak{R}$  e  $Q > 0$ , como a LMI

$$F(x) = \begin{bmatrix} Q^{-1} & Ax + b \\ (Ax + b)' & c'x + d \end{bmatrix} \geq 0 \quad (5.7)$$

cujos componentes básicos são dados por

$$F_0 = \begin{bmatrix} Q^{-1} & b \\ b' & d \end{bmatrix} \text{ e } F_i = \begin{bmatrix} \mathbf{0} & A_i \\ A_i' & c_i \end{bmatrix} \quad (5.8)$$

Segundo esse mesmo raciocínio, o problema não linear convexo

$$\begin{aligned} &\text{minimizar} && (c'x)^2/d'x \\ &\text{s.a.} && \begin{cases} Ax + b > 0 \\ d'x > 0 \end{cases} \end{aligned} \quad (5.9)$$

em que  $x \in \mathfrak{R}^l$ ,  $A \in \mathfrak{R}^{m \times l}$ ,  $b \in \mathfrak{R}^m$ ,  $c$  e  $d \in \mathfrak{R}^l$ , pode também ser escrito como uma LMI. Com a introdução da variável extra  $\tau \in \mathfrak{R}$ , transforma-se esse problema no problema convexo com objetivo linear

$$\begin{aligned} &\text{minimizar} && \tau \\ &\text{s.a.} && \begin{cases} Ax + b > 0 \\ (d'x)\tau > (c'x)^2 \\ \tau > 0 \end{cases} \end{aligned} \quad (5.10)$$

cujos conjunto de restrições é a LMI

$$F(x) = \begin{bmatrix} \text{diag}(Ax + b) & 0 & 0 \\ 0 & \tau & c'x \\ 0 & c'x & d'x \end{bmatrix} > 0 \quad (5.11)$$

que tem por base as matrizes

$$F_0 = \begin{bmatrix} b & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad F_\tau = \begin{bmatrix} \mathbf{0} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (5.12)$$

e

$$F_{i=\{1,\dots,l\}} = \begin{bmatrix} \text{diag}(A_i) & 0 & 0 \\ 0 & 0 & c_i \\ 0 & c_i & d_i \end{bmatrix} \quad (5.13)$$

Obviamente, sabe-se que existem algoritmos<sup>1</sup> que, ao explorar inteligentemente as estruturas particulares dos problemas anteriores, são capazes de resolvê-los com bastante eficiência. A vantagem advinda da conversão ao formato LMI, no entanto, é a possibilidade de se resolverem problemas sujeitos a restrições de diversas naturezas, porém passíveis de conversão a LMI. Para se ilustrar esse fato, toma-se como exemplo um problema discutido em Vandenberghe & Boyd [24]. O problema consiste em separar dois conjuntos de pontos no espaço  $\mathbb{R}^l$ ,  $X_1 = \{x_1, \dots, x_m\}$  e  $X_2 = \{x_{m+1}, \dots, x_n\}$ , por superfícies quadráticas. Pode-se fazer uma analogia entre esse problema e o problema de separação por hiperplanos em função do vetor  $a \in \mathbb{R}^l$  e do escalar  $b$ . Caso existam  $a$  e  $b$  de tal forma que as inequações lineares

$$\begin{cases} x'_i a + b \leq 0 & , \text{ para } i = 1, \dots, m \\ x'_i a + b \geq 0 & , \text{ para } i = m + 1, \dots, n \end{cases} \quad (5.14)$$

sejam satisfeitas, garante-se que o hiperplano  $\{x \mid a'x + b = 0\}$  separa os conjuntos  $X_1$  e  $X_2$ . No caso quadrático, o conjunto de restrições que deve ser atendido para que haja a separação é dado por

$$\begin{cases} x'_i A x_i + x'_i b + c \leq 0 & , \text{ para } i = 1, \dots, m \\ x'_i A x_i + x'_i b + c \geq 0 & , \text{ para } i = m + 1, \dots, n \end{cases} \quad (5.15)$$

em que  $A = A' \in \mathbb{R}^{l \times l}$ ,  $b \in \mathbb{R}^l$  e  $c$  é um escalar. Embora esse conjunto de restrições seja novamente linear nas variáveis  $A$ ,  $b$  e  $c$ , a obtenção de diversas características interessantes para a superfície separadora exige a presença de restrições adicionais não lineares, o que torna adequado um tratamento no formato LMI. É possível, por exemplo, impor a restrição  $A > 0$  de tal forma que a separação seja feita apenas por superfícies elipsoidais, confinando-se os pontos pertencentes a  $X_1$  ao interior do elipsóide separador. Além disso, pode-se adicionar um par de restrições do tipo  $\mathbf{I} \leq A \leq \alpha \mathbf{I}$ , no qual  $\alpha$  limita a razão<sup>2</sup> entre o maior e o menor autovalor de  $A$ , isto é, o *condition number* de  $A$ , que pode ser interpretado como a razão entre o raio maior e o raio menor da elipse separadora no caso de pontos no espaço  $\mathbb{R}^2$ . Eventualmente, pode-se até mesmo considerar  $\alpha$  como variável e minimizá-la, o que resulta no elipsóide separador *mais esférico* possível dos conjuntos  $X_1$  e  $X_2$ .

Retomando-se a exposição, o problema convexo de minimização do autovalor máximo de uma matriz real e simétrica escrita como uma combinação afim  $A(x) = A_0 + \sum_{i=1}^l A_i x_i$  pode ser formulado como

$$\begin{aligned} & \text{minimizar } \tau \\ & \text{s.a. } \tau \mathbf{I} - A(x) \geq 0 \end{aligned} \quad (5.16)$$

que é um problema no formato de (2.1).

Para esta mesma combinação afim de matrizes pode-se formular o problema

$$\text{minimizar } \|A(x)\| = \max_i \sigma_i(A(x)) \quad (5.17)$$

<sup>1</sup>Não obrigatoriamente de pontos interiores.

<sup>2</sup>Pois  $c$  também é uma variável do problema.

em que  $\sigma_i(A(x))$  denota o  $i$ ésimo valor singular da matriz  $A(x)$ . Com a introdução da variável auxiliar  $\tau$  e a ajuda do lema 5.1, consegue-se reescrever esse problema como

$$\begin{aligned} & \text{minimizar } \tau \\ \text{s.a. } & F(x) = \begin{bmatrix} \tau \mathbf{I} & A(x) \\ A(x)' & \tau \mathbf{I} \end{bmatrix} \geq 0 \end{aligned} \quad (5.18)$$

Note-se que no problema de minimização do autovalor máximo os componentes básicos da combinação de matrizes  $A(x)$  devem ser simétricos, enquanto que no problema de minimização da norma isso não é mais obrigatório.

Outros exemplos de problemas convexos (e algumas aproximações para problemas não convexos) formulados como LMI podem ser encontrados em Vandenberghe & Boyd [24] e [5]. Problemas específicos da área de controle que podem ser resolvidos no contexto de LMI podem ser encontrados na próxima seção deste trabalho e em Boyd *et alii* [3] e suas referências.

## 5.2 LMI e problemas de controle

Nesta seção serão discutidos alguns (dentre vários!) problemas de controle que podem ser formulados e resolvidos eficientemente como LMI.

Inicialmente, permita-se que se abra espaço para um breve comentário: é no mínimo surpreendente observar que o teste de determinação da estabilidade de sistemas lineares segundo a teoria de Lyapunov<sup>3</sup>, que acaba de completar um século, consiste em nada menos do que verificar a existência ou não de uma matriz factível num conjunto escrito como uma LMI. Isto é, dado o sistema autônomo, contínuo, linear e invariante no tempo descrito pelo sistema de equações diferenciais lineares

$$\dot{x} = Ax \quad (5.19)$$

em que  $x \in \mathfrak{R}^n$ , esse sistema é estável<sup>4</sup> se, e somente se, existir uma matriz  $P$  que satisfaça estritamente a seguinte LMI

$$\begin{bmatrix} -(A'P + PA) & \mathbf{0} \\ \mathbf{0} & P \end{bmatrix} \geq 0 \quad (5.20)$$

Obviamente as condições de Lyapunov (de fato muito mais abrangentes do que o exposto!) não consistem na resolução explícita da LMI mas, sim, na obtenção de uma relação entre os autovalores de  $A$  e a existência, ou não, de uma solução desse problema.

No restante desta seção discute-se o chamado problema de síntese  $\mathcal{H}_2$  de controladores. Não serão apresentadas provas formais dos resultados, o que pode ser fartamente encontrado nas referências bibliográficas, visto que o objetivo desta apresentação consiste apenas em ilustrar a versatilidade advinda do uso da formulação LMI em problemas de controle. Assim,

<sup>3</sup>Veja por exemplo, Vidyasagar [25].

<sup>4</sup>Para qualquer condição inicial dada, as trajetórias de  $x$  ao longo do tempo convergem sempre para a origem.

dado o sistema não autônomo, contínuo, linear e invariante no tempo descrito pelas equações de estado

$$\begin{cases} \dot{x} &= Ax + Bw \\ z &= Cx \end{cases} \quad (5.21)$$

em que  $x \in \mathbb{R}^n$ ,  $w \in \mathbb{R}^r$  e  $z \in \mathbb{R}^s$ , a norma  $\mathcal{H}_2$  desse sistema, definida somente caso a matriz  $A$  seja estável, é dada por

$$\|H_{wz}\|_2^2 = \frac{1}{\pi} \int_0^\infty \text{tr}(H_{wz}^*(j\omega)H_{wz}(j\omega)) d\omega \quad (5.22)$$

em que  $H_{wz}(s)$  é a função de transferência da entrada  $w$  para a saída  $z$  e  $(\cdot)^*$  denota o complexo conjugado transposto de  $(\cdot)$ . Pode-se provar que este cálculo, efetuado em termos do operador função de transferência  $H_{wz}$  no domínio *freqüencial*, é equivalente ao cálculo no domínio *temporal* de

$$\|H_{wz}\|_2^2 = \int_0^\infty h(t)'h(t)dt \quad (5.23)$$

em que  $h(t)$  é a resposta impulsiva do sistema linear acima. Em decorrência disso e da linearidade do sistema, torna-se possível calcular<sup>5</sup> essa norma mediante a avaliação dos chamados gramianos de observabilidade e controlabilidade, de tal forma que

$$\|H_{wz}\|_2^2 = \text{tr}(B'L_oB) = \text{tr}(CL_cC') \quad (5.24)$$

em que os gramianos  $L_o$  e  $L_c$  são obtidos mediante a resolução das seguintes equações matriciais lineares

$$AL_c + L_cA' + BB' = \mathbf{0} \quad (5.25)$$

$$A'L_o + L_oA + C'C = \mathbf{0} \quad (5.26)$$

Por estas equações percebe-se que, sempre que  $B'B$  e  $C'C$  forem matrizes estritamente definidas positivas, os gramianos satisfazem às inequações de Lyapunov  $(AL_c + L_cA') < 0$  e  $(A'L_o + L_oA) < 0$  e, portanto, conforme a breve discussão apresentada no início desta seção,  $L_c$  e  $L_o$  são matrizes estritamente definidas positivas.

Esse problema pode ser também formulado como um problema linear sujeito a restrições do tipo LMI. Mostra-se, avaliando-se as condições de otimalidade, por exemplo, que as soluções dos problemas

$$\begin{aligned} &\text{minimizar} \quad \langle C'C, P \rangle \\ &\text{s.a.} \quad \begin{cases} AP + PA' + BB' \leq 0 \\ P > 0 \end{cases} \end{aligned} \quad (5.27)$$

e

$$\begin{aligned} &\text{minimizar} \quad \langle BB', Q \rangle \\ &\text{s.a.} \quad \begin{cases} A'Q + QA + C'C \leq 0 \\ Q > 0 \end{cases} \end{aligned} \quad (5.28)$$

são exatamente  $P = L_c$  e  $Q = L_o$ . Essa formulação, visivelmente mais complicada do que a resolução da equação linear matricial anterior, permite, entretanto, que o problema de

<sup>5</sup>Satisfeitas as condições de controlabilidade e observabilidade.

síntese de controladores seja formulado de maneira bastante conveniente. Para o sistema linear descrito pelas equações

$$\begin{cases} \dot{x} = Ax + B_1 w + B_2 u \\ z = Cx + Du \end{cases} \quad (5.29)$$

em que as matrizes  $C$  e  $D$  são ortogonais<sup>6</sup>  $C'D = \mathbf{0}$ ,  $D'D > 0$  e  $B_1 B_1' > 0$ ,  $x \in \mathbb{R}^q$ ,  $w \in \mathbb{R}^r$ ,  $z \in \mathbb{R}^s$  e  $u \in \mathbb{R}^v$ , o problema de realimentação de estados  $\mathcal{H}_2$  pode ser enunciado como

*Dado o sistema linear descrito pelo sistema de equações (5.29), encontrar a matriz de realimentação de estados  $K$  de tal forma que o sistema realimentado pela lei de controle  $u = Kx$  seja estável e tenha a mínima norma  $\|H_{wz}\|_2$ .*

Esta formulação do problema em termos de uma realimentação estática do vetor de estados se justifica uma vez que é possível mostrar que nenhuma outra lei de controle linear ou não linear consegue obter um menor custo  $\mathcal{H}_2$  do que o obtido por meio desta configuração. Desta forma, após a eliminação da entrada  $u$  mediante a substituição da lei de controle, obtém-se o sistema, também linear,

$$\begin{cases} \dot{x} = A_{mf}x + B_1 w \\ z = C_{mf}x \end{cases} \quad (5.30)$$

em que  $A_{mf} = (A + B_2 K)$  e  $C_{mf} = (C + DK)$ . Para esse sistema, o cálculo da norma  $\mathcal{H}_2$  por meio dos gramianos resulta nos seguintes problemas não convexos

$$\begin{aligned} &\text{minimizar} \quad \text{tr}(CPC' + DKPK'D') \\ &\text{s.a.} \quad \begin{cases} AP + PA' + B_2 KP + PK'B_2' + B_1 B_1' \leq 0 \\ P > 0 \end{cases} \end{aligned} \quad (5.31)$$

e

$$\begin{aligned} &\text{minimizar} \quad \text{tr}(B'QB) \\ &\text{s.a.} \quad \begin{cases} A'Q + QA + QB_2K + K'B_2'Q + C'C + K'D'DK \leq 0 \\ Q > 0 \end{cases} \end{aligned} \quad (5.32)$$

A chave da transformação desses problemas em problemas convexos na forma de LMI consiste na mudança de variáveis descrita em Geromel *et alii* [9]. Para o primeiro problema, definem-se as variáveis  $Y = KP$  e  $Z$  de tal forma que seja possível reescrevê-lo como

$$\begin{aligned} &\text{minimizar} \quad \text{tr}(CPC' + DZD') \\ &\text{s.a.} \quad \begin{cases} AP + PA' + B_2 Y + Y'B_2' + B_1 B_1' \leq 0 \\ \begin{bmatrix} P & Y' \\ Y & Z \end{bmatrix} \geq 0 \end{cases} \end{aligned} \quad (5.33)$$

Note-se que a função objetivo foi *linearizada* através da introdução da variável  $Z$  na forma do complemento de Schur  $Z \geq YP^{-1}Y'$ . Esse novo problema é um problema linear com restrições LMI e a prova formal da equivalência pode ser encontrada no citado artigo

<sup>6</sup>Sem perda de generalidade, veja por exemplo Anderson & Moore [1].

e suas referências. Fato de vital importância é que a solução  $(P, Y)$ , caso exista<sup>7</sup>, permite determinar unicamente a matriz de ganho ótima como  $K^* = YP^{-1} = K^*PP^{-1}$ .

Para o segundo problema, antes de se efetuar a troca de variáveis deve-se multiplicar por  $W = Q^{-1} > 0$  e pelos dois lados a primeira restrição, o que resulta numa nova restrição na forma

$$WA' + AW + B_2KW + WK'B_2' + (CW + DKW)'(CW + DKW) \leq 0 \quad (5.34)$$

logo, introduzindo-se as variáveis  $S = KW$  e  $R$  e aplicando-se o lema 5.1, obtém-se o problema equivalente

$$\begin{array}{l} \text{minimizar} \quad \text{tr}(R) \\ \text{s.a.} \quad \left\{ \begin{array}{l} \left[ \begin{array}{cc} WA' + AW + B_2S + S'B_2' & WC' + S'D' \\ CW + DS & -I \end{array} \right] \leq 0 \\ \left[ \begin{array}{cc} R & B' \\ B & W \end{array} \right] \geq 0 \end{array} \right. \end{array} \quad (5.35)$$

e, novamente,  $K^*$  pode ser obtido como  $K^* = SW^{-1}$ .

Pelas discussões anteriores, a solução do problema realimentação de estados  $\mathcal{H}_2$  é de tal forma que

$$\|H_{zw}\|_2^2 = \min_{u(t)} \int_0^\infty x(t)'C'Cx(t) + u(t)'D'Du(t) dt = \text{tr}(B_1L_oB_1') \quad (5.36)$$

portanto, é completamente equivalente à solução do chamado problema linear quadrático<sup>8</sup> no qual as matrizes de ponderação dos estados e das entradas são dadas, respectivamente, por  $C'C$  e  $D'D$ , e a matriz  $B_1$  pode ser identificada com as condições iniciais. Conseqüentemente, existe também a equivalência com o problema de controle de ótimo Gaussiano cujas matrizes de covariância dos ruídos valem  $C'C$  e  $D'D$ . é sabido, portanto, que as soluções de todos esses problemas podem ser obtidas de modo muito eficiente mediante a resolução das chamadas equações de Riccati. Todavia, a grande vantagem da abordagem LMI é que ela torna possível tratar, e com igual facilidade, o chamado problema de realimentação de estados  $\mathcal{H}_2$  com custo garantido. Nesse problema, deseja-se obter uma matriz de realimentação de estados que assegure estabilidade a todos os sistemas lineares resultantes da realimentação de  $N$  sistemas cuja dinâmica seja descrita por equações de estado como

$$\dot{x} = A_i x + B_1 w + (B_2)_i u \quad (5.37)$$

e que, ao mesmo tempo, minimize um critério do tipo  $\mathcal{H}_2$ , que pode ser novamente calculado mediante a introdução da ponderação representada pela “saída”

$$z = Cx + Du \quad (5.38)$$

Elaborando-se um pouco mais: denotando-se por  $\mathcal{K}$  o conjunto de todos os ganhos que estabilizam simultaneamente todos os sistemas em (5.37), por  $\mathcal{P}_A$  o conjunto de todos as

<sup>7</sup>Veja condições nas referências.

<sup>8</sup>Veja novamente Anderson & Moore [1] e Peres & Geromel [19].

matrizes  $A_i$  e por  $\mathcal{P}_{B_2}$  o conjunto de todas as matrizes  $(B_2)_i$  em (5.37), deseja-se encontrar a lei de controle  $u = Kx$  que minimize

$$\begin{aligned} & \inf_i \quad \|(H_{zw})_i\|_2^2 \\ & \text{s.a.} \quad \begin{cases} K \in \mathcal{K} \\ A_i \in \mathcal{P}_A \\ (B_2)_i \in \mathcal{P}_{B_2} \end{cases} \end{aligned} \quad (5.39)$$

Nas referências bibliográficas encontram-se provas de que isso pode ser obtido pela resolução de um dos dois problemas lineares

$$\begin{aligned} & \text{minimizar} \quad \text{tr}(CPC' + DZD') \\ & \text{s.a.} \quad \begin{cases} \text{bl}(A_iP + PA'_i + (B_2)_iY + Y'(B_2)'_i + B_1B'_1)_{i=\{1,\dots,N\}} \leq 0 \\ \begin{bmatrix} P & Y' \\ Y & Z \end{bmatrix} \geq 0 \end{cases} \end{aligned} \quad (5.40)$$

ou

$$\begin{aligned} & \text{minimizar} \quad \text{tr}(R) \\ & \text{s.a.} \quad \begin{cases} \text{bl} \left( \begin{bmatrix} WA'_i + A_iW + (B_2)_iS + S'(B_2)'_i & WC' + S'D' \\ CW + DS & -I \end{bmatrix} \right)_{i=\{1,\dots,N\}} \leq 0 \\ \begin{bmatrix} R & B' \\ B & W \end{bmatrix} \geq 0 \end{cases} \end{aligned} \quad (5.41)$$

em que  $\text{bl}(\cdot)$  representa a matriz resultante da composição das  $i$  matrizes  $(\cdot)$  em formato bloco-diagonal.

Em Geromel *et alii* [9], prova-se que, se esses problemas apresentarem soluções, o ganho  $K^* = YP^{-1} = SW^{-1}$  estabiliza simultaneamente os  $N$  sistemas em (5.37). Não é difícil visualizar (principalmente no primeiro problema) que, para todo ganho  $K$  associado a pares  $(P, Y)$  ou  $(W, S)$  factíveis, e em particular o ganho ótimo, as restrições dos problemas acima são tais que, definindo-se  $(A_{mf})_i = A_i + (B_2)_iK$ ,

$$(A_{mf})_iP + P(A_{mf})'_i < 0 \quad (5.42)$$

pois  $B_1B'_1 > 0$ , e

$$(A_{mf})_iW + W(A_{mf})'_i < 0 \quad (5.43)$$

pois  $D'D > 0$ . De fato, devido à “linearidade” em relação às matrizes  $A_i$  e  $(B_2)_i$ , esse resultado chega a garantir que  $K^*$  estabiliza todo sistema linear invariante no tempo cujo modelo possa ser obtido como combinação convexa das  $N$  matrizes  $A_i$  e  $(B_2)_i$ , o que permite, portanto, delimitar uma região politópica de estabilidade no espaço dos parâmetros  $A$  e  $B_2$ . Mais ainda, considere-se o sistema linear, contínuo e variante no tempo descrito pela equação de estados

$$\dot{x} = A_{mf}(t)x \quad (5.44)$$

em que

$$A_{mf}(t) = \sum_{i=1}^N \alpha_i(t)(A_{mf})_i, \quad \sum_{i=1}^N \alpha_i(t) = 1 \quad \text{e} \quad \alpha_i(t) \geq 0 \quad (5.45)$$

Construindo-se para esse sistema a função de Lyapunov  $V(t) = x(t)'P^{-1}x(t)$ ,  $P > 0$ , suas derivadas são dadas por

$$\dot{V}(t) = x(t)' \left( A_{mf}(t)'P^{-1} + P^{-1}A_{mf}(t) \right) x(t) \quad (5.46)$$

e, portanto, o sistema (5.44) será estável se

$$PA_{mf}(t)' + A_{mf}(t)P = P \sum_{i=1}^N \alpha_i(t)(A_{mf})'_i + \sum_{i=1}^N \alpha_i(t)(A_{mf})_i P \quad (5.47)$$

$$= \sum_{i=1}^N \alpha_i(t) (P(A_{mf})'_i + (A_{mf})_i P) < 0 \quad (5.48)$$

Surpreendentemente, percebe-se que a propriedade expressa em (5.42) assegura que a solução do problema (5.39) atende a uma condição suficiente de estabilidade do sistema linear variante no tempo (5.44). Esse é um resultado bastante expressivo e interessante, pois assegura que a lei de controle linear obtida através da minimização do custo  $\mathcal{H}_2$  para  $N$  sistemas lineares invariantes no tempo garante robustez (no sentido da estabilidade) e otimalidade (no sentido  $\mathcal{H}_2$ ) mesmo na presença de incertezas paramétricas variantes no tempo restritas a uma região poliedral do espaço de parâmetros. Isso consiste, sem dúvida, numa excelente maneira de se obter o controle de sistemas que não são bem descritos por modelos lineares devido a mudanças das condições de operação ocasionada por variações das entradas ou mesmo variações intrínsecas dos parâmetros ao longo do tempo.

Uma outra vantagem desta formulação consiste na possibilidade de imposição de restrições estruturais à matriz de realimentação  $K$ . Isso permite obter soluções, por exemplo, que implementem o chamado controle descentralizado<sup>9</sup>. Faz-se isso impondo-se estruturas adequadas às variáveis  $(P, Y)$  ou  $(W, S)$ , de tal forma que, ao se efetuar o produto  $K = YP^{-1} = SW^{-1}$ , obtenha-se a desejada estrutura em  $K$ . Note-se que essas restrições são sempre lineares nas variáveis do problema e implicam redução do número de variáveis independentes e que esse tipo de abordagem é somente suficiente, uma vez que a inexistência de produtos  $YP^{-1}$  ou  $SP^{-1}$  estruturados e factíveis não implica inexistência de  $K$  estruturado e factível.

Finalmente, a formulação LMI permite obter eficientemente a solução do chamado problema  $\mathcal{H}_2$  generalizado, introduzido em Rotea [22]. Este problema consiste na minimização de funções obtidas a partir de

$$f_p = \sup_{\int_0^\infty w(t)'w(t)dt \leq 1} \|z(t)\|_{\infty, p} \quad (5.49)$$

em que a norma  $(\infty, p)$  do sinal  $z(t)$  é dada por

$$\|z(t)\|_{\infty, p} = \sup_{t \geq 0} \|z(t)\|_p \quad (5.50)$$

Essas normas correspondem ao máximo valor de pico da saída para entradas de energia limitada, ou seja, o *overshoot*, que freqüentemente consiste em uma restrição de projeto.

<sup>9</sup>Veja Geromel *et alii* [7].

Em Wilson [26] e Geromel & Oliveira [8] demonstra-se que, para alguns valores de  $p$ , essas normas podem ser calculadas em função do gramiano de controlabilidade como

$$f_1 = \max_i d_i' C L_o C' d_i \quad (5.51)$$

$$f_2 = \max_i \lambda_i(C L_o C') \quad (5.52)$$

$$f_\infty = \max \text{diag}(C L_o C') \quad (5.53)$$

em que  $d_i$  são todos os vetores pertencentes a  $\mathbb{R}^q$  cujos componentes valem ou um, ou menos um. Verifica-se que, em todos esses casos, essas funções são monotônicas, convexas, e podem ser colocadas no formato LMI mediante a introdução de uma variável escalar adicional. Pode-se minimizar a norma  $f_1$ , por exemplo, adicionando-se ao mesmo<sup>10</sup> conjunto de restrições do caso  $\mathcal{H}_2$  tradicional restrições do tipo

$$\tau \geq d_i' C P C' d_i \quad (5.54)$$

e minimizando-se  $\tau$ . Pode-se fazer o mesmo em relação ao objetivo  $f_\infty$  definindo-se apropriadamente uma base de vetores que extraia os valores da diagonal de  $(C P C')$ . Quanto ao objetivo  $f_2$ , pode-se tratá-lo como exemplificado no problema (5.16).

Encerrando-se as discussões, a formulação dual do problema de controle por realimentação de estados  $\mathcal{H}_2$  — o problema de estimação de estados — pode ser também abordada no contexto LMI de modo que se possa tirar proveito das considerações desta seção. Observe-se, ainda, que a conhecida propriedade de separação dos problemas de controle e estimação permanece válida mesmo para o problema  $\mathcal{H}_2$  generalizado<sup>11</sup>, e isso permite que os resultados discutidos nesta seção sejam implementados por meio de controladores dinâmicos com a mesma ordem do sistema original, sempre que houver impossibilidade de se obter todo o vetor de estados.

### 5.3 Notas finais e perspectivas

Encerrando este trabalho, a fim de situar o leitor desta dissertação em relação às referências bibliográficas, apresentam-se nas próximas linhas breves comentários bibliográficos, organizados por capítulos.

- O capítulo 2 apresenta os fundamentos teóricos necessários ao desenvolvimento dos algoritmos de pontos interiores para desigualdades matriciais lineares dando um tratamento semelhante ao utilizado no contexto de programação linear, veja, por exemplo, [10] e as diversas referências nele citadas. Todo o desenvolvimento foi feito pelo autor. Resultados relativos ao conceito de “autoconcordância” podem ser encontrados em [14]. O nome *transformação simétrica* foi cunhado pelo próprio autor por analogia funcional ao *scaling*, em programação linear.
- Inicia-se o capítulo 3 discutindo-se o problema da medida de centralização, inspirado no tratamento dado em [10] para programação linear. Os resultados referentes ao

<sup>10</sup>A monotonicidade em  $P$  garante que no ponto ótimo obter-se-á, assim como no caso  $\mathcal{H}_2$  tradicional,  $P = L_c$ .

<sup>11</sup>Veja Rotea [22].

passo ótimo são originais. Novamente baseado em [10], descrevem-se os algoritmos seguidores de trajetória. A descrição dos algoritmos primais-duais mescla de maneira original informações obtidas principalmente em [10, 24, 23, 14]. Quanto ao algoritmo projetivo, a apresentação é inteiramente original, revelando aspectos não explorados em [15, 16, 14], como a determinação do limitante  $f_k^*$  e fornecendo uma interpretação geométrica mais sólida e correta. Finalmente, o breve tratamento do problema de factibilidade é baseado em [24, 2].

- No capítulo 4, as discussões acerca da avaliação do operador projeção são inspiradas em [16]. O tratamento da determinação do passo ótimo unidimensional é original. A busca plana, no caso dos algoritmos primais-duais, reflete a abordagem encontrada em [24].
- A primeira parte do capítulo 5 é baseada em [24, 5] enquanto a segunda parte foi inteiramente idealizada e escrita pelo próprio autor.

Finalmente, enumeram-se alguns pontos que poderiam ser ainda explorados, dando prosseguimento a este trabalho.

- Extensão do algoritmo projetivo para que se possa tratar o problema quase-convexo (3.107).
- Implementação computacional eficiente dos algoritmos (em andamento).
- Comparação da eficiência numérica dos vários procedimentos propostos e das diversas famílias de algoritmos (em andamento).



# Referências bibliográficas

- [1] B.D.O. Anderson, J. B. Moore, *Optimal Control : linear quadratic methods*, Prentice Hall, 1989.
- [2] M.S. Bazaraa, H.D. Sherali, C.M. Shetty, *Nonlinear Programming : theory and algorithms*, J. Wiley, 1993.
- [3] S.P. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, PA, 1994.
- [4] S.P. Boyd, C.H. Barrat, *Linear Controller Design : limits of performance*, Prentice-Hall, NJ, 1991.
- [5] S.P. Boyd, L. Vandenberghe, *Notas de aula*, Stanford University, 1995.
- [6] A.V. Fiacco, G.P. McCormick, *Nonlinear Programming : sequential unconstrained minimization techniques*, J. Wiley, 1968.
- [7] J.C. Geromel, J. Bernussou, P. L. D. Peres, “Decentralized Control Through Parameter Space Optimization”, *Automatica*, v. 30, n. 10, out. 1994.
- [8] J.C. Geromel, M.C. Oliveira, “Output Peak Value Determination for Linear Time-invariant Systems”, [artigo submetido].
- [9] J.C. Geromel, P.L.D. Peres, J. Bernussou, “On a Convex Parameter Space Method for Linear Control Design of Uncertain Systems”, *SIAM Journal on Control and Optimization*, v. 29, n. 2, p. 381-402, mar. 1991.
- [10] C.C. Gonzaga, “Path-following Methods for Linear Programming”, *SIAM Review*, v. 34, n. 2, p. 167-224, jun. 1992.
- [11] P. Huard, “Resolution of Mathematical Programming with non-linear Constraints by the Method of Centers”, *Nonlinear Programming*, J. Abadie, ed., Amsterdan, 1967.
- [12] N. Karmakar, “A New Polinomial Time Algorithm for Linear Programming”, *Combinatorica*, v. 4, p. 373-395, 1984.
- [13] D.G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA, 1984.
- [14] Y.E. Nesterov, A. Nemirovskii, *Interior-point Polynomial Methods in Convex Programming*, SIAM, Philadelphia, PA, 1994.

- 
- [15] A. Nemirovskii, P. Gahinet, “The Projective Method for Solving Linear Matrix Inequalities”, *Proceedings of the American Control Conference*, Maryland, jun. 1994.
- [16] A. Nemirovskii, P. Gahinet, “The Projective Method for Solving Linear Matrix Inequalities”, a ser publicado in: *Mathematical Programming Series B*.
- [17] M.C. Oliveira, J.C. Geromel, *Convex Toolbox : user’s guide*, versão 2.0 disponível em <ftp://ftp.dt.fee.unicamp.br>, jul. 1994.
- [18] M.C. Oliveira, J.C. Geromel, “Interior Point Methods and LMI : improvements and benchmark”, [artigo submetido].
- [19] P.L.D. Peres, J.C. Geromel, “An Alternate Numerical Solution to the Linear Quadratic Problem”, *IEEE - Transactions on Automatic Control*, v. 39, p. 198-203, jan. 1994.
- [20] J. Renegar, “A Polynomial Time Algorithm based on Newton’s Method for Linear Programming”, *Mathematical Programming*, v. 40, p. 59-94, 1988.
- [21] T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.
- [22] M.A. Rotea, “The Generalized  $\mathcal{H}_2$  Problem”, *Automatica*, v. 29, n. 2, p. 373-385, 1993.
- [23] L. Vandenberghe, S. Boyd, “A Primal-dual Potential Reduction Method for Problems involving Matrix Inequalities”, *Mathematical Programming*, v. 69, p. 205-236, 1995.
- [24] L. Vandenberghe, S. Boyd, “Semidefinite Programming”, a ser publicado in: *SIAM Review*.
- [25] M. Vidyasagar, *Nonlinear Systems Analysis*, Second Edition, Prentice-Hall, NJ, 1993.
- [26] D.A. Wilson, “Convolution and Hankel Operator Norms for Linear Systems”, *IEEE Transactions on Automatic Control*, v. 34, n. 1, p. 94-97, jan. 1989.