UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

REFORMULAÇÕES DA TRAJETÓRIA CENTRAL E MÉTODO TIPO SUAVILIZAÇÃO PARA PSD

Por

Marcelo Francisco de Andrade

Orientador: Prof. Dr. Geci José Pereira da Silva

DISSERTAÇÃO DE MESTRADO EM MATEMÁTICA

Goiânia, Goiás
2007

Universidade Federal de Goiás
Instituto de Matemática e Estatística
Coordenação de Pós-Graduação em Matemática

Reformulações da trajetória central e método tipo suavilização para PSD

por

Marcelo Francisco de Andrade

Área de concentração: Matemática Aplicada
Orientador: Prof. Dr. Geci José Pereira da Silva

Dissertação submetida à Banca Examinadora designada pelo Conselho Diretor do Instituto de Matemática e Estatística, como parte dos requisitos necessários à obtenção do grau de **Mestre em Matemática**.

Goiânia, Goiás
2007

Dados Internacionais de Catalogação-na-Publicação (CIP)

(GPT/BC/UFG)

Andrade, Marcelo Francisco de.

A553r Reformulações da trajetória central e método tipo suavilização para PSD / Marcelo Francisco de Andrade. – 2007.

69f.

Orientador: Prof. Dr. Geci José Pereira da Silva.

Dissertação (Mestrado) – Universidade Federal de Goiás. Instituto de Matemática e Estatística, 2007.

Bibliografia: f. 68-69.

1. Matemática computacional – Programação semidefinida (PSD) 2 Programação semidefinida (PSD) – Trajetória central. 3. Matemática aplicada – Método tipo suavilização – Convergência I. Silva, Geci José Pereira da II. Universidade Federal de Goiás. **Instituto de Matemática e Estatística.** III. Título.

CDU: 519.6

Resumo

Neste trabalho, através das funções de Fischer-Burmeister e de mínimo, apresentaremos uma reformulação nas condições da trajetória central e um algoritmo tipo suavilização para o Problema de Programação Semidefinida, PSD, generalizando assim alguns métodos de Programação Linear e Problema de Complementariedade Linear. Mostraremos também que o método proposto apresenta convergência global e superlinear local.

Abstract

In this work, by Fischer-Burmeister and minimum functions, we introduce a reformulation of the central path conditions and a smoothing-type algorithm to the Problem of Semidefinite Programming, PSD, generalizing thus some methods of Linear Programming and Linear Complementarity Problem. We also will show that proposed method holds global and local superlinear convergence.

Sumário

Introdução

Um Problema de Programação Semidefinida, PSD, é um problema de otimização convexa no espaço das matrizes simétricas de ordem $n \times n$, $S^{n\times n}$, dado na forma primal por

$$\begin{array}{rl} \min & C \bullet X \\ \text{sujeito a} & A_i \bullet X = b_i, \quad i = 1, \ldots, m, \\ & X \succeq 0 \end{array} \tag{1}$$

onde X, A_i e $C \in S^{n\times n}$ com $i = 1, \ldots, n$, $b \in \mathbb{R}^m$, e $C \bullet X$ =traço(CX). Na forma dual o PSD é dado por

$$\begin{array}{rl} \max & b^T\lambda \\ \text{sujeito a} & \sum_{i=1}^{m} \lambda_i A_i + S = C, \\ & S \succeq 0 \end{array} \tag{2}$$

onde $S, A_i, C \in S^{n\times n}$ com $i = 1, \ldots, n$ e $\lambda, b \in \mathbb{R}^m$.

Utilizando os resultados de otimalidade para otimização convexa temos que as condições de otimalidade para o PSD acima são dadas por

$$\begin{array}{rcl} \sum_{i=1}^{m} \lambda_i A_i + S & = & C, \\ A_i \bullet X & = & b_i, \quad \forall\ i = 1, \ldots m \\ X \succeq 0, \quad S \succeq 0, & & XS = 0. \end{array} \tag{3}$$

Assim, as idéias de trajetória central de Programação Linear podem ser generalizadas para PSD através da perturbação das condições de otimalidade (**??**) e deste modo podemos obter que as condições de trajetória central para o PSD são dadas por:

$$\begin{aligned} \sum_{i=1}^{m} \lambda_i A_i + S &= C, \\ A_i \bullet X &= b_i, \quad \forall\ i = 1, \dots m \\ X \succ 0,\ S \succ 0,\ XS &= \tau^2 I, \quad \tau > 0. \end{aligned} \tag{4}$$

Deste modo, nesta dissertação, um dos nossos objetivos é através da Função de Fischer-Burmeister suavilizada, que é dada por

$$\begin{aligned} \phi_\tau : S^{n\times n} \times S^{n\times n} &\rightarrow S^{n\times n} \\ (X, S) &\mapsto \phi_\tau(X, S) = X + S - (X^2 + S^2 + 2\tau^2 I)^{1/2}, \end{aligned} \tag{5}$$

e da função-mínimo suavilizada, dada por

$$\begin{aligned} \phi_\tau : S^{n\times n} \times S^{n\times n} &\rightarrow S^{n\times n} \\ (X, S) &\mapsto \phi_\tau(X, S) = X + S - ((X - S)^2 + 4\tau^2 I)^{1/2}. \end{aligned} \tag{6}$$

dar uma nova caracterização para as condições da trajetória central (**??**) através do sistema de equações não-lineares dado a seguir

$$\Phi_\tau(X, \lambda, S) \ := \left(\begin{array}{c} \displaystyle\sum_{i=1}^{m} \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi_\tau(X, S) \end{array} \right) = 0.$$

O outro objetivo é utilizar o sistema acima para apresentar um método tipo suavilização, que pode ser visto como uma generalização de alguns métodos para Programação Linear e Problemas de Complementariedade Linear, para a estrutura de

Programação Semidefinida, para mais detalhes, veja [?]. E, além disso, provaremos para tal método, a convergência global e convergência superlinear local.

Nesta dissertação, no Capítulo 1 são apresentadas algumas noções preliminares quanto à notação e resultados básicos utilizados na dissertação.

No Capítulo 2, abordaremos inicialmente sobre as reformulações na trajetória central sugeridas pelas condições de otimalidade utilizando as funções de Fischer-Burmeister e mínimo suavilizadas.

No Capítulo 3, mostraremos que as funções de Fischer-Burmeister e mínimo suavilizadas são diferenciáveis.

No Capítulo 4, descreveremos o Algoritmo e mostraremos que está bem definido.

E, finalmente no Capítulo 5 mostraremos convergência global e superlinear local do Algoritmo.

Capítulo 1

Preliminares

Neste capítulo apresentamos definições e resultados importantes que serão necessários no desenvolvimento deste trabalho.

1.1 Resultados básicos

A seguir algumas definições e observações sobre matrizes.

Definição 1.1.1. *Uma matriz A de ordem $n \times n$ é dita simétrica se $A = A^T$ (A^T é a matriz transposta de A).*

Indicaremos por $S^{n\times n}$ o conjunto das matrizes simétricas de ordem $n \times n$.

Definição 1.1.2. *Uma matriz simétrica A é dita semidefinida positiva (definida positiva) se $z^TAz \geq 0$ $(z^TAz > 0)$, $\forall\ z \in R^n$, $z \neq 0$, ou equivalentemente, se todos os autovalores de A são não-negativos (positivos).*

Usaremos a notação $A \succeq 0$ para dizer que a matriz A é simétrica e semidefinida positiva. Analogamente, $A \succ 0$ significa que a matriz A é simétrica e definida positiva. O conjunto das matrizes simétricas semidefinidas positivas será indicado por $S_+^{n\times n}$ e o das definidas positivas por $S_{++}^{n\times n}$. Dessa forma, quando afirmarmos, por exemplo, que $A \succeq 0$ (ou $A \succ 0$,) isso significa que $A \in S_{+}^{n\times n}$ (e respectivamente $A \in S_{++}^{n\times n}$).

As desigualdades $A \succeq B$ ou $A \succ B$ querem dizer que $A - B \succeq 0$ ou que $A - B \succ 0$.

Observação 1.1.1. *O símbolo padrão de Kronecker será denotado por δ_{ij}, para todo $i, j = 1, \ldots, n$ e representado por*

$$\delta_{ij} := \begin{cases} 1, & se\ i = j \\ 0, & se\ i \neq j. \end{cases}$$

Definição 1.1.3. *Uma matriz D é dita diagonal, se $d_{ij} = 0, \quad \forall i \neq j$.*

Quando indicarmos $D = diag(d_{11}, d_{22}, \ldots, d_{nn})$, estaremos nos referindo ao fato de que os elementos $d_{11}, d_{22}, \ldots, d_{nn}$ compõem a diagonal principal da matriz D, e os demais elementos de D são nulos.

Definição 1.1.4. *Uma matriz A não-singular é dita ortogonal, se $A^T = A^{-1}$.*

Observamos que se $A \in S^{n\times n}$, então A é diagonalizável, ou seja, existe uma matriz ortogonal P tal que $A = PDP^T$, onde D é uma matriz diagonal. Neste caso, dizemos que a matriz D diagonaliza a matriz A.

A Observação a seguir trata da decomposição espectral (de autovalores) de uma dada matriz.

Observação 1.1.2. *Considere $A \in S^{n\times n}$. Assim, a matriz A pode ser decomposta como*

$$A = PDP^T = \sum_{i=1}^{n} \lambda_i(A) p_i p_i^T$$

onde D é uma matriz diagonal com os autovalores $\lambda_i(A)$ $(i = 1, \ldots, n)$ de A na diagonal, e P é uma matriz ortogonal com um conjunto correspondente de autovetores *ortonormais* $p_1, \ldots, p_n$ *de* A *como colunas. Observe também que toda matriz* $X \succ 0$ *é não-singular, ou seja, admite inversa.*

Supondo que $A \succeq 0$, *teremos então que* $\lambda_i(A) \geq 0$ $(i = 1, \ldots, n)$ *e dessa forma podemos definir a única raiz quadrada semidefinida positiva de* A :

$$A^{1/2} = \sum_{i=1}^{n} \sqrt{\lambda_i(A)} p_i p_i^T.$$

Observe ainda que $A^{1/2}A^{1/2} = A$ *e* $A^{1/2}$ *é a única matriz com tal propriedade.*

Definição 1.1.5. *O traço de uma matriz* A *de ordem* $n \times n$, *denotado por* $tr(A)$, *é definido como a soma dos elementos da diagonal principal de* A, *isto é,*

$$tr(A) = \sum_{i=1}^{n} a_{ii}.$$

No espaço vetorial $S^{n\times n} \times S^{n\times n}$, podemos definir o seguinte produto escalar

$$A \bullet B = \langle A, B\rangle = tr(AB^T), \; A, B \in S^{n\times n}. \tag{1.1}$$

É bom ressaltar que o símbolo $\bullet$ representa a composição de duas aplicações e não o produto usual entre duas matrizes.

Teorema 1.1.1. *Seja* $A \succ 0$. *Para cada* $H \in \mathbb{R}^{n\times n}$ *dada, há uma única solução* G *para*

$$GA + AG = H$$

Demonstração. Veja [?], Teorema 2.2.3, p. 98. □

Dada uma matriz $A \in \mathbb{R}^{n\times n}$, denotaremos por $vec(A)$ o vetor formado pelas colunas da matriz A, ou seja

$$vec\ (A) = [a_{11}, a_{21}, \ldots, a_{n1}, a_{12}, a_{22}, \ldots, a_{nn}]^T$$

Adiante, a definição de operador auto-adjunto.

Definição 1.1.6. *Um operador linear $T : E \to E$, num espaço vetorial munido de produto interno, chama-se auto-adjunto quando*

$$\langle Tu, v\rangle = \langle u, Tv\rangle,$$

para quaisquer $u, v \in T$.

Agora, vamos relacionar fortemente convexa com fortemente monótona. Observe:

Definição 1.1.7. *Seja $f : \mathbb{R}^n \to \mathbb{R}$*

Dizemos que f é fortemente convexa (com módulo $\gamma > 0$), quando para quaisquer $x, y \in D \subset \mathbb{R}^n$ e $\alpha \in [0, 1]$, tem-se

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) - \gamma\alpha(1-\alpha)\|x - y\|^2.$$

Acrescentamos a isso o fato de que toda função fortemente convexa possui gradiente fortemente monótono, uma vez que

$$\langle \nabla f(x) - \nabla f(y), x - y\rangle \geq \gamma\|x - y\|^2,$$

onde γ é uma constante.

Dada $F(x) : \mathbb{R}^n \to \mathbb{R}^n$, o Método de Newton é usado para resolver um sistema de equações não-lineares:

$$\text{achar } x^* \in \mathbb{R}^n \text{ tal que } F(x^*) = 0.$$

O Método gera uma seqüência iterativa x^k onde dada a iterada atual x^k, a iterada seguinte é obtida como a raiz de uma aproximação afim de F numa vizinhança de x^k. Mais precisamente, o Método de Newton é o seguinte: dada a iterada x^k, resolvendo o seguinte sistema

$$F'(x^k)s^k = -F(x^k)$$

e obtemos que $x^{k+1} = x^k + s^k$. O Método de Newton apresenta convergência quadrática local.

Observação 1.1.3. *Sendo V um espaço vetorial e*

$$T : V \to V$$

um operador linear, podemos definir a norma espectral $\| \cdot \|_2$ no espaço $\mathbb{R}^{n\times n}$ por

$$\| A \|_2 = \max_{x\in\mathbb{R}^n} \frac{\| Ax \|}{\| x \|}, \quad A \in \mathbb{R}^{n\times n}.$$

O produto escalar apresentado na Definição 1.1 induz a tão chamada norma de Frobenius (ou Euclideana) em $\mathbb{R}^{n\times n}$ e é dada por

$$\| A \|_F = \langle A, A\rangle = tr(AA^T) = \sum_{i,j=1}^{n} a_{ij}^2.$$

Dessa forma, o espaço vetorial $\mathbb{R}^{n\times n} \times \mathbb{R}^m \times \mathbb{R}^{n\times n}$ será denotado com a seguinte norma:

$$||| (X, \lambda, S) ||| = \sqrt{\| X \|_F^2 + \| \lambda \|_2^2 + \| S \|_F^2}.$$

e para o espaço vetorial $\mathbb{R}^{n\times n} \times \mathbb{R}^m \times \mathbb{R}^{n\times n} \times \mathbb{R}$ usaremos a norma

$$||| (X, \lambda, S, \tau) ||| = \sqrt{\| X \|_F^2 + \| \lambda \|_2^2 + \| S \|_F^2 + \tau^2}$$

É bom lembrar que as duas normas citadas anteriormente correspondem à norma Euclideana padrão se todas as entradas nos espaços vetoriais forem vistas como vetores.

Dessa forma, após essas considerações iniciais, podemos enfim expor o problema central deste trabalho.

Capítulo 2

Reformulações na trajetória central

Neste Capítulo, daremos duas novas reformulações das condições da trajetória central para PSD's. Nesse estudo, nos ocorrerá o seguinte sistema de equações não-lineares:

$$\begin{cases} \displaystyle\sum_{i=1}^{m} \lambda_i A_i + S - C \\ A_i \bullet X - b_i, \;\; (i = 1, \cdots, m) \\ \phi_\tau(X, S) \end{cases}$$

Teremos que caracterizar essas funções ϕ de maneira que resolver as condições da trajetória central (4) seja equivalente a resolver o sistema descrito anteriormente.

Para então iniciar nossa abordagem, definimos a aplicação

$$\begin{aligned} \varphi : \mathbb{R} \times \mathbb{R} \;&\rightarrow\; \mathbb{R} \\ (a, b) \;&\mapsto\; \varphi(a, b) = a + b - \sqrt{a^2 + b^2} \end{aligned} \tag{2.1}$$

Tal aplicação é conhecida por *função de Fischer-Burmeister*. É bem definida e possui a seguinte propriedade:

Propriedade 2.0.1. $\varphi(a, b) = 0$ *se, e somente se,* $ab = 0, \; a \geq 0, \; b \geq 0.$

Demonstração. Vemos facilmente que se $\varphi(a,b) = 0$, temos que

$$a + b = \sqrt{a^2 + b^2}.$$

Elevando ambos os lados ao quadrado, obtemos que

$$a^2 + 2ab + b^2 = a^2 + b^2,$$

e, daí segue que $ab = 0$. Com isso, temos que $a = 0$ ou $b = 0$. Agora, se $a = 0$, segue que

$$b = \sqrt{b^2}$$

e assim $b \geq 0$. Analogamente, se $b = 0$, segue que $a \geq 0$. o que prova a primeira implicação. Já para a segunda, temos que

$$\begin{aligned}
\varphi(a,b) &= a + b - \sqrt{a^2 + b^2} \\
&= a + b - \sqrt{a^2 + 2ab + b^2}, \text{ pois } ab = 0 \\
&= a + b - \sqrt{(a+b)^2}, \text{ e como } a, b \geq 0 \\
&= a + b - (a + b) = 0.
\end{aligned}$$

$\square$

Agora, abrangendo o universo da Programação Semidefinida, definamos a aplicação

$$\begin{aligned}
\phi : S^{n\times n} \times S^{n\times n} &\rightarrow S^{n\times n} \\
(X,S) &\mapsto \phi(X,S) = X + S - (X^2 + S^2)^{1/2}
\end{aligned} \tag{2.2}$$

A aplicação acima é uma extensão da função de Fischer-Burmeister (2.1), para $S^{n\times n}$, a diferença é que, na primeira, os argumentos da função são números reais e nesta são matrizes semidefinidas.

A função de Fischer-Burmeister dada em (2.2) também tem a seguinte propriedade:

Propriedade 2.0.2. $\phi(X,S) = 0$ *se, e somente se,*$X \succeq 0,\ S \succeq 0,\ XS = 0.$

Adiante provaremos tal propriedade utilizando-se de um resultado encontrado em [?]. Mas, antes disso, enunciaremos um lema que contribuirá para a demonstração do resultado citado.

Lema 2.0.1. *Seja* $A \succeq 0$ *de ordem* $n \times n$.
Então, $a_{ii} \geq 0,\ \forall i = 1, 2, \ldots, n.$
Além disso, se $a_{tt} = 0$, *então todos os elementos da t-ésima linha e t-ésima coluna de* A *são iguais a zero.*

Lema 2.0.2. *Sejam* $X, S \in S_+^{n\times n}$. *Então,* $XS = 0$ *se, e somente se,* $X \bullet S = 0$.

Demonstração. Como $XS = 0$, então temos que

$$0 = tr(XS) = tr(XS^T) = X \bullet S.$$

Reciprocamente, agora façamos a decomposição espectral (de autovalores) de S :

$$S = UDU^T$$

sendo U ortogonal, isto é,

$$U^T = U^{-1} \text{ e } D = diag(\lambda_i), \lambda_i \geq 0,\ i = 1, \ldots, n.$$

Considere

$$F = U^T XU;\ F \succeq 0;\ f_{ii} \geq 0.$$

Vamos mostrar que $FD = 0$.

Como $X \bullet S = 0$, então $0 = tr(XS) = tr(UFU^TUDU^T)$. Claramente, vemos que

$$tr(UFU^TUDU^T) = tr(UFDU^T) = tr(FD) = 0,$$

e chegamos então em $F \bullet D = 0$.

Podemos ainda mencionar que se $tr(FD) = 0$, então

$$\sum_{i=1}^{n} f_{ii}\lambda_i = 0,$$

o que implica em $f_{ii}\lambda_i = 0, \ \forall i = 1, \ldots, n$.

Para a igualdade anterior, estabelecemos duas possibilidades:

i) Se $f_{ii} > 0$, então $\lambda_i = 0$. Daí temos então que $D = 0$ e consequentemente $FD = 0$.

ii) Se $\lambda_i > 0$, então $f_{ii} = 0$. Dessa forma, $F = 0$ e pelo lema anterior, $FD = 0$.(*)

E daí, segue abaixo que

$$FD = U^T XUU^T SU = 0$$

Com isso, temos que $UU^T XSUU^T = 0$ e concluimos que $XS = 0$.

Agora, suponha por absurdo, que ocorra $(FD)_{ij} \neq 0$ para algum i, j.

Então

$$f_{ij}\lambda_j \neq 0,$$

o que, por $(*)$, nos leva a concluir que todo elemento da j-ésima coluna é igual a zero, e então, $f_{ij} = 0$, uma contradição. $\square$

O Lema demonstrado acima nos possibilita enunciar a seguinte proposição:

Proposição 2.0.1. *Seja ϕ a função de Fischer-Burmeister definida em* (2.2)*. Então*

$$\phi(X, S) = 0 \textit{ se, e somente se, } X \succeq 0, \ S \succeq 0, \ XS = 0$$

Demonstração. Por hipótese, temos que

$$X \succeq 0, \ S \succeq 0, \ XS = 0.$$

Note que, pelo Lema (2.0.2), temos que

$$XS = 0 = tr(XS) = tr(SX) = S \bullet X = SX$$

Isso implica que:

$$(X + S)^2 = X^2 + XS + SX + S^2 = X^2 + S^2$$

Extraindo a raiz quadrada de ambos os lados, vem

$$X + S = (X^2 + S^2)^{1/2},$$

que implica em $\phi(X, S) = 0$.

Reciprocamente, temos, por hipótese, que $\phi(X, S) = 0$, ou seja

$$X + S = (X^2 + S^2)^{1/2}, \quad X, S \in S^{n \times n}$$

Elevando ambos os lados ao quadrado, temos que

$$(X + S)^2 = X^2 + S^2, \quad X + S \in S^{n \times n}_{+},$$

e equivalentemente $XS + SX = 0, \quad X + S \in S^{n \times n}_{+}$.

Seja $X = Q^T DQ,\ Q \in \mathbb{R}^{n \times n}$ matriz ortogonal (isto é, $Q^{-1} = Q^T$ e $D = diag(\lambda_1, \ldots, \lambda_n)$), a decomposição espectral (de autovalores) da matriz simétrica X. Reescrevendo:

$$\begin{aligned} XS + SX &= 0, \quad X + S \in S^{n \times n}_{+} \\ Q^T DQS + SQ^T DQ &= 0, \quad Q^T DQ + S \in S^{n \times n}_{+} \\ Q(Q^T DQS + SQ^T DQ)Q^T &= 0, \quad Q(Q^T DQ + S)Q^T \in S^{n \times n}_{+} \\ DQSQ^T + QSQ^T D &= 0, \quad D + QSQ^T \in S^{n \times n}_{+} \end{aligned}$$

Façamos agora a substituição $A = QSQ^T$ e vem

$$DA + AD = 0, \quad D + A \in S^{n \times n}_{+}$$

Componente a componente, reescrevemos

$$(\lambda_i + \lambda_j)a_{ij} = 0, \quad D + A \in S_+^{n\times n}, \quad \forall i, j = 1, \ldots, n.$$

Em particular, obtemos para $i = j$

$$2\lambda_i a_{ii} = 0 \text{ e } \lambda_i + a_{ii} \geq 0, \ \forall i = 1, \ldots, n.$$

Isso implica que $\lambda_i \geq 0, \quad \forall i = 1, \ldots, n$ e consequentemente, $X \succeq 0$. Usando um argumento análogo (baseado agora na decomposição espectral de S), vemos que S também é semidefinida positiva.

Agora, para verificar que $XS = 0$, observemos que

$$\begin{aligned} XS + SX &= 0 \\ X \bullet S &= tr(XS) = \frac{1}{2}[tr(XS + SX)] = 0 \end{aligned}$$

E, pelo lema (2.2), temos que $X \bullet S = 0$ se, e somente se, $XS = 0$. Isso demonstra a propriedade. □

A partir de agora, com o objetivo, de caracterizar as condições da trajetória central, vamos fazer uma modificação na definição da função ϕ. Considere $\tau \geq 0$ um parâmetro. Defina a aplicação

$$\begin{aligned} \varphi_\tau : \mathbb{R} \times \mathbb{R} &\rightarrow \mathbb{R} \\ (a, b) &\mapsto \varphi_\tau(a, b) = a + b - \sqrt{a^2 + b^2 + 2\tau^2} \end{aligned}$$

A aplicação acima é conhecida como função de Fischer-Burmeister suavizada e é continuamente diferenciável para todo $\tau > o$ e φ_τ coincide com φ quando $\tau = 0$. Vale para esta função φ_τ uma propriedade análoga à que foi dada a função φ anteriormente em (2.0.1). Observe a seguir:

Propriedade 2.0.3. *$\varphi_\tau(a,b)=0$ se, e somente se, $a \geq 0,\ b \geq 0,\ ab=\tau^2$.*

Demonstração. Para a primeira implicação, temos que se $\varphi_\tau(a,b)=0$, então

$$(a+b)=\sqrt{a^2+b^2+\tau^2}$$

Elevando ambos os termos ao quadrado, vem

$$(a+b)^2=(\sqrt{a^2+b^2+\tau^2})^2$$

E então,

$$a^2+2ab+b^2=a^2+b^2+2\tau^2$$

E consequentemente, $ab=\tau^2$.

Suponha, por absurdo, que $a<0$ ou $b<0$. Dessa forma, $\tau^2=ab<0$, que é uma contradição, visto que τ é um parâmetro positivo. Suponha também, por absurdo, que $a<0$ e $b<0$. Por hipótese, temos que

$$\sqrt{a^2+b^2+\tau^2}=a+b$$

e concluimos então que

$$\sqrt{a^2+b^2+\tau^2}<0.$$

Um absurdo. Logo, $a,b \geq 0$.

Para a segunda implicação, temos, por hipótese, que $ab=\tau^2$ e implica que

$$\varphi_\tau(a,b)=a+b-\sqrt{a^2+b^2+2\tau^2}=a+b-\sqrt{a^2+b^2+2ab}=a+b-\sqrt{(a+b)^2}.$$

E, como por hipótese, $a,b \geq 0$, temos que

$$\varphi_\tau(a,b)=a+b-(a+b)=0,$$

o que completa a demonstração. $\square$

Generalizaremos agora tal aplicação para o universo das matrizes simétricas.

Defina a seguinte aplicação, também conhecida como função de Fischer-Burmeister suavizada:

$$\begin{aligned} \phi_\tau : S^{n\times n} \times S^{n\times n} &\rightarrow S^{n\times n} \\ (X,S) &\mapsto \phi_\tau(X,S) = X + S - (X^2 + S^2 + 2\tau^2 I)^{1/2} \end{aligned} \tag{2.3}$$

Proposição 2.0.2. *Seja* $\tau > 0$ *e* ϕ *definida como em* (2.3)*. Então,*

$$\phi_\tau(X,S) = 0 \textit{ se, e somente se, } X \succ 0,\ S \succ 0,\ XS = \tau^2 I.$$

Demonstração. Por hipótese, temos que

$$X \succ 0,\ S \succ 0,\ XS = \tau^2 I.$$

Temos que se $X \succ 0$, então $\exists\ X^{-1}$. Tomando então $XS = \tau^2 I$ e multiplicando por X^{-1} do lado esquerdo da igualdade e por X o lado direito, vem

$$X^{-1}(XS)X = X^{-1}(\tau^2 I)X$$

E, então,

$$SX = X^{-1}\tau^2 X = \tau^2 X^{-1} X = \tau^2 I \tag{2.4}$$

Dessa forma, temos que

$$XS + SX = 2\tau^2 I$$

E daí

$$(X+S)^2 = X^2 + XS + SX + S^2 = X^2 + S^2 = X^2 + S^2 + 2\tau^2 I$$

Extraindo a raiz quadrada de ambos os lados, temos que

$$X + S = (X^2 + S^2 + 2\tau^2 I)^{1/2}$$

e então, $\phi_\tau(X,S) = 0$.

Temos, por hipótese, que $\phi_\tau(X,S) = 0$, isto é,

$$X + S = (X^2 + S^2 + 2\tau^2 I)^{1/2}$$

Elevando ambos os membros ao quadrado, vem

$$\begin{aligned} (X+S)^2 &= X^2 + S^2 + 2\tau^2 I, \quad X + S \in S_{++}^{n\times n} \\ \text{e equivalentemente} \quad XS + SX &= 2\tau^2 I, \quad X + S \in S_{++}^{n\times n} \end{aligned} \tag{2.5}$$

Seja $X = Q^T DQ$, $Q \in \mathbb{R}^{n\times n}$ matriz ortogonal (isto é, $Q^{-1} = Q^T$ e $D = diag(\lambda_1, \ldots, \lambda_n)$), a decomposição espectral da matriz simétrica X. Daí, fazendo a substituição solicitada, (2.5) fica assim reescrita:

$$\begin{aligned} XS + SX &= 2\tau^2 I, \quad X + S \in S_{++}^{n\times n} \\ Q^T DQS + SQ^T DQ &= 2\tau^2 I, \quad Q^T DQ + S \in S_{++}^{n\times n} \\ Q(Q^T DQS + SQ^T DQ)Q^T &= Q \cdot 2\tau^2 I \cdot Q^T, \quad Q(Q^T DQ + S)Q^T \in S_{++}^{n\times n} \\ DQSQ^T + QSQ^T D &= 2\tau^2 I, \quad D + QSQ^T \in S_{++}^{n\times n} \end{aligned}$$

Fazendo a substituição $A = QSQ^T$, temos que

$$DA + AD = 2\tau^2 I, \quad D + A \in S_{++}^{n\times n} \tag{2.6}$$

Componente a componente, reescrevemos

$$(\lambda_i + \lambda_j)a_{ij} = 2\tau^2 \delta_{ij}, \quad D + A \in S_{++}^{n\times n}, \quad \forall i,j = 1,\ldots,n. \tag{2.7}$$

A notação δ_{ij} representa os símbolos do Kronecker, já citados neste trabalho.

Em particular, para $i = j$ temos

$$2\lambda_i a_{ii} = 2\tau^2 \ (\tau > 0) \ \text{ e } \ \lambda_i + a_{ii} \geq 0, \quad \forall i = 1,\ldots,n.$$

E, assim, temos então que $\lambda_i > 0, \forall i = 1, \ldots, n$ o que implica em $X \succ 0$.

Usando um argumento análogo (baseado agora na decomposição espectral de S), vemos que S também é definida positiva.

Agora, para finalizar, verifiquemos que $XS = \tau^2 I$.

Para isso, observemos que de (2.7), considerando $i \neq j$, obtemos

$$(\lambda_i + \lambda_j)a_{ij} = 2\tau^2\delta_{ij}$$

Daí, chegamos em $a_{ij} = 0$, desde que $\lambda_i + \lambda_j > 0$. Portanto A é matriz diagonal. Em particular, por essa razão, temos que

$$DA = AD$$

Assim, de (2.6), como $DA = \tau^2 I$, então

$$\begin{aligned} Q^T(DA)Q &= Q^T(\tau^2 I)Q \\ Q^T DQSQ^T Q &= \tau^2 I \\ XS &= \tau^2 I \end{aligned}$$

Isso demonstra a propriedade. $\square$

A partir de agora, vamos introduzir uma nova função cujas propriedades são similares às de Fischer-Burmeister suavizada. Considere então

$$\varphi(a, b) := 2 \min \{a, b\}; \ a, b \in \mathbb{R}$$

Tal função será conhecida por função mínimo.

Observação 2.0.4. *Vale também para esta função a seguinte equivalência:*

$$\varphi(a, b) = 0 \text{ se, e somente se, } a \geq 0, \ b \geq 0, \ ab = 0.$$

Demonstração. Por hipótese, temos que $\varphi(a,b) = 0$. Logo, $2\min\{a,b\} = 0$ e, com isso, $\min\{a,b\} = 0$.

Se $a < b$, temos que $\min\{a,b\} = a = 0$ e portanto $b \geq 0$. Porém, se $b < a$, então $\min\{a,b\} = b = 0$ e teremos $a \geq 0$. Com tudo isso, se $\min\{a,b\} = 0$, então a ou $b = 0$ o que acarreta em $ab = 0$.

Agora, supondo que $ab = 0$, temos que $a = 0$ ou $b = 0$.

Se $a = 0$, teremos $b \geq 0 = \min\{a,b\} = 2\min\{a,b\} = \varphi(a,b)$. Agora, se $b = 0$, então $a \geq 0 = \min\{a,b\} = 2\min\{a,b\} = \varphi(a,b)$, o que completa a demonstração. $\square$

Com o objetivo de estender a definição da função mínimo à classe das matrizes simétricas, faremos uma reformulação.

Observação 2.0.5.

$$\varphi(a,b) = 2\min\ \{a,b\} = a + b - |a-b| = a + b - \sqrt{(a-b)^2}$$

Demonstração. Temos por hipótese que,

$$\varphi(a,b) = 2\min\ \{a,b\}$$

Se $a < b$, então $\varphi(a,b) = 2a = a + b + a - b = a + b - (b-a)$. Contudo, se $b < a$, logo $\varphi(a,b) = 2b = a + b - a + b = a + b - (a-b)$.

Pela própria definição de módulo, vê-se claramente que

$$\varphi(a,b) = a + b - |a-b|$$

Considerando enfim $\sqrt{(a-b)^2}$ percebemos que

$$\sqrt{(a-b)^2}\begin{cases} a-b, \text{ se } a \geq b \\ b-a, \text{ se } a < b \end{cases}$$

Vemos também que a função exposta acima é equivalente à função-módulo, logo

$$\varphi(a,b) = 2\min\{a,b\} = a + b - |a-b| = a + b - \sqrt{(a-b)^2}$$

$\square$

Através da expressão citada anteriormente, definamos a função mínimo agora no universo das matrizes simétricas:

$$\begin{aligned} \phi : S^{n\times n} \times S^{n\times n} &\rightarrow S^{n\times n} \\ (X,S) &\mapsto \phi(X,S) = X + S - ((X-S)^2)^{\frac{1}{2}} \end{aligned} \tag{2.8}$$

Proposição 2.0.3. *Seja ϕ a função mínimo definida anteriormente em (2.8). Então,*

$$\phi(X,S) = 0 \text{ se, e somente se } X \succeq 0,\ S \succeq 0,\ XS = 0.$$

Demonstração. Temos inicialmente por hipótese que

$$X \succeq 0,\ S \succeq,\ XS = 0$$

implica em $0 = tr(XS) = tr(XS^T) = X \bullet S$.

Dessa forma, de $XS = 0$ temos que

$$tr(XS) = 0 = tr(SX) = S \bullet X = SX,$$

pelo Lema (2.0.2).

Isso implica que $XS + SX = 0$. E dessa forma,

$$(X+S)^2 = X^2 + XS + SX + S^2 = X^2 + S^2 = (X-S)^2,$$

e vemos então que $(X+S)^2 = (X-S)^2$ que implica em

$$X + S = ((X-S)^2)^{\frac{1}{2}},$$

que equivale a $\phi(X, S) = 0$.

De maneira recíproca agora, temos por hipótese que $\phi(X, S) = 0$. Com isso,

$$X + S = ((X - S)^2)^{\frac{1}{2}},$$

e elevando ambos os termos ao quadrado, vem

$$(X + S)^2 = (X - S)^2 \text{ e } X + S \in S_+^{n \times n},$$

ou equivalentemente $XS + XS = 0$ e $X + S \in S_+^{n \times n}$. O término da demonstração é inteiramente análogo ao da Proposição 2.0.1. □

Queremos agora modificar a definição da função mínimo a fim de que satisfaça uma caracterizacao das condições da trajetória central dadas em (**??**). Para isso, inicialmente, faremos uma modificação da função mínimo para variáveis escalares como segue:

$$\varphi_\tau(a, b) := a + b - \sqrt{(a - b)^2 + 4\tau^2}, \ \tau \geq 0$$

A função mínimo descrita acima é conhecida por função mínimo suavizada ou ainda por função suave de Chen-Harker-Kanzow-Smale e também possui a seguinte propriedade:

Propriedade 2.0.4.

$$\varphi_\tau(a, b) = 0 \textit{ se, e somente se, } a > 0, \ b > 0, \ ab = \tau^2, \ \tau > 0$$

Demonstração. Por hipótese, temos que $a + b = \sqrt{(a - b)^2 + 4\tau^2}$. Elevando ambos os termos ao quadrado, vem

$$a^2 + 2ab + b^2 = (a - b)^2 + 4\tau^2 = a^2 - 2ab + b^2 + 4\tau^2,$$

que recai em $4ab = 4\tau^2$ e portanto $ab = \tau^2$.

Veja:

$$\text{se } a \text{ ou } b \leq 0, \text{ contradiz } ab = \tau^2$$

$$\text{se } a, b < 0, \text{ contradiz } a + b = \sqrt{(a-b)^2 + 4\tau^2}$$

Logo, $a, b > 0$.

Reciprocamente, considerando $a, b, \tau > 0$ e $ab = \tau^2$, vem

$$\begin{aligned}
\varphi_\tau(a,b) = a + b - \sqrt{(a-b)^2 + 4\tau^2} &= a + b - \sqrt{a^2 - 2ab + b^2 + 4ab} \\
&= a + b - \sqrt{a^2 + 2ab + b^2} \\
&= a + b - \sqrt{(a+b)^2} \\
&= a + b - (a+b) \\
&= 0.
\end{aligned}$$

□

Isso nos motiva à definição da aplicação a seguir para a classe das matrizes simétricas

$$\begin{aligned}
\phi_\tau : S^{n\times n} \times S^{n\times n} &\rightarrow S^{n\times n} \\
(X, S) &\mapsto \phi_\tau(X,S) = X + S - ((X-S)^2 + 4\tau^2 I)^{1/2}
\end{aligned} \tag{2.9}$$

E assim, enunciamos a seguinte proposição.

Proposição 2.0.4. *Seja* $\tau > 0$ *e definida como em* (2.9). *Então*

$$\phi_\tau(X,S) = 0 \textit{ se, e somente se } X \succ 0, \; S \succ 0, \; XS = \tau^2 I$$

Demonstração. Considere

$$X \succ 0, \; S \succ 0, \; XS = \tau^2 I$$

Por 2.4, temos que $XS + SX = 2\tau^2 I$, e assim,

$$\begin{array}{ll} (1) & (X+S)^2 = X^2 + XS + SX + S^2 \\ (2) & (X-S)^2 = X^2 - XS - SX + S^2 \end{array} \tag{2.10}$$

Fazendo (1) $-$ (2) temos que $(X+S)^2 - (X-S)^2 = 2(XS+SX)$, e então

$$(X+S)^2 = (X-S)^2 + 4\tau^2 I$$

Extraindo a raiz quadrada de ambos os lados, vem

$$X + S = ((X-S)^2 + 4\tau^2 I)^{\frac{1}{2}}$$

que implica em $\phi_\tau(X, S) = 0$.

Reciprocamente, se

$$\phi_\tau(X, S) = 0, \text{ para } X, S \in S^{n\times n}$$

temos que

$$X + S = ((X-S)^2 + 4\tau^2 I)^{\frac{1}{2}}$$

e a demonstração segue inteiramente análoga à da Proposição 2.0.2. □

De agora em diante, quando mencionarmos à função ϕ_τ, estaremos nos referindo tanto à função de Fischer-Burmeister suave

$$\phi_\tau(X, S) := X + S - (X^2 + S^2 + 2\tau^2 I)^{\frac{1}{2}} \tag{2.11}$$

quanto à função mínimo suave

$$\phi_\tau(X, S) := X + S - ((X-S)^2 + 4\tau^2 I)^{\frac{1}{2}} \tag{2.12}$$

pois vimos anteriormente que as propriedades e proposições pertinentes à uma podem ser estendidas também à outra.

Dessa forma, então, defina a seguinte aplicação

$$\begin{aligned} \Phi_\tau : S^{n\times n} \times \mathbb{R}^m \times S^{n\times n} &\rightarrow S^{n\times n} \times \mathbb{R}^m \times S^{n\times n} \\ \Phi_\tau(X, \lambda, S) &:= \begin{pmatrix} \sum_{i=1}^{m} \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi_\tau(X, S) \end{pmatrix} \end{aligned} \tag{2.13}$$

Tal aplicação anterior generaliza as condições de otimalidade do problema sujeitas às caracterizações das funções ϕ_τ que já tratam das condições da trajetória central.

Tendo como base as proposições 2.0.2 e 2.0.4, podemos então agora enunciar o seguinte teorema (cuja demonstração é imediata) que representa uma nova caracterização das condições da trajetória central dadas em (**??**) para PSD's:

Teorema 2.0.2. *Seja* Φ_τ *dada em* (2.13) *com a função* ϕ *dada por* (2.11) *ou* (2.12) *e tome* $\tau > 0$.
Então as seguintes afirmações são equivalentes:

(a) (X, λ, S) *satisfaz as condições da trajetória central dadas em* (**??**)

(b) (X, λ, S) *é uma solução do sistema de equações não-lineares* $\Phi_\tau(X, \lambda, S) = 0$.

No Capítulo seguinte, trataremos de propriedades mais específicas das funções ϕ, principalmente no que dizem respeito à sua diferenciabilidade.

Capítulo 3

Propriedades específicas da função ϕ

No presente Capítulo, abordaremos sobre algumas propriedades das funções ϕ já introduzidas anteriormente e analisaremos particularmente a diferenciabilidade das mesmas.

Anteriormente, tratávamos τ simplesmente como um parâmetro não-negativo ou em alguns casos, positivo. A partir de agora, sempre que mencionarmos τ estaremos nos referindo a uma variável independente. Tal modificação é necessária por algumas razões computacionais que serão detalhadas posteriormente quando tratarmos da solução das condições de otimalidade dadas em (**??**) através de um método do tipo suave que será ainda apresentado.

Para deixar mais clara também a notação a ser utilizada de agora em diante, façamos a seguinte substituição:

$$\phi(X, S, \tau) := \phi_\tau(X, S)$$

Então, estabelecemos

$$\phi(X, S, \tau) = X + S - (X^2 + S^2 + 2\tau^2 I)^{1/2} \tag{3.1}$$

para a função de Fischer-Burmeister suavizada de (2.11) e

$$\phi(X, S, \tau) = X + S - ((X - S)^2 + 4\tau^2 I)^{1/2} \tag{3.2}$$

para a função mínimo suavizada de (2.12).

Com as sucintas modificações, iniciaremos nossa análise às funções ϕ apresentando o Lema seguinte, que mostra que a função ϕ citada em (3.1) ou (3.2) é contínua em τ.

Lema 3.0.3. *Considere a função ϕ expressa em* (3.1) *ou* (3.2)*. Então, para quaisquer $X, S \in S^{n\times n}$ e $\tau > \nu > 0$, temos*

$$\kappa(\tau - \nu)I \succeq \phi(X, S, \nu) - \phi(X, S, \tau) \succ 0,$$

$$\kappa\tau I \succeq \phi(X, S, 0) - \phi(X, S, \tau) \succ 0.$$

onde κ denota uma constante positiva e independente de X, S, τ e ν e seu valor é

Observação 3.0.6. *A constante κ citada no Lema* (3.0.3) *tem seu valor conhecido:*

$\kappa = \sqrt{2}$ *para a função de Fischer-Burmeister suavizada dada em* (3.1) *e*

$\kappa = 2$ *para a função mínimo suavizada dada em* (3.2).

Demonstração. Fixe $X, S \in S^{n\times n}$ e $\tau, \nu > 0$. Agora, considere a função de Fischer-Burmeister suavizada dada em (3.1), ou seja,

$$\phi(X, S, \tau) = X + S - (X^2 + S^2 + 2\tau^2 I)^{1/2}$$

Seja $G := X^2 + S^2$ e escolha qualquer matriz P ortogonal (isto é $P^T = P^{-1}$) e $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tais que satisfaçam

$$G = P^T \, diag[\lambda_1, \ldots, \lambda_n] \, P$$

Então,

$$\begin{aligned}\phi(X,S,\tau) &= X+S-(G+2\tau^2 I)^{1/2}\\ &= X+S-P^T\ diag[(\lambda_1+2\tau^2)^{1/2},\ldots,(\lambda_n+2\tau^2)^{1/2}]\ P\end{aligned}$$

e assim, de maneira inteiramente análoga para $\phi(X,S,\nu)$, ou seja,

$$\phi(X,S,\nu) = X+S-P^T\ diag[(\lambda_1+2\nu^2)^{1/2},\ldots,(\lambda_n+2\nu^2)^{1/2}]\ P$$

Dessa forma, temos que

$$\phi(X,S,\nu)-\phi(X,S,\tau) = P^T\ diag[(\lambda_i+2\tau^2)^{1/2}-(\lambda_i+2\nu^2)^{1/2}]_{i=1}^n\ P$$

Além disso, como $G \succeq 0$ de tal maneira que $\lambda_i \geq 0$ para cada i, temos então que

$$0 < (\lambda_i+2\tau^2)^{1/2}-(\lambda_i+2\nu^2)^{1/2} \leq \sqrt{2}\ (\tau-\nu)$$

A segunda desigualdade anterior utiliza-se da seguinte observação:

Observação 3.0.7. *Considere* $h(\tau) := (\lambda+2\tau^2)^{1/2}$.

Tal função é diferenciável e convexa em $\mathbb{R}_{++}$ *(reais positivos), para qualquer* $\lambda \in \mathbb{R}_+$*. Como, Observe agora que* $h(\tau) = (\lambda+2\tau^2)^{1/2}$*, observe agora que*

$$\begin{aligned}h'(\tau) &= \frac{1}{2}(\lambda+2\tau^2)^{-1/2}\ \cdot\ 4\tau\\ &= 2(\lambda+2\tau^2)^{-1/2}\ \tau\\ &\leq \sqrt{2}\end{aligned}$$

A conclusão de que $h'(\tau) \leq \sqrt{2}$ *vem do fato que o único ponto crítico de* $h'(\tau)$ *(nesse caso, ponto de máximo global) ocorre quando* $\lambda = 0$ *e seu valor é* $\sqrt{2}$.

E assim,

$$h(\tau)-h(\nu) = (\lambda+2\tau^2)^{1/2}-(\lambda+2\nu^2)^{1/2} \leq h'(\tau) \leq \sqrt{2}$$

Daí, como $\tau > \nu > 0$, *vem*

$$h(\tau) - h(\nu) \leq h'(\tau)\ (\tau - \nu) \leq \sqrt{2}\ (\tau - \nu)$$

Então concluimos que

$$\sqrt{2}\ (\tau - \nu)\ I \succeq \phi(X, S, \nu) - \phi(X, S, \tau) \succ 0$$

Isso prova a primeira relação do Lema com $\kappa = \sqrt{2}$.

Considerando que a relação anterior valha para qualquer $\nu \in (0, \tau)$, e tomando $\nu \to 0$, chegamos na segunda relação a ser demonstrada. Além disso, temos que

$$\begin{aligned} \phi(X, S, 0) = \lim_{\tau \to 0} \phi(X, S, \tau) &= X + S - P^T\ diag[\lambda_1^{1/2}, \ldots, \lambda_n^{1/2}]\ P \\ &= X + S - G^{1/2} \\ &= X + S - (X^2 + S^2)^{1/2} \end{aligned}$$

Temos então que

$$0 < \phi(X, S, \nu) - \phi(X, S, \tau) \leq \sqrt{2}\ (\tau - \nu)$$

Com $\nu \to 0$, teremos então

$$0 < \phi(X, S, 0) - \phi(X, S, \tau) \leq \sqrt{2}\ \tau$$

E então fica

$$\sqrt{2}\ \tau\ I \succeq \phi(X, S, 0) - \phi(X, S, \tau) \succ 0$$

o que prova a segunda relação do Lema. A demonstração das duas relações considerando-se a função mínimo suavizada dada em (3.2) é inteiramente análoga.

□

Como consequência do Lema 3.0.3, segue-se o corolário:

Corolário 3.0.1. *Seja* ϕ *definida como em* (3.1) *ou em* (3.2) *e a constante* κ *do lema* (3.0.3)*. Dessa forma, as seguintes afirmações são verdadeiras:*

1.

$$\|\phi(X,S,\nu) - \phi(X,S,\tau)\|_F \leq \kappa\sqrt{n}\ (\tau - \nu)$$

vale para quaisquer $X, S \in S^{n\times n}$ *e* $\tau > \nu > 0$.

2.

$$\|\phi(X,S,0) - \phi(X,S,\tau)\|_F \leq \kappa\sqrt{n}\ \tau$$

vale para quaisquer $X, S \in S^{n\times n}$ *e* $\tau > 0$.

Demonstração. Sejam $\lambda_1, \ldots, \lambda_n$ os autovalores da matriz simétrica $\phi(X,S,\nu) - \phi(X,S,\tau)$. Pelo Lema 3.0.3, temos que

$$\kappa\ (\tau - \nu) \geq \lambda_i > 0$$

Fazendo a substituição $Z := \phi(X,S,\nu) - \phi(X,S,\tau)$ e aplicando a norma de Frobenius temos que

$$\begin{aligned}
\|\phi(X,S,\nu) - \phi(X,S,\tau)\|_F = \|Z\|_F &= \langle Z, Z\rangle = tr(ZZ^T) \\
&= \left[\sum_{i,j=1}^{n} z_{ij}^2\right]^{1/2} \\
&= (\lambda_1^2 + \ldots + \lambda_n^2)^{1/2}
\end{aligned}$$

Temos que $0 < \lambda_i \leq \kappa\ (\tau - \nu)$ implica em

$$\lambda_i^2 \leq [\kappa\ (\tau - \nu)]^2$$

E então,

$$\lambda_1^2 + \ldots + \lambda_n^2 \leq n\ [\kappa\ (\tau - \nu)]^2$$

que ainda implica que $(\lambda_1^2 + \ldots + \lambda_n^2)^{1/2} \leq \sqrt{n}\ \kappa\ (\tau - \nu)$ e obtemos que

$$\|Z\|_F = (\lambda_1^2 + \ldots + \lambda_n^2)^{1/2} \leq \sqrt{n}\ \kappa\ (\tau - \nu), \quad \tau > \nu > 0,$$

o que prova o item 1.

Para a demonstração do item 2, temos, do item 1, que

$$\|\phi(X, S, \nu) - \phi(X, S, \tau)\|_F = (\lambda_1^2 + \ldots + \lambda_n^2)^{1/2} \leq \sqrt{n}\ \kappa\ (\tau - \nu)$$

E para concluir, vem

$$\|\phi(X, S, 0) - \phi(X, S, \tau)\|_F \leq \sqrt{n}\ \kappa\ (\tau - 0) = \sqrt{n}\ \kappa\ \tau, \quad \tau > 0.$$

o que completa a demonstração do item 2. $\square$

Nosso principal objetivo agora é mostrar que as funções de Fischer-Burmeister e de mínimo suavizadas dadas em (3.1) e (3.2) são continuamente diferenciáveis em relação a X, S e τ. Esse importante resultado foi demonstrado em [**?**].

A prova que será dada nesse trabalho, em relação à diferenciabilidade das funções ϕ, é um tanto quanto diferente da apresentada em [**?**]. O lema seguinte inicia esse processo.

Lema 3.0.4. *Considere as matrizes* $A \in S_{++}^{n\times n}$, $B \in S_{+}^{n\times n}$. *Então,*

$$\|A^{1/2} - B^{1/2}\|_2 \leq \|A^{-1/2}\|_2 \cdot \|A - B\|_2$$

Demonstração. Veja [**?**], Seção 7.2, p.411. $\square$

Depois da apresentação dos resultados anteriores dessa Seção, torna-se mais apropriada a obtenção da fórmula para as derivadas das aplicações ϕ. Sem perda de generalidade, apenas para tornar nosso estudo mais específico, vamos direcionar nossa análise apenas à função de Fischer-Burmeister suavizada dada em (3.1) (o estudo para a função mínimo suavizada de (3.2) é realizado de maneira análoga).

Para esse fim, temos que provar que

$$\|\phi(X+U, S+V, \tau+\mu) - \phi(X, S, \tau) - \nabla\phi(X, S, \tau)(U, V, \mu)_2\| = o(|\|(U, V, \mu)\||) \quad (3.3)$$

vale para todo $(U, V, \mu) \in S^{n\times n} \times S^{n\times n} \times \mathbb{R} \to (0, 0, 0)$, e $\nabla\phi(X, S, \tau)$ designa um operador linear conveniente para a derivada de ϕ em (X, S, τ).

Para facilitar, vamos decompor a função ϕ da seguinte maneira:

$$\phi(X, S, \tau) = \phi_1(X, S, \tau) - \phi_2(X, S, \tau)$$

onde

$$\begin{aligned} \phi_1 &:= X + S, \\ \phi_2 &:= (X^2 + S^2 + 2\tau^2 I)^{1/2} \end{aligned}$$

É imediato ver que ϕ_1 é diferenciável e

$$\nabla\phi_1(X, S, \tau)(U, V, \mu) = U + V$$

Já para ϕ_2, o trabalho é mais minucioso. Definamos então

$$E := (X^2 + S^2 + 2\tau^2 I)^{1/2} \text{ com } E \succ 0$$

Considere agora

$$L_E[X] := EX + XE \quad (3.4)$$

como sendo o operador de Lyapunov correspondente a X. Devido ao fato de $E \succ 0$, a equação seguinte (conhecida como equação de Lyapunov)

$$L_E[X] = H$$

possui solução única em $S^{n\times n}$ para cada matriz H nesse conjunto, pelo Teorema 1.1.1. Por isso, podemos definir L_E^{-1}, a inversa de L_E (em outras palavras, $L_E^{-1}(H)$ denota a única matriz X que satisfaz $EX + XE = H$).

Para simplificar notações, defina então

$$D := ((X+U)^2 + (S+V)^2 + 2(\tau+\mu)^2 I)^{1/2}$$

E a partir daí, temos que

$$L_E[D-E] := E(D-E) + (D-E)E = ED + DE - 2E^2$$

e então,

$$\begin{aligned} L_E[D-E] + (D-E)^2 &= ED + DE - 2E^2 + D^2 - DE - ED + E^2 \\ &= D^2 - E^2 \end{aligned}$$

Aplicando a inversa de L_E, L_E^{-1}, em ambos os lados temos

$$\begin{aligned} L_E^{-1}(D^2 - E^2) &= L_E^{-1}(L_E[D-E] + (D-E)^2) \\ &= (D-E) + L_E^{-1}(D-E)^2 \end{aligned}$$

e ordenando os termos vem,

$$\begin{aligned} & E - D = L_E^{-1}[(D-E)^2 - (D^2 - E^2)] \\ = \; & L_E^{-1}[(E-D)^2 - [(X+U)^2 + (S+V)^2 + 2(\tau+\mu)^2 I - (X^2 + S^2 + 2\tau^2 I)]] \\ = \; & L_E^{-1}[(E-D)^2 - (X^2 + UX + XU + U^2 + S^2 + SV + VS + V^2 + 2\tau^2 I + \\ & +2\;\tau\mu I + 2\mu\tau I + 2\mu^2 I - X^2 - S^2 - 2\tau^2 I)] \\ = \; & L_E^{-1}[(E-D)^2 - (XU + UX + SV + VS + 4\tau\mu I + U^2 + V^2 + 2\mu^2 I)] \end{aligned}$$

Como L_E^{-1} é um operador linear, podemos ainda fazer

$$\begin{aligned} & \phi_2(X+U, S+V, \tau+\mu) - \phi_2(X,S,\tau) - \nabla\phi_2(X,S,\tau)(U,V,\mu) \\ = \; & -\nabla\phi_2(X,S,\tau)(U,V,\mu) - (E-D) \\ = \; & -\nabla\phi_2(X,S,\tau)(U,V,\mu) + L_E^{-1}[XU+UX+SV+VS+4\tau\mu I] + \\ + \; & L_E^{-1}[U^2+V^2+2\mu^2 I] - L_E^{-1}[(E-D)^2] \end{aligned} \tag{3.5}$$

Com tudo isso concluimos então que,

$$\|L_E^{-1}[U^2+V^2+2\mu^2 I]\|_F = O(\|\|(U,V,\mu)\|\|^2).$$

Agora, observe que

$$\|(E-D)^2\| = \|(E-D)(E-D)\| \le \|E-D\| \cdot \|E-D\| = \|E-D\|^2$$

E pelo Lema 3.0.4 temos ainda que

$$\begin{aligned} \|E-D\|_F^2 & \le \|E^{-1}\|_F^2 \; \|E^2-D^2\|_F^2 \\ & = \gamma_1 \|E^2-D^2\|_F^2 \\ & = \gamma_1 \|XU+UX+SV+VS+4\tau\mu I + U^2+V^2+2\mu^2 I\|_F^2 \\ & = O(\|\|(U,V,\mu)\|\|^2) \end{aligned}$$

onde $\gamma_1 = \|E^{-1}\|_F^2 > 0$, isto é, uma constante positiva, independente dos parâmetros U, V e μ. Toda essa situação recai em

$$\|L_E^{-1}[(E-D)^2]\|_F = O(\|\|(U,V,\mu)\|\|^2). \tag{3.6}$$

Dessa forma, se considerarmos

$$\nabla\phi_2(X,S,\tau)(U,V,\mu) := L_E^{-1}[XU+UX+SV+VS+4\tau\mu I],$$

chegamos à conclusão que, de (3.5) a aplicação ϕ_2 é diferenciável em (X, S, τ). Como já se constatou que ϕ_1 também é diferenciável nesse ponto, logo ϕ é diferenciável em (X, S, τ).

Considerando (3.6) e a função ϕ dada por (3.1), temos que

$$\nabla\phi(X, S, \tau)(U, V, \mu) := U + V - L_E^{-1}[XU + UX + SV + VS + 4\tau\mu I]$$

é contínua em (X, S, τ).

O fato de ϕ ser diferenciável em (X, S, τ). e o resultado dado no Lema anterior nos serão muito úteis para a demonstração do Teorema que segue.

Teorema 3.0.3. *Seja $X, S \in S^{n\times n}$ e $\tau \in \mathbb{R}_+$.*

1. *(a) Se ϕ é dada por (3.1) e $X^2 + S^2 + 2\tau^2 I \succ 0$, então ϕ é contínua diferenciável em (X, S, τ) com*

$$\nabla\phi(X, S, \tau)(U, V, \mu) = U + V - L_E^{-1}[XU + UX + SV + VS + 4\tau\mu I], \tag{3.7}$$

onde $E := (X^2 + S^2 + 2\tau^2 I)^{1/2}$.

2. *(b) Se ϕ é dada por (3.2) e $(X - S)^2 + 4\tau^2 I \succ 0$, então ϕ é contínua diferenciável em (X, S, τ) com*

$$\nabla\phi(X, S, \tau)(U, V, \mu) = U + V - L_E^{-1}[(X - S)(U - V) + (U - V)(X - S) + 8\tau\mu I], \tag{3.8}$$

onde $E := ((X - S)^2 + 4\tau^2 I)^{1/2}$.

Demonstração. A análise que fizemos anteriormente já garante a diferenciabilidade da função ϕ dada pela função de Fischer-Burmeister suavizada de (3.1).

Reciprocamente, considere

$$D := ((X - S + U - V)^2 + 4(\tau + \mu)^2 I)^{1/2}.$$

Como agora se trata da função mínimo, modificaremos a definição de E como segue a seguir:

$$E := ((X-S)^2 + 4\tau^2 I)^{1/2}.$$

Usando o fato de que

$$D^2 - E^2 = L_E[D-E] + (D-E)^2$$

apliquemos então a inversa de L_E, L_E^{-1}, em ambos os lados. Daí,

$$\begin{aligned} L_E^{-1}(D^2 - E^2) &= L_E^{-1}(L_E[D-E] + (D-E)^2) \\ &= (D-E) + L_E^{-1}(D-E)^2 \end{aligned}$$

e ordenando os termos vem,

$$\begin{aligned} E - D &= L_E^{-1}[(D-E)^2 - (D^2 - E^2)] \\ &= L_E^{-1}[(E-D)^2 - [(X-S+U-V)^2 + 4(\tau+\mu)^2 I - ((X-S)^2 + 4\tau^2 I)]] \\ &= L_E^{-1}[(E-D)^2 - (X^2 - XS + XU - XV - SX + S^2 - SU + SV \\ &\quad + UX - US + U^2 - UV - VX + VS - VU + V^2 + 4\tau^2 I + 8\tau\mu I + \\ &\quad + 4\mu^2 I - X^2 + XS + SX - S^2 - 4\tau^2 I)] \\ &= L_E^{-1}[(E-D)^2 - ((XU - XV - SU + SV) + (UX - US - VX + VS) + \\ &\quad + (U^2 - UV - VU + V^2) + 8\tau\mu I + 4\mu^2 I)] \\ &= L_E^{-1}[(E-D)^2 - ((X-S)(U-V) + (U-V)(X-S) + (U-V)^2 + 8\tau\mu I + 4\mu^2 I)] \end{aligned}$$

Como L_E^{-1} é um operador linear, podemos ainda fazer

$$\begin{aligned} &\phi(X+U, S+V, \tau+\mu) - \phi(X,S,\tau) - \nabla\phi(X,S,\tau)(U,V,\mu) \\ = \;& U + V - D + E - \nabla\phi(X,S,\tau)(U,V,\mu) \\ = \;& E - D + L_E^{-1}[(X-S)(U-V) + (U-V)(X-S) + 8\tau\mu I \\ = \;& L_E^{-1}[(E-D)^2] - L_E^{-1}[(U-V)^2 + 4\mu^2 I]. \end{aligned}$$

De acordo com o Lema (3.0.4) temos que

$$\begin{aligned}
\|(E-D)^2\|_F &\leq \|E-D\|_F^2 \\
&\leq \gamma_2\|E^2-D^2\|_F^2 \\
&= \gamma_2\|(X-S)(U-V)+(U-V)(X-S)+(U-V)^2+8\tau\mu I+4\mu^2 I\|_F^2 \\
&= O(\||(U,V,\mu)\||^2)
\end{aligned}$$

onde γ_2 corresponde a uma constante positiva independente de U, V e τ. Tudo isso implica que

$$\begin{aligned}
&\|\phi(X+U,S+V,\tau+\mu)-\phi(X,S,\tau)-\nabla\phi(X,S,\tau)(U,V,\mu)\|_F \\
\leq\ & \gamma_3\|(E-D)^2\|_F+\gamma_3\|(U-V)^2+4\mu^2 I\|_F \\
=\ & O(\||(U,V,\mu)\||^2)
\end{aligned}$$

onde γ_3 também diz respeito a uma constante positiva independente de U, V e τ. Essa análise nos garante a diferenciabilidade de ϕ considerando $\nabla\phi$ dado por (3.8). Analogamente à conclusão do primeiro item deste Teorema, se considerarmos

$$E = ((X-S)^2+4\tau^2 I)^{1/2} \succ 0$$

contínua em (X, S, τ), vemos de imediato que $\nabla\phi(X, S, \tau)$ também é contínuo em (X, S, τ). $\square$

Pelo Teorema 3.0.3 vimos que as funções ϕ dadas em (3.1) e em (3.2) são ambas contínuas diferenciáveis em (X, S, τ), com $\tau > 0$. Foi discutido e demonstrado que

$$\|\phi(X+U,S+V,\tau+\mu)-\phi(X,S,\tau)-\nabla\phi(X,S,\tau)(U,V,\mu)_2 = O(\||(U,V,\mu)\||^2), \quad (3.9)$$

apesar de termos sugerido a prova de (3.3). No nosso caso, a função ϕ, provar (3.9) foi suficiente para garantir a diferenciabilidade dessa função. No entanto, não podemos generalizar tal resultado. Um bom contra-exemplo é exposto a seguir:

$$f(x) := \begin{cases} x^2 \ \operatorname{sen}(\frac{1}{x}), \ se \ x \neq 0 \\ 0, \ se \ x = 0 \end{cases}$$

Veja o que acontece:

$$|f(x) - f(0) - f'(0)x| = O(|x|^2).$$

com $f'(0) = 0$, ou seja, satisfaz (3.9), porém tal função só é diferenciável na origem, o que não a torna contínua diferenciável.

Capítulo 4

O Algoritmo

Neste Capítulo, faremos uma abordagem sobre o aspecto computacional do nosso trabalho. O algoritmo que será apresentado neste Capítulo deverá resolver as condições de otimalidade dadas em (**??**), bem como a solução do PSD dado nas formas primal e dual. Para isso, utilizaremos muitos dos resultados citados e analisados até agora.

Optamos por uma idéia que está inserida no contexto de problemas de complementariedade não-linear. Sobre problemas dessa natureza, veja [**?**]. Além disso, τ será visto como uma variável independente. Então, definamos a aplicação

$$\begin{aligned} \Theta : S^{n\times n} \times \mathbb{R}^m \times S^{n\times n} \times \mathbb{R} \;&\rightarrow\; S^{n\times n} \times \mathbb{R}^m \times S^{n\times n} \times \mathbb{R} \\ \Theta(X, \lambda, S, \tau) \;&:=\; \begin{pmatrix} \displaystyle\sum_{i=1}^{m} \lambda_i A_i + S - C \\ A_i \bullet X - b_i, \;\; (i = 1, \ldots, m) \\ \phi(X, S, \tau) \\ \tau \end{pmatrix} \end{aligned} \tag{4.1}$$

onde ϕ é expressa por (3.1) ou por (3.2).

Além do fato de termos considerado τ como uma variável independente ao invés de um parâmetro, a nossa escolha também se diferenciou da feita em [**?**] pelo fato de termos acrescentado uma linha ao sistema a fim de que o mesmo se tornasse um sistema de equações não-lineares quadrado, como se vê a seguir

$$\Theta(X, \lambda, S, \tau) = 0 \tag{4.2}$$

e dessa forma $(X^*, \lambda^*, S^*, \tau^*)$ é solução de (4.2). Além de fazer com que o sistema de equações se tornasse quadrado, a linha adicional colocada nele também implica que resolver o sistema (4.2) é equivalente a resolver as condições de otimalidade dadas em (**??**), e não às da trajetória central. O método escolhido por nós continuará garantindo a continuidade e diferenciabilidade da função ϕ dadas no Teorema 3.0.3 o que, com certeza, não deixa de ser uma vantagem em relação a trabalhar com a função ϕ não-suavizada.

O algoritmo que seguirá adiante deverá ser solucionado pelo método de Newton (veja também [**?**]) e sua convergência global será obtida de acordo com uma vizinhança conveniente da trajetória central que é dada a seguir

$$N(\beta) = \left\{ (X, \lambda, S, \tau) \mid A_i \bullet X = b_i,\ \forall i = 1, \ldots, m,\ \sum_{i=1}^{m} \lambda_i A_i + S = C, \|\phi(X, S, \tau)\|_F \leq \beta\tau \right\},$$

onde $\beta > 0$.

Já em se tratando da convergência local rápida, a mesma será obtida utilizando-se um conveniente passo preditor (será abordado futuramente).

O algoritmo que será abordado a seguir trata-se do tipo preditor-corretor. Fazendo um paralelo com a Programação Linear, tal algoritmo consiste em se considerar um único passo "corretor"para a trajetória central logo após cada passo "preditor"para reduzir τ. O passo preditor, após a inicialização da execução do algoritmo, quando $k = 0$ (k, índice de iterações), move y_0 numa vizinhança limitada da trajetória central na fronteira de uma vizinhança mais ampla e o passo corretor então parte deste ponto e move para a próxima iteração, y_1, numa vizinhança limitada sucessivas vezes (variando

k) até finalmente concluir o processo. Sem contar que esse métodos preditor-corretor são uns dos mais usados atualmente na categoria de métodos primal-dual, devido ao fato de proporcionarem implementações com êxito.

Para simplificar notações, então considere:

$$W := (X, \lambda, S) \text{ e } W^k := (X^k, \lambda^k, S^k),$$

onde k denota o índice de iterações. Com tudo isso, considere a seguir o nosso método de suavilização para a solução do PSD:

Algoritmo 4.0.1. 1. *(P.0) Inicialização*

Escolha

$$W^0 = (X^0, \lambda^0, S^0) \in S^{n\times n} \times \mathbb{R}^m \times S^{n\times n}$$

com

$$\sum_{i=1}^{m} \lambda_i^0 A_i + S^0 = C \text{ e } A_i \bullet X^0 = b_i, \ (i = 1, \ldots, m)$$

Escolha $\tau_0, \beta_0 > 0$ *com* $\|\phi(X^0, S^0, \tau^0\|_F \leq \beta\tau_0$ *e coloque* $k := 0$.

Escolha $\hat{\sigma}, \alpha_1, \alpha_2 \in (0, 1)$.

2. *(P.1) Passo Preditor*

Seja

$$(\Delta W^k, \Delta\tau_k) = (\Delta X^k, \Delta\lambda^k, \Delta S^k, \Delta\tau_k) \in S^{n\times n} \times \mathbb{R}^m \times S^{n\times n} \times \mathbb{R}$$

solução do sistema

$$\nabla\Theta(W^k, \tau_k) \begin{pmatrix} \Delta W \\ \Delta\tau \end{pmatrix} = -\Theta(W^k, \tau_k). \tag{4.3}$$

Se

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F = 0 : \ \textit{pare}.$$

Caso contrário,

(a) (i) Se $\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k)_F > \beta\tau_k$, *então tome*

$$\hat{W}^k := W^k, \quad \hat{\tau}^k := \tau^k \quad e \quad \eta_k := 1$$

e vá para o passo (P.2)

(b) (ii) ou do contrário, seja $\eta_k = \alpha_1^s$, *onde* s *é um número natural com*

$$\begin{aligned} \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^r \tau_k)\|_F &\leq \beta\tau_k\alpha_1^r, \ r = 0, 1, \ldots, s \\ \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^{s+1} \tau_k)\|_F &> \beta\tau_k\alpha_1^{s+1}, \end{aligned}$$

e coloque

$$\hat{\tau}_k := \eta_k \tau_k \ e \ \hat{W}^k := \begin{cases} W^k, \ se \ s = 0 \\ W^k + \Delta W^k, \ caso \ contrário. \end{cases}$$

e vá para o passo (P.2).

3. *(P.2) Passo Corretor*

   Seja

$$(\Delta\hat{W}^k, \Delta\hat{\tau}_k) = (\Delta\hat{X}^k, \Delta\hat{\lambda}^k, \Delta\hat{S}^k, \Delta\hat{\tau}_k)$$

   a solução de

$$\nabla\Theta(\hat{W}^k, \hat{\tau}_k) \begin{pmatrix} \Delta\hat{W} \\ \Delta\hat{\tau} \end{pmatrix} = -\Theta(\hat{W}^k, \hat{\tau}_k) + \begin{pmatrix} 0 \\ (1 - \hat{\sigma}) \ \hat{\tau}_k \end{pmatrix} \tag{4.4}$$

   Seja $\hat{\eta}_k = \max\ \{1, \alpha_2, \alpha_2^2, \ldots\}$ *com*

$$\|\phi(\hat{X}^k + e\hat{t}a_k \Delta\hat{X}^k, \hat{S}^k + e\hat{t}a_k \Delta\hat{S}^k, \hat{\tau}_k + e\hat{t}a_k \Delta\hat{\tau}_k)\|_F \leq (1 - \hat{\sigma}\hat{\eta}_k) \ \beta\hat{\tau}_k. \tag{4.5}$$

   Coloque

$$W^{k+1} := \hat{W}^k + \hat{\eta}_k \Delta\hat{W}^k, \quad \tau_{k+1} := (1 - \hat{\sigma}\hat{\eta}_k)\hat{\tau}_k$$

   Faça $k := k + 1$ *e volte ao passo (P.1).*

As iteradas (X^k, λ^k, S^k) e $(\hat{X}^k, \hat{\lambda}^k, \hat{S}^k)$ geradas pelo Algoritmo são viáveis para as condições de otimalidade dadas em (**??**) no sentido que

$$\sum_{i=1}^{m} \lambda_i^k A_i + S^k = C, \quad A_i \bullet X^k = b_i, \ (i = 1, \ldots, m) \tag{4.6}$$

e

$$\sum_{i=1}^{m} \hat{\lambda}_i^k A_i + \hat{S}^k = C, \quad A_i \bullet \hat{X}^k = b_i, \ (i = 1, \ldots, m)$$

vale para todo $k \in \mathbb{N}$.

A viabilidade das iteradas (comentadas há pouco) juntamente com as Proposições 2.0.1 e 2.0.3 justificam o critério de terminação abordado no Passo (P.1) e tais resultados implicam que

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F = 0 \text{ se, e somente se, } W^k + \Delta W^k$$

é solução de (**??**).

Para a análise teórica que faremos sobre o Algoritmo, suponhamos, por hipótese, que esse critério não vale se a sequência gerada pelo Algoritmo for infinita (pois dessa forma, o método não convergiria). Além do mais, a regra de atualização para τ_{k+1} no Passo (P.2) segue imediatamente da última linha do sistema linear (4.4) no Passo corretor que resulta em $\Delta\hat{\tau}_k = -\hat{\sigma}\hat{\tau}_k$. Tal regra é equivalente a

$$\tau_{k+1} = \hat{\tau}_k + \hat{\eta}_k \Delta\hat{\tau}_k$$

Para finalizar, frisamos que o sistema de equações lineares deve ser resolvido em ambos os Passos, preditor e corretor, e ainda com possíveis matrizes diferentes $\nabla\Theta(W, \tau)$. Todavia, sem comprometer os resultados sobre sua convergência, propomos uma sutil modificação no Algoritmo que acarretará em resolver unicamente um sistema de equações lineares no passo preditor ou dois sistemas com a mesma matriz coeficiente. A modificação requerida consiste no seguinte:

Se o Passo preditor for aceito com $\eta_k < 1$, então pule o Passo corretor, ou seja, coloque:

$$\begin{aligned} W^{k+1} &:= W^k + \Delta W^k, \\ \tau_{k+1} &:= \eta_k \tau_k \\ \text{Faça } k &:= k+1 \text{ e volte ao Passo (P.1).} \end{aligned}$$

Agora, formalmente, analisaremos as Propriedades do Algoritmo, mostrando inicialmente que o mesmo é bem definido. Para isso, queremos provar que os sistemas lineares (4.3) e (4.4) têm uma única solução. Com esse objetivo, aprimoraremos algumas propriedades específicas do operador de Lyapunov dado em (3.4), iniciando pelo Lema seguinte.

Lema 4.0.5. *Sejam $A, B \in S^{n\times n}_{++}$. Então, valem as seguintes afirmações:*

1. *L_A e L_B são auto-adjuntas*
2. *L_A^{-1} e L_B^{-1} são auto-adjuntas*
3. *$L_A \circ L_B$ e $L_B \circ L_A$ são fortemente monótonas*
4. *$L_A^{-1} \circ L_B$ e $L_B^{-1} \circ L_A$ são fortemente monótonas*

Demonstração. (1)

É fácil ver que L_A é auto-adjunta. Observe:

$$\begin{aligned}
L_A[X] \bullet Y &= tr(L_A[X]\ Y) \\
&= tr((AX + XA)Y) \\
&= tr(AXY) + tr(XAY) \\
&= tr(XYA) + tr(XAY) \\
&= tr(X(AY + YA)) \\
&= tr(X\ L_A[Y]) \\
&= X \bullet L_A[Y]
\end{aligned}$$

para quaisquer $X, Y \in S^{n \times n}$. Para L_B, a demonstração é inteiramente análoga.

(2)

Para demonstrar que L_A^{-1} é auto-adjunta, usaremos o item (1) e o fato de que L_A^{-1} é a inversa de L_A. Veja:

$$\begin{aligned}
L_A^{-1}[X] \bullet Y &= \langle L_A^{-1}[X], Y \rangle \\
&= \langle L_A^{-1}[X], L_A(L_A^{-1}[Y]) \rangle, \text{ e pelo item anterior, vem} \\
&= \langle L_A(L_A^{-1}[X]), L_A^{-1}[Y] \rangle \\
&= \langle X, L_A^{-1}[Y] \rangle \\
&= X \bullet L_A^{-1}[Y]
\end{aligned}$$

para quaisquer $X, Y \in S^{n \times n}$. A demonstração para L_B^{-1} é inteiramente análoga.

(3)

Utilizando o resultado do item (1), temos que

$$\begin{aligned}
(L_A \circ L_B[X]) \bullet X &= \langle L_A \circ L_B[X], X\rangle \\
&= \langle L_B[X], L_A[X]\rangle, \text{ pelo item (1)} \\
&= \langle BX + XB, AX + XA\rangle \\
&= tr(BX + XB)(AX + XA) \qquad (4.7)\\
&= tr(BXAX + XBAX + BXXA + XBXA) \\
&= 2\ tr(BXAX) + tr(X^2(BA + AB)) \\
&= 2\|B^{1/2}XA^{1/2}\|_F^2 + tr[X(BA + AB)X],\ \forall X \in S^{n\times n}
\end{aligned}$$

Como $BA(AB)$ é similar a $B^{1/2}AB^{1/2}(A^{1/2}BA^{1/2})$, pois

$$BA = B^{1/2}(B^{1/2}AB^{1/2})B^{-1/2}$$

e ainda $A, B \succeq 0$, segue que BA e AB tem autovalores reais não-negativos. Portanto,

$$BA + AB \succeq 0 \tag{4.8}$$

Além disso tudo, considerando a aplicação

$$X \mapsto \|B^{1/2}XA^{1/2}\|_F$$

que define uma norma e sabendo que em espaços de dimensão finita, todas as normas são equivalentes, $\exists\ \mu > 0$ tal que

$$\|B^{1/2}XA^{1/2}\|_F \geq \mu\ \|X\|_F,\ \forall\ X \in S^{n\times n} \tag{4.9}$$

Relacionando as passagens (4.7), (4.8) e (4.9), obtemos

$$(L_A \circ L_B[X]) \bullet X \geq 2\ \|B^{1/2}XA^{1/2}\|_F^2 \geq 2\mu^2\|X\|_F^2$$

ou seja, $L_A \circ L_B$ é fortemente monótona em $S^{n\times n}$ pois

$$\langle L_A \circ L_B[X] - L_A \circ L_B[Y], X - Y\rangle = \langle L_A \circ L_B[X-Y], X-Y\rangle$$

Permutando A e B provamos que $L_B \circ L_A$ é fortemente monótona de forma análoga.

(4)

Fazendo $Y = L_A^{-1}[X]$ para $\forall X \in S^{n\times n}$ e por (1) temos que

$$\begin{aligned}
(L_A^{-1} \circ L_B[X]) \bullet X &= \langle L_A^{-1} \circ L_B[X], X\rangle \\
&= \langle L_A^{-1} \circ L_B \circ L_A[Y], L_A[Y]\rangle \\
&= \langle L_B \circ L_A[Y], Y\rangle \\
&= (L_B \circ L_A[Y]) \bullet Y
\end{aligned}$$

e como $L_B \circ L_A$ é fortemente monótona pelo item (3), concluimos, a partir disso, que $L_A^{-1} \circ L_B$ é fortemente monótona. Para $L_B^{-1} \circ L_A$, a prova é inteiramente análoga. $\square$

Nosso objetivo agora é mostrar que os sistemas lineares (4.3) e (4.4) citados no Algoritmo possuem solução única. Mas, para isso, mostraremos que a aplicação linear $\nabla\Theta(X, \lambda, S, \tau)$ é inversível. Iniciemos pela seguinte Hipótese:

Hipótese 4.0.1. *As matrizes A_i, $(i = 1, \ldots, m)$ são linearmente independentes, isto é,*

$$\sum_{i=1}^{m} \alpha_i A_i = 0, \quad \alpha_i \in \mathbb{R} \Rightarrow \alpha_i = 0, \quad \forall\ i = 1, \ldots, m$$

Agora com base no Lema (4.0.5) e na Hipótese (4.0.1) já podemos provar que a aplicação $\nabla\Theta(X, \lambda, S, \tau)$ é bijetiva e isso implica que as direções preditora $(\Delta X^k, \Delta\lambda^k, \Delta S^k, \Delta\tau_k)$ e corretora $(\Delta\hat{X}^k, \Delta\hat{\lambda}^k, \Delta\hat{S}^k, \Delta\hat{\tau}_k)$ estão bem definidas, como enuncia a Proposição a seguir.

Proposição 4.0.5. *Suponhamos que vale a Hipótese* (4.0.1)*. Então a aplicação linear* $\nabla\Theta(X, \lambda, S, \tau)$ *com* ϕ *dada por* (3.1) *ou* (3.2) *é bijetiva para todo* $(X, \lambda, S, \tau) \in S^{n\times n} \times \mathbb{R}^m \times S^{n\times n} \times \mathbb{R}_{++}$.

Demonstração. Para a prova, vamos considerar a função ϕ dada em (3.1). Já para a função mínimo suavizada dada em (3.2) a prova é inteiramente análoga.

Seja $(X, \lambda, S, \tau) \in S^{n\times n} \times \mathbb{R}^m \times S^{n\times n} \times \mathbb{R}_{++}$. Como $\nabla\Theta(X, \lambda, S, \tau)$ é uma aplicação linear de $S^{n\times n} \times \mathbb{R}^m \times S^{n\times n} \times \mathbb{R}$ sobre ele mesmo, devemos mostrar apenas que tal aplicação é injetiva. Para isto, é suficiente mostrar que o sistema

$$\nabla\Theta(X, \lambda, S, \tau)(\Delta X, \Delta\lambda, \Delta S, \Delta\tau) = (0, 0, 0, 0)$$

ou equivalentemente o sistema

$$\begin{aligned} \sum_{i=1}^{m} \Delta\lambda_i A_i + \Delta S &= 0 && (4.10) \\ A_i \bullet \Delta X &= 0, \quad (i = 1, \ldots, m) && (4.11) \\ \nabla\phi(X, S, \tau)(\Delta X, \Delta S, \Delta\tau) &= 0 && (4.12) \\ \Delta\tau &= 0 && (4.13) \end{aligned}$$

tem $\Delta X, \Delta\lambda, \Delta S, \Delta\tau) = (0, 0, 0, 0)$ como sua única solução a trivial. Por (4.13), temos que $\Delta\tau = 0$.

Fazendo $E := (X^2 + S^2 + 2\tau^2 I)^{1/2}$, temos por (4.12) e pelo Teorema (3.0.3) que

$$\nabla\phi(X, S, \tau)(\Delta X, \Delta S, \Delta\tau) = 0$$

E, então, temos que

$$\Delta X + \Delta S - L_E^{-1}[X\ \Delta X + \Delta X\ X + S\ \Delta S + \Delta S\ S] = 0.$$

Agora, aplicando L_E em ambos os lados da equação anterior, vem

$$L_E[\Delta X + \Delta S - L_E^{-1}[X\ \Delta X + \Delta X\ X + S\Delta S + \Delta S\ S]] = L_E[0] = 0$$

E rearranjando os termos, fica

$$\begin{aligned}
0 &= E\ \Delta X + \Delta X\ E + E\ \Delta S + \Delta S\ E - X\ \Delta X - \Delta X\ X - S\ \Delta S - \Delta S\ S \\
&= (E\ \Delta X - X\ \Delta X) + (\Delta X\ E - \Delta X\ X) + (E\ \Delta S - S\ \Delta S) + (\Delta S\ E - \Delta S\ S) \\
&= (E - X)\Delta X + \Delta X\ (E - X) + (E - S)\Delta S + \Delta S\ (E - S) \\
&= L_{E-X}[\Delta X] + L_{E-S}[\Delta S]
\end{aligned}$$

Note agora que

$$E^2 - S^2 = [(X^2 + S^2 +^2 \tau^2 I)^{1/2}]^2 - S^2 = X^2 + 2\tau^2 I \succeq 0.$$

Observação 4.0.8. *Considere E da mesma forma como foi tomado nesta Proposição e $E \succ 0$ (pela definição do operador de Lyapunov). Se vale*

$$E^2 - S^2 \succeq 0,$$

então vale

$$E - S \succeq 0.$$

Demonstração. Temos, por hipótese, que $E^2 - S^2 \succeq 0$. Temos que mostrar que $E - |S| \succeq 0$, onde denotaremos $|S| := (S^2)^{1/2}$, o que implica que $E - S \succeq 0$.

Vamos então supor, por absurdo, que $E - S \not\succeq 0$, ou seja, existe $v \in \mathbb{R}^n$ não-nulo e $\lambda \in (-\infty, 0)$ com

$$(E - |S|)v = \lambda v.$$

Como $E \succeq 0$, então temos para P uma matriz ortogonal que

$$PEP^T = \begin{pmatrix} \tilde{E} & 0 \\ 0 & 0 \end{pmatrix}$$

onde $\tilde{E}$ é alguma sub-matriz definida positiva. Como $E^2 - S^2 \succeq 0$, devemos ter também

$$P|S|P^T = \begin{pmatrix} \tilde{S} & 0 \\ 0 & 0 \end{pmatrix}$$

onde $\tilde{S}$ é alguma sub-matriz semi-definida positiva. Então

$$\begin{pmatrix} (\tilde{E} - \tilde{S})(Pv_1) \\ 0 \end{pmatrix} = P(E - |S|)P^T Pv = P(E - |S|)v = \lambda Pv,$$

o que implica em $Pv_1 \neq 0$. Como

$$(E + |S|)(E - |S|) + (E - |S|)(E + |S|) = E^2 - S^2 + E^2 - S^2 = 2E^2 - 2S^2 \succeq 0,$$

podemos então obter

$$\begin{aligned} 0 &\leq v^T((E + |S|)(E - |S|) + (E - |S|)(E + |S|))v \\ &= 2\ \lambda v^T(E + |S|)v \\ &= 2\ \lambda(Pv)^T P(E + |S|)P^T Pv \\ &= 2\ \lambda(Pv_1)^T(\tilde{E} + \tilde{S})(Pv_1) < 0, \end{aligned}$$

onde a última desigualdade segue do fato de que $\tilde{E} + \tilde{S} \succ 0$, $(Pv_1) \neq 0$, e $\lambda < 0$, o que caracteriza uma contradição. Logo, $E - S \succeq 0$. □

Como pela Observação anterior, $E - S \succ 0$, existe L^{-1}_{E-S}, a inversa de L_{E-S} e fica então

$$L^{-1}_{E-S} \circ L_{E-X}[\Delta X] + \Delta S = 0 \tag{4.14}$$

De (4.10) e (4.14) temos que

$$L^{-1}_{E-S} \circ L_{E-X}[\Delta X] = \sum_{i=1}^{m} \Delta\lambda_i A_i$$

Tomando o produto escalar com ΔX, obtemos

$$L_{E-S}^{-1} \circ L_{E-X}[\Delta X] \bullet \Delta X = \sum_{i=1}^{m} \Delta\lambda_i A_i \bullet \Delta X$$

E então usando (4.11) vem

$$0 = L_{E-S}^{-1} \circ L_{E-X}[\Delta X] \bullet \Delta X - \sum_{i=1}^{m} \Delta\lambda_i \underbrace{A_i \bullet \Delta X}_{=0} = L_{E-S}^{-1} \circ L_{E-X}[\Delta X] \bullet \Delta X \quad (4.15)$$

Utilizando-se do fato que $E - S \succ 0$ e $E - X \succ 0$ (pela própria definição do operador de Lyapunov correspondente), segue do Lema (4.0.5), item (4), que o operador $L_{E-S}^{-1} \circ L_{E-X}$ é fortemente monótono. E então, por (4.15) obtemos $\Delta X = 0$. Concluímos também que $\Delta S = 0$, por (4.14). E finalmente, por (4.10) e pela Hipótese (4.0.1) chegamos em $\Delta\lambda = 0$, o que completa a demonstração. □

Com base nesse último resultado, podemos então enunciar o seguinte Teorema:

Teorema 4.0.4. *O Algoritmo é bem definido sobre a Hipótese* (4.0.1). *Além disso, as iteradas $W^k = (X^k, \lambda^k, S^k)$ e τ_k e também $\hat{W}^k = (\hat{X}^k, \hat{\lambda}^k, \hat{S}^k)$ e $\hat{\tau}_k$ pertencem à vizinhança $N(\beta)$.*

Demonstração. Com o que já foi analisado e demonstrado na Proposição (4.0.5) nos resta mostrar que as duas estratégias de backtracking dadas nos Passos (P.1) e (P.2) do Algoritmo estão bem definidas.

Para isso, iniciemos pelo Passo (P.1). Considerando-se a hipótese que o Algoritmo gera uma sequência infinita, o critério de terminação no Passo (P.1) não seria satisfeito, $\forall k \in \mathbb{N}$. Portanto, temos o seguinte:

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F > 0$$

Considerando contínua a aplicação

$$\tau \mapsto \|\phi(X, S, \cdot)\|_F$$

implica em afirmarmos que a estratégia de backtracking no Passo $(P.1)$ termina num número finito de loops internos. Com isso, $\hat{W}^k$ e $\hat{\tau}_k$ estão bem definidos e satisfazem a condição de vizinhança estabelecida em $N(\beta)$, isto é, $(\hat{W}^k, \hat{\tau}_k) \in N(\beta)$.

Agora, consideremos o procedimento do comprimento do Passo $(P.2)$. Definamos então a aplicação

$$\psi(X, S, \tau) := \|\phi(X, S, \tau)\|_F$$

Usando a teoria de cálculo diferencial, vem

$$\begin{aligned} \psi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)(\Delta\hat{X}^k, \Delta\hat{S}^k, \Delta\hat{\tau}_k) &= \frac{\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k) \bullet \nabla\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)(\Delta\hat{X}^k, \Delta\hat{S}^k, \Delta\hat{\tau}_k)}{\|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F} \\ &= -\|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F, \quad \text{por (4.4).} \end{aligned} \tag{4.16}$$

Agora, vamos supor, que os cálculos para o comprimento do Passo $\hat{\eta}_k$ não terminem em um número finito de loops em $(P.2)$. Temos então

$$\|\phi(\hat{X}^k + \alpha_2^t\Delta\hat{X}^k, \hat{S}^k + \alpha_2^t\Delta\hat{S}^k, \hat{\tau}_k + \alpha_2^t\Delta\hat{\tau}_k)\|_F > (1 - \hat{\sigma}\alpha_2^t)\ \beta\hat{\tau}_k, \quad \forall\ t \in \mathbb{N}.$$

Considerando $\beta\hat{\tau}_k \geq \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F$, então implica que

$$\|\phi(\hat{X}^k + \alpha_2^t\Delta\hat{X}^k, \hat{S}^k + \alpha_2^t\Delta\hat{S}^k, \hat{\tau}_k + \alpha_2^t\Delta\hat{\tau}_k)\|_F > (1 - \hat{\sigma}\alpha_2^t)\ \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F$$

e multiplicando ambos os lados da desigualdade por $\dfrac{1}{\alpha_2^t}$ pois $\alpha_2^t > 0$, logo, $\forall\ t \in \mathbb{N}$, vem

$$\frac{\|\phi(\hat{X}^k + \alpha_2^t\Delta\hat{X}^k, \hat{S}^k + \alpha_2^t\Delta\hat{S}^k, \hat{\tau}_k + \alpha_2^t\Delta\hat{\tau}_k)\|_F - \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F}{\alpha_2^t} > -\hat{\sigma}\ \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F$$

Tomando o limite com $t \to \infty$ e considerando (4.16), obtemos

$$-\|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F = \psi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)(\Delta\hat{X}^k, \Delta\hat{S}^k, \Delta\hat{\tau}_k) \geq -\hat{\sigma}\ \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F$$

Observe também que se $\hat{\sigma} \in (0, 1)$, então $\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k) = 0$.

Portanto, temos então que,

$$\|\phi(\hat{X}^k + \alpha_2^t \Delta \hat{X}^k, \hat{S}^k + \alpha_2^t \Delta \hat{S}^k, \hat{\tau}_k + \alpha_2^t \Delta \hat{\tau}_k)\|_F \to 0, \quad \text{quando } t \to \infty,$$

$$\text{enquanto } \ (1 - \hat{\sigma}\alpha_2^t)\ \beta\hat{\tau}_k \to \beta\hat{\tau}_k, \quad \text{quando } t \to \infty,$$

o que contradiz a hipótese de que os cálculos para o comprimento do Passo $\hat{\eta}_k$ não terminam em um número finito de loops em $(P.2)$, o que acarreta afirmar que também está bem definida a busca linear dada nesse Passo.

E, para finalizar, as regras de atualização do Algoritmo garantem que as iteradas $W^k = (X^k, \lambda^k, S^k)$ e τ_k e também $\hat{W}^k = (\hat{X}^k, \hat{\lambda}^k, \hat{S}^k)$ e $\hat{\tau}_k$ pertencem à vizinhança $N(\beta)$. $\square$

A seguir, abordaremos sobre a convergência do presente Algoritmo, tanto da global quanto da superlinear local.

Capítulo 5

Convergência

Como já vimos que o Algoritmo exposto na seção anterior está bem definido, agora nosso trabalho será analisar sua convergência. Iniciemos abordando a convergência global.

5.1 Convergência global

Como hipótese, para essa seção, assumiremos que o Algoritmo gera uma sequência infinita. Com respeito à tal hipótese, provaremos nesta seção, que todo ponto de acumulação da sequência $W^k = (X^k, \lambda^k, S^k)$ gerada pelo Algoritmo é uma solução das condições de otimalidade dadas em (**??**). Para esse fim, precisaremos da seguinte Proposição:

Proposição 5.1.1. *Se a sequência $W^k = (X^k, \lambda^k, S^k)$ gerada pelo Algoritmo tem um ponto de acumulação, então a sequência τ_k converge para zero.*

Demonstração. Pelo fato da sequência τ_k ser monótona decrescente e limitada inferiormente por zero, ela converge a um número não-negativo, que chamaremos de τ_*. Se $\tau_* = 0$, nada há a demonstrar.

Então, suponhamos que $\tau_* > 0$. Com isso, a regra de atualização no Passo $(P.1)$ do Algoritmo dá

$$\hat{W}^k = W^k, \quad \hat{\tau}_k = \tau_k, \quad \eta_k = 1 \ \text{ para todo k suficientemente grande} \tag{5.1}$$

Subsequencionando se necessário, suponhamos sem perda de generalidade que podemos estender (5.1) para $\forall\ k \in \mathbb{N}$. Dessa forma, as regras de atualização no Passo $(P.2)$ ficam assim obtidas:

$$\tau_k = \tau_0 \prod_{j=0}^{k-1} (1 - \hat{\sigma}\hat{\eta}_j).$$

Como, por hipótese, temos que $\tau_k \to \tau_* > 0$, concluimos então que

$$\lim_{k\to\infty} \hat{\eta}_k = 0.$$

Devido a isso, o comprimento do Passo $\hat{\rho}_k := \dfrac{\hat{\eta}_k}{\alpha_2}$ não satisfaz a busca linear dada em (4.5) $\forall\ k \in \mathbb{N}$ suficientemente grande, o que acarreta em

$$\|\phi(\hat{X}^k + \hat{\rho}_k \Delta\hat{X}^k, \hat{S}^k + \hat{\rho}_k \Delta\hat{S}^k, \hat{\tau}_k + \hat{\rho}_k \Delta\hat{\tau}_k)\|_F > (1 - \hat{\sigma}\hat{\rho}_k)\ \beta\hat{\tau}_k \tag{5.2}$$

para todos estes $k \in \mathbb{N}$.

Agora, considere $W^* = (X^*, \lambda^*, S^*)$ um ponto de acumulação da sequência W^k e tome também uma subsequência ${W^k}_K$ de tal maneira que

$${W^k}_K \to W^*.$$

Como $\tau_* > 0$, baseando-se em (5.1) e na Proposição (4.0.5) temos que a subsequência correspondente

$$(\Delta\hat{W}^k, \Delta\hat{\tau}_k)_K \to (\Delta\hat{W}^*, \Delta\hat{\tau}_*) = (\Delta\hat{X}^*, \Delta\hat{\lambda}^*, \Delta\hat{S}^*, \Delta\hat{\tau}_*),$$

onde $(\Delta\hat{W}^*, \Delta\hat{\tau}_*)$ é uma solução do sistema linear

$$\nabla\Theta(W^*, \tau_*) \begin{pmatrix} \Delta\hat{W} \\ \Delta\hat{\tau} \end{pmatrix} = -\Theta(W^*, \tau_*) + \begin{pmatrix} 0 \\ (1 - \hat{\sigma})\ \tau_* \end{pmatrix} \tag{5.3}$$

Em particular, a sequência $(\Delta\hat{W}^k, \Delta\hat{\tau}_k)_K$ é limitada. Baseando-se no fato de que $\{\hat{\rho}_k\}_K \to 0$ e fazendo $k \to \infty$ em K, obtemos de (5.1), de (5.2) e da continuidade da função $\phi(\cdot,\cdot,\cdot)$ que

$$\|\phi(X^*, S^*, \tau_*)\|_F \geq \beta\tau_*. \tag{5.4}$$

Por outro lado, também nos referindo a (5.1) e (5.2), e além disso à definição da vizinhança $N(\beta)$ e ao Teorema (4.0.4) obtemos

$$\begin{aligned}
\|\phi(\hat{X}^k + \hat{\rho}_k\Delta\hat{X}^k, \hat{S}^k + \hat{\rho}_k\Delta\hat{S}^k, \hat{\tau}_k + \hat{\rho}_k\Delta\hat{\tau}_k)\|_F &> (1 - \hat{\sigma}\hat{\rho}_k)\ \beta\hat{\tau}_k \\
&= (1 - \hat{\sigma}\hat{\rho}_k)\ \beta\tau_k \\
&\geq (1 - \hat{\sigma}\hat{\rho}_k)\ \|\phi(X^k, S^k, \tau_k)\|_F
\end{aligned}$$

para todo $k \in \mathbb{N}$ suficientemente grande. Com vista em (5.1), e multiplicando ambos os lados da desigualdade estrita anterior por $\dfrac{1}{\hat{\rho}_k}$ pois $\hat{\rho}_k > 0$, vem

$$\frac{\|\phi(\hat{X}^k + \hat{\rho}_k\Delta\hat{X}^k, \hat{S}^k + \hat{\rho}_k\Delta\hat{S}^k, \hat{\tau}_k + \hat{\rho}_k\Delta\hat{\tau}_k)\|_F - \|\phi(X^k, S^k, \tau_k)\|_F}{\hat{\rho}_k} > -\hat{\sigma}\ \|\phi(X^k, S^k, \tau_k)\|_F.$$

Para finalizar, consideremos $\psi(X, S, \tau) := \|\phi(X, S, \tau)\|_F$ e utilizando (4.16), obtemos para $k \to \infty$ que

$$-\|\phi(X^*, S^*, \tau_*)\|_F \geq -\hat{\sigma}\|\phi(X^*, S^*, \tau_*)\|_F,$$

admitindo evidentemente que ψ seja continuamente diferenciável em (X^*, S^*, τ_*). E, além disso, com $\hat{\sigma} \in (0, 1)$, isso implica que

$$\|\phi(X^*, S^*, \tau_*)\|_F = 0,$$

ou seja, uma contradição com respeito a (5.4), o que demonstra a Proposição. □

Uma consequência simples da Proposição (5.1.1) já nos faz obter o resultado seguinte que garante a convergência global do Algoritmo. É bom ressaltar que tal convergência (a global) não tem nenhuma influência do Passo Preditor do Algoritmo, dependendo exclusivamente do Passo Corretor. Eis o Teorema:

Teorema 5.1.1. *Todo ponto de acumulação da sequência*

$$W^k = (X^k, \lambda^k, S^k)$$

gerada pelo Algoritmo é uma solução das condições de otimalidade dadas em (**??**)

Demonstração. Considere $W^* = (X^*, \lambda^*, S^*)$ um ponto de acumulação da sequência $W^k = (X^k, \lambda^k, S^k)$ gerada pelo Algoritmo e seja ${W^k}_K$ uma subsequência convergindo para W^*.

Pela Proposição (5.1.1), temos que

$$\lim_{k\to\infty} \tau_k = 0.$$

Como todas as iteradas pertencem à vizinhança $N(\beta)$ (pelo Teorema (4.0.4)), obtemos ainda

$$\|\phi(X^*, S^*, 0)\|_F = \lim_{k\in K} \|\phi(X^k, S^k, \tau_k)\|_F \leq \lim_{k\in K} \beta\tau_k = 0.$$

Com isso, por (4.6) e pelas Proposições (2.0.1) e (2.0.3), temos que $W^* = (X^*, \lambda^*, S^*)$ é uma solução das condições de otimalidade dadas em (**??**). $\square$

5.2 Convergência superlinear local

O Algoritmo e suas propriedades específicas serão analisadas nesta Seção. Nela mostraremos que a sequência τ_k converge a zero superlinearmente. Mas para isso, alguns resultados precedentes são necessários. Iniciemos pela Hipótese a seguir:

Hipótese 5.2.1. *A sequência τ_k gerada pelo Algoritmo converge para zero e ainda*

$$\left|\left|\left| \begin{pmatrix} \Delta W^k \\ \Delta \tau_k \end{pmatrix} \right|\right|\right| = O(\tau_k), \tag{5.5}$$

onde $(\Delta W^k, \Delta \tau_k)$ *denota a direção de busca calculada em* (4.3)

Pelo fato de a Proposição (5.1.1) dar uma condição suficiente para a sequência τ_k convergir a zero, justifica a Hipótese (5.2.1). Analisemos agora a segunda condição abordada. Assumimos inicialmente que a sequência de operadores inversos

$$\nabla \Theta(W^k, \tau_k)^{-1}$$

permanece limitada quando $k \to \infty$. Dessa forma, tomando (4.3) obtemos de tal sistema linear que (5.5) vale se considerarmos que o lado direito de (4.3) é da ordem $O(\tau_k)$. Todavia, observando a viabilidade das iteradas (por (4.6)) aliado ao fato de que todas elas pertencem à vizinhança $N(\beta)$ (pelo Teorema (4.0.4)), isso se torna mais óbvio. Tudo isso implica em

$$\begin{aligned} |||\Theta(W^k, \tau_k)||| &= \sqrt{\|\phi(X^k, S^k, \tau_k)\|_F^2 + \tau_k^2} \\ &\leq \|\phi(X^k, S^k, \tau_k)\|_F + \tau_k \\ &\leq \beta\tau_k + \tau_k \\ &= O(\tau_k). \end{aligned}$$

Em suma, tal relação também vale se substituirmos o lado direito de (4.3) por $-\Theta(W^k, 0)$ e dessa forma então, a partir do Corolário (3.0.1) e do Teorema (4.0.4) temos que

$$\begin{aligned} |||\Theta(W^k, 0)||| &= \|\phi(X^k, S^k, 0)\|_F \\ &\leq \|\phi(X^k, S^k, \tau_k) - \phi(X^k, S^k, 0)\|_F + \|\phi(X^k, S^k, \tau_k)\|_F \\ &\leq k\,\sqrt{n}\,\tau_k + \beta\,\tau_k \\ &= O(\tau_k), \end{aligned}$$

onde k denta a constante do Lema (3.0.3). Particularmente, tal modificação citada anteriormente faz permanecerem válidas as propriedades de convergência do Algoritmo, tanto a global quanto a superlinear. A Hipótese a seguir dá condições suficientes da Hipótese (5.2.1) se tornar satisfeita.

Hipótese 5.2.2. *Seja* (X^*, λ^*, S^*) *uma solução das condições de otimalidade* (**??**) *tal que*

1. *(Complementariedade estrita)*

   $X^* + S^* \succ 0$

   Soluções primal-dual do tipo (X^*, λ^*, S^*) *com tal propriedade são conhecidas como soluções estritamente complementares. O Teorema* 2.4 *de [*?*] (p.*28*) garante que no mínimo uma dessas soluções existe. Num dado problema em Programação Linear, por exemplo, pode haver soluções primal-dual múltiplas, algumas estritamente complementares e outras não.*

2. *(Não-degenerância)*

   Para qualquer $(\Delta X, \Delta\lambda, \Delta S)$ *satisfazendo*

   $$\sum_{i=1}^{m} \Delta\lambda_i A_i + \Delta S = 0 \quad e \quad A_i \bullet \Delta X = 0 \ (i = 1, \ldots, m),$$

   vale a seguinte implicação:

   $$X^* \ \Delta S + \Delta X \ S^* = 0 \Rightarrow (\Delta X, \Delta S) = (0, 0).$$

Em se tratando da Hipótese (5.2.2), vemos que o item (1) é clássico e o item (2) foi apresentado por Kojima et alii, veja [**?**]. Em relação à essa referência, Haeberly mostrou que tal Hipótese é equivalente à condição de não-generância primal e dual abordada por Alizadeh et alii, veja [**?**].

Teorema 5.2.1. *Suponha que as Hipóteses* (4.0.1) *e* (5.2.2) *valem na solução* (X^*, λ^*, S^*) *de* (**??**). *Assim, a aplicação linear* $\nabla\Theta(X^*, \lambda^*, S^*, 0)$ *é bijetiva.*

Antes da demonstração, é bom ressaltarmos que tal Teorema mostra que a Hipótese (5.2.1) vale sobre as duas citadas desde que as iteradas (X^k, λ^k, S^k) geradas pelo Algoritmo convirjam a uma solução (X^*, λ^*, S^*) satisfazendo estas condições. O fato de as iteradas convergirem para este único ponto não é restritivo, visto que juntas, as duas Hipóteses (4.0.1) e (5.2.2) implicam que (X^*, λ^*, S^*) é a solução única das condições de otimalidade dadas em (**??**). A seguir, a demonstração do Teorema (5.2.1).

Demonstração. Consideremos ϕ definida por (3.1). O caso para (3.2) é análogo. Defina

$$E := ((X^*)^2 + (S^*)^2)^{1/2}$$

Pela complementariedade estrita, temos que $E \succ 0$. Portanto, pelo Teorema (3.0.3) temos que Θ é continuamente diferenciável em $(X^*, \lambda^*, S^*, 0)$. Para ver que $\nabla\Theta(X^*, \lambda^*, S^*)$ é bijetiva, temos que provar apenas que ela é injetiva. Então, considere

$$\nabla\Theta(X^*, \lambda^*, S^*, 0)\begin{pmatrix} \Delta X \\ \Delta\lambda \\ \Delta S \\ \Delta\tau \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Temos que mostrar que $(\Delta X, \Delta\lambda, \Delta S, \Delta\tau) = (0, 0, 0, 0)$ é a única solução. Note que a última linha dá imediatamente

$$\Delta\tau = 0 \tag{5.6}$$

Tomando (5.6) e usando também o Teorema (3.0.3), podemos reescrever

$$\begin{aligned} \sum_{i=1}^{m} \Delta\lambda_i A_i + \Delta S &= 0, && (5.7) \\ A_i \bullet \Delta X &= 0 \quad (i = 1, \ldots, m) && (5.8) \\ \Delta X + \Delta S - L_E^{-1}[X^* \, \Delta X + \Delta X \; X^* + S^* \, \Delta S + \Delta S \; S^*] &= 0. && (5.9) \end{aligned}$$

De (5.9) temos que

$$L_{E-X^*}[\Delta X] + L_{E-S^*}[\Delta S] = 0, \tag{5.10}$$

pela Proposição (4.0.5). Agora, utilizando-se do fato de que (X^*, λ^*, S^*) é uma solução estritamente complementar de (**??**), em particular, temos que

$$X^* S^* = 0$$

ou seja, X^* e S^* comutam. Dessa informação, segue que, essas duas matrizes podem ser simultaneamente diagonalizáveis por uma transformação ortogonal. Isso significa que podemos encontrar uma única matriz ortogonal $Q \in \mathbb{R}^{n\times n}$ e matrizes diagonais $D_X, D_S \in \mathbb{R}^{n\times n}$ de tal maneira que

$$X^* = Q^T D_X Q \quad \text{e} \quad S^* = Q^T D_S Q$$

Observe ainda que

$$(X^* + S^*)^2 =^(X^*)^2 + X^*S^* + S^*X^* + (S^*)^2 = (X^*)^2 + (S^*)^2,$$

e daí então temos que,

$$E = ((X^*)^2 + (S^*)^2)^{1/2} = ((X^* + S^*)^2)^{1/2} = X^* + S^*$$

e então $E - X^* = S^*$ e $E - S^* = X^*$, daí se $L_{S^*}[\Delta X] + L_{X^*}[\Delta S] = 0$, (5.10) fica assim reescrita

$$S^*\ \Delta X + \Delta X\ S^* + X^*\Delta S + \Delta S\ X^* = 0.$$

Para concluir que, a partir da passagem anterior, chegamos em $(\Delta X, \Delta S) = (0, 0)$, usaremos a seguinte notação, para simplificar notações

$$J := \Delta X$$

$$M := \Delta S$$

Com $S^*, X^* \succeq 0$ e $S^*X^* = X^*S^*$, podemos tomar uma matriz ortogonal Q que diagonaliza S^* e X^* simultaneamente, isto é,

$$Q^T S^* Q = D_S \ \text{ e } \ Q^T X^* Q = D_X$$

onde D_S e D_X são matrizes diagonais $n \times n$. Como Q é ortogonal, tomando $D_S = Q^T S^* Q$ e $D_X = Q^T X^* Q$, temos então que:

$$S^* = Q^T D_S Q \ \text{ e } \ X^* = Q^T D_X Q$$

E além disso, vem

$$\begin{aligned} D_S \bullet D_X &= tr(D_S D_X) \\ &= tr((Q^T S^* Q)(Q^T X^* Q)) \\ &= tr(S^* X^*) \\ &= 0, \quad \text{pela Proposição (2.0.1)}, \end{aligned}$$

e como $S^* \bullet X^* = tr(S^*X^*)$ temos que

$$S^* \bullet X^* = D_S \bullet D_X = 0.$$

Temos também que $D_S + D_X = Q^T(S^* + X^*)Q \succ 0$. Então, sem perda de generalidade, podemos considerar que as matrizes diagonais D_S e D_X possuem as seguintes formas:

$$D_S = \begin{pmatrix} D_{X_{11}} & 0 \\ 0 & 0 \end{pmatrix} \ \text{ e } \ D_X = \begin{pmatrix} 0 & 0 \\ 0 & D_{X_{22}} \end{pmatrix},$$

respectivamente. Aqui $D_{S_{11}}$ e $D_{X_{22}}$ são matrizes diagonais positivas de ordem $m \times m$ e $(n-m) \times (n-m)$ respectivamente e $0 \leq m \leq n$. Seja

$$M' = Q^T M Q = \begin{pmatrix} M'_{11} & M'_{12} \\ (M'_{12})^T & M'_{22} \end{pmatrix} \ \text{ e } \ J' = Q^T J Q = \begin{pmatrix} J'_{11} & J'_{12} \\ (J'_{12})^T & J'_{22} \end{pmatrix}.$$

Então, temos por hipótese, que

$$
\begin{aligned}
D_{S_{11}} J'_{11} + J'_{11} D_{S_{11}} &= 0, \\
D_{X_{22}} M'_{22} + M'_{22} D_{X_{22}} &= 0 \text{ e} \\
D_{S_{11}} J'_{12} + M'_{12} D_{X_{22}} &= 0.
\end{aligned}
\tag{5.11}
$$

Definição 5.2.1. *O produto de Kronecker de duas matrizes* $A = [a_{ij}] \in \mathbb{R}^{n\times n}$ *e* $B = [b_{ij}] \in \mathbb{R}^{n\times n}$ *é denotado por* $A \otimes B$ *e é definido como o bloco de matriz*

$$
A \otimes B \equiv \begin{bmatrix} a_{11}\ B & \dots & a_{1n}\ B \\ \vdots & \ddots & \vdots \\ a_{n1}\ B & \dots & a_{nn}\ B \end{bmatrix} \in \mathbb{R}^{n\times n}.
$$

Usando o produto Kronecker de matrizes, podemos reescrever a primeira e a segunda igualdades anteriores da seguinte forma:

$$
\begin{aligned}
(I \otimes D_{S_{11}} + D_{S_{11}} \otimes I)(vec\ J'_{11}) &= 0 \\
(I \otimes D_{X_{22}} + D_{X_{22}} \otimes I)(vec\ M'_{22}) &= 0
\end{aligned}
$$

Como $D_{S_{11}}, D_{X_{22}} \succ 0$, portanto são também as matrizes

$$
(I \otimes D_{S_{11}} + D_{S_{11}} \otimes I) \text{ e}
$$

$$
(I \otimes D_{X_{22}} + D_{X_{22}} \otimes I)
$$

Com isso, temos então que

$$
J'_{11} = 0 \text{ e } M'_{22} = 0 \tag{5.12}
$$

Segue então de (5.11) e de (5.12) que

$$
\begin{aligned}
S^* J + M X^* &= Q(D_S J' + M' D_X) Q^T \\
&= Q \begin{pmatrix} D_{S_{11}} J'_{11} & D_{S_{11}} J'_{12} + M'_{12} D_{X_{22}} \\ 0 & M'_{22} D_{X_{22}} \end{pmatrix} Q^T \\
&= 0.
\end{aligned}
$$

Portanto, obtemos pela Hipótese (5.2.2), item (2) que $(M, J) = (0, 0)$, isto é $(\Delta X, \Delta S) = (0, 0)$.

Como, pela Hipótese (4.0.1), as matrizes A_i são linearmente independentes, concluimos então, de (5.7), que $\Delta\lambda = 0$, o que encerra a demonstração. □

Já afirmamos anteriormente que o Teorema (5.2.1) dá apenas uma condição suficiente para a Hipótese (5.2.1) ser satisfeita. Como a hipótese usada nesse Teorema implica que a solução das condições de otimalidade (**??**) é única, tal Teorema se torna um tanto restritivo. Porém, alguns trabalhos recentes nas áreas de Programação Linear e Problemas de Complementariedade indicam que tal Hipótese pode ser enfraquecida e não implica necessariamente a solubilidade única das condições de otimalidade dadas em (**??**).

Analisaremos agora o comportamento local do Algoritmo iniciando pelo Lema seguinte.

Lema 5.2.1. *Suponha que a Hipótese* (5.2.1) *vale. Então temos que*

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k + \Delta\tau_k)\|_F = o(\tau_k).$$

Demonstração. Para simplificar notações, seja:

$$\begin{pmatrix} \Delta X^k \\ \Delta S^k \\ \Delta\tau_k \end{pmatrix} := (\Delta K)$$

Temos que, de (4.3)

$$\nabla\phi(X^k, S^k, \tau_k)(\Delta K) = -\phi(X^k, S^k, \tau_k),$$

obtemos do Teorema do Valor Médio

$$\begin{aligned}
& \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k + \Delta\tau_k)\|_F \\
= & \left\| \int_0^1 \nabla\phi(X^k + \eta\Delta X^k, S^k + \eta\Delta S^k, \tau_k + \eta\Delta\tau_k)(\Delta K)\ d\eta + \phi(X^k, S^k, \tau_k) \right\|_F \\
= & \left\| \int_0^1 \nabla\phi(X^k + \eta\Delta X^k, S^k + \eta\Delta S^k, \tau_k + \eta\Delta\tau_k)(\Delta K)\ d\eta - \nabla\phi(X^k, S^k, \tau_k)(\Delta K) \right\|_F \\
\leq & \int_0^1 \left\| [\nabla\phi(X^k + \eta\Delta X^k, S^k + \eta\Delta S^k, \tau_k + \eta\Delta\tau_k) - \nabla\phi(X^k, S^k, \tau_k)](\Delta K) \right\|_F\ d\eta \\
= & o\left(\||(\Delta K)\|| \right).
\end{aligned}$$

Essa igualdade anterior provém do fato da aplicação ϕ ser continuamente diferenciável. Levando em consideração a Hipótese (5.2.1), obtemos finalmente

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k + \Delta\tau_k)\|_F = o(\tau_k).$$

□

O próximo resultado concentra a etapa principal para se obter a convergência superlinear da sequência τ_k.

Lema 5.2.2. *Suponha que valem a Hipótese* (5.2.1) *e a desigualdade*

$$\beta > \kappa\sqrt{n},$$

onde κ *denota a constante do Lema* (3.0.3). *Assim, a sequência* η_k *converge para zero.*

Demonstração. Seja $\epsilon > 0$ dado arbitrário. Por (4.3) temos que

$$\Delta\tau_k = -\tau_k,$$

e obtemos do Lema (5.2.1)

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\| = \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k + \Delta\tau_k)\|_F = o(\tau_k).$$

Dessa forma, existe um índice $K_\epsilon \in \mathbb{N}$ tal que

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F \leq \epsilon\tau_k, \quad \forall\ k \geq K_\epsilon.$$

Assim, temos então, para $\forall\ \eta > 0$ e também para $\forall\ k \geq K_\epsilon$ que

$$\begin{aligned}
\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \eta\tau_k \ &\leq\ \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F + \\
&+\ \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \eta\tau_k) - \\
&-\ \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F \\
&\leq\ \epsilon\tau_k + \kappa\sqrt{n}\ \eta\tau_k,
\end{aligned}$$

e observe também que a última desigualdade anterior segue do Corolário (3.0.1). É verdade que

$$\epsilon\tau_k + \kappa\sqrt{n}\ \eta\tau_k \leq \beta\eta\tau_k$$

vale para todo $\eta \geq \dfrac{\epsilon}{\beta - \kappa\sqrt{n}}$, vemos que a definição de η_k mostra que $\eta_k\alpha_1$ contradiz tal desigualdade, ou seja,

$$\eta_k < \frac{\epsilon}{(\beta - \kappa\sqrt{n})\alpha_1}$$

Como $\beta - \kappa\sqrt{n} > 0$ por hipótese e $\epsilon > 0$ por escolha arbitrária, isso implica que

$$\eta_k \to 0,$$

concluindo a prova. $\square$

Para concluir então enunciemos o Teorema que comprova a convergência superlinear local do Algoritmo:

Teorema 5.2.2. *Sobre a Hipótese* (5.2.1) *temos que*

$$\tau_{k+1} = o(\tau_k),$$

ou seja, o parâmetro de suavilização converge superlinearmente localmente para zero.

Demonstração. De acordo com o Lema (5.2.2) e pela definição de τ_{k+1} e $\hat{\tau}_k$ no Algoritmo, obtemos

$$\tau_{k+1} = (1 - \hat{\sigma}\hat{\eta}_k)\hat{\eta}_k \leq \hat{\tau}_k = \eta_k \tau_k = o(\tau_k)$$

ou seja, $\tau_k \to 0$ superlinearmente, como queríamos demonstrar. □

Finalizando, afirmamos que em [**?**] observou-se que o Teorema (5.2.2) também vale se em (4.3) trocarmos o lado direito da igualdade por $-\Theta(W^k, 0)$. Tal questionamento não se adequa ao que fizemos aqui, pois nas nossas discussões, em ambos os Lemas (5.2.1) e (5.2.2), a substituição feita do lado direito da passagem (4.3) era por $-\Theta(W^k, \tau_k)$.

Conclusão

Neste trabalho, apresentamos uma reformulação da trajetória central para o problema de Programação Semidefinida, através das funções de Fischer-Burmeister e função mínimo suavilizadas. Nesta nova reformulação a trajetória central pertubada é apresentada através de um sistema de equações não-lineares $\Phi_\tau(X, \lambda, S) = 0$, onde a função $\Phi_\tau(X, \lambda, S)$ é continuamente diferenciável, facilitando assim a utilização do método de Newton no algoritmo tipo suavilização apresentado.

Para o algoritmo apresentado, provamos convergência global e superlinear local sob hipóstese convenientes. É importante salientar que complementariede estrita é uma condição suficiente para que a hipótese (5.2.1), utilizada nas demonstrações de convergência, seja satisfeita, assim, como perspectivas futuras de estudo seria interessante apresentar exemplos onde vale a hipótese (5.2.1) e não vale complementariedade estrita.

Referências Bibliográficas

[1] Alizadeh, F., *Interior point methocs in semidefinite programming with applications to combinatorial optimization.* SIAM Journal on Optimization 5, pp.13-51, (1995).

[2] Alizadeh, F., Haeberly, J.-P and Overton, M.L., *Primal-dual inteior-point methods for semidefinite programming: Convergence rates, stability and numerical results.* SIAM Journal on Optimization 8, pp.746-768, (1998)

[3] Chen, X. and Tseng, P., *Non-interior continuation methods for solving semidefinite complementarity problems.* Technical Report, Department of Mathematics, University of Washington, Seattle, (1999).

[4] Graybill, F. A., *Matrices with Applications in Statistics.* Colorado State University, Belmont, California, (1983)

[5] Horn, R.A. and Johnson, C.R., *Matrix Analysis.* Cambridge University Press, (1985).

[6] Horn, R.A. and Johson, C.R., *Topics in Matrix Analysis.* Cambridge University Press, (1991).

[7] Kanzow, Christian and Nagel, Christian., *Semidefinite programs: new search directions, smoothing-type methods, and numerical results.* Department of Mathematics.

University of Hamburg. Center of Optimization and Approximation. Hamburg, Germany, (2001).

[8] Klerk, E. de., *Aspects of semidefinite programming.* Delft University of Technology, Kluwer Academic Publishers, The Netherlands, (2002).

[9] Kojima, M., Shida, M. and Shindoh, S., *Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs.* Mathematical Programming 80, pp.129-160, (1998).

[10] Kojima, M., Shida, M. and Shindoh, S., *A predictor-corrector interior-point algorithm for the semidefinite linear complementarity problem using the Alizadeh-Haeberly-Overton search direction.* SIAM Journal on Optimization 9, pp.444-465, (1999)

[11] Wright, S. J., *Primal-dual interior-point methods.* SIAM, Philadelphia, (1997).

[12] Ye, Y., *Interior Point Alghoritms.* Wiley Interscience Series in Discrete Mathematics ans Optimization. Toronto, Canada, (1997).

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)