FUNÇÕES DE MÉRITO PARA O PROBLEMA DE COMPLEMENTARIEDADE SEMIDEFINIDO

Rosely Maria Barbosa Goes

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Aprovada por:

Prof. Paulo Roberto Oliveira, Dr. Ing.

Prof. Luiz Satoru Ochi, D. Sc.

Prof. Márcia Helena Costa Fampa, D. Sc.

Prof. Maria Helena Cautiero Horta Jardim, D. Sc.

Prof. Orizon Pereira Ferreira, D. Sc.

Prof. Susana Scheimberg de Makler, D. Sc.

RIO DE JANEIRO, RJ - BRASIL

OUTUBRO DE 2001

GOES, ROSELY MARIA BARBOSA

Funções de Mérito para o Problema de Complementariedade Semidefinido [Rio de Janeiro] 2001

IX, 94 p. 29,7 cm (COPPE/UFRJ, D.Sc., Engenharia de Sistemas e Computacão, 2001)

Tese – Universidade Federal do Rio de Janeiro, COPPE

1 - Problema de Complementariedade Semidefinido

I. COPPE/UFRJ II. Título (série)

- Aos meus pais, ao meu esposo Geci e ao meu filho Bruno.

- À memória de minha tia Nanci.

Agradecimentos

- Ao meu orientador Paulo Roberto Oliveira pela sua orientação, incentivo e crédito em mim depositado.

- Em especial ao colega Geci José Pereira da Silva pela sua paciência em ouvir-me e pelas sugestões dadas neste trabalho, que foi de fundamental importância.

- Aos professores do Programa de Engenharia de Sistemas e Computação da COPPE/UFRJ que também contribuíram diretamente ou indiretamente para que este trabalho se realizasse. Em especial a professora Regina Burachik que enriqueceu-me não só em conteúdos mas também me fez ver, indiretamente, o quanto é importante darmos chances para que possamos chegar a um objetivo.

- A banca examinadora: Satoru, Márcia Fampa, Maria Helena, Orizon e Susana pelas sugestões dadas que contribuíram para melhoria dela. Em especial ao Orizon que apoiou a minha vinda para a UFRJ.

- Aos colegas do Departamento de Matemática e Estatística da UFG, que assumiram minha carga horária durante o tempo que estive realizando o doutorado. Em especial ao nosso Diretor Ronaldo Garcia.

- A CAPES, pelo apoio financeiro.

- Não posso deixar de mencionar que, teria sido impossível chegar ao fim da luta sem o constante suporte emocional e afetivo do meu companheiro, Geci, para vencer os inúmeros momentos de incertezas. E o carinho do meu filho Bruno.

- Aos meus pais: Edivaldo e Elia e aos meus irmãos: Reginaldo e Rogério que sempre me apoiaram, dando seu afeto e carinho.

- Aos familiares do meu esposo: Minha sogra Maria e meu sogro Gerson e aos seus irmãos, em especial a Sueli e ao Luiz.

- Aos meus parentes, em especial a Família Everado e Evanda e seus filhos que são meus primos-irmãos, a Tia Elisdete, minha tia-irmã, e ao meu Tio Eduardo.

- A todos os amigos que direta ou indiretamente me ajudaram em mais esta conquista.

  - A Viviane, Marina, minhas irmãs de coração, que vieram me prestigiar de Natal- RN e Campinas-SP, respectivamente

  - Aos amigos da UFRJ, em especial Vera, Gilvan e Rodrigo, que sempre estiveram presentes.

  - Aos amigos do Impa, em especial ao Ari, nosso irmão de coração, e ao Rolando, amigo eterno.

  - A Alexandrina por cuidar do meu filho Bruno de forma tão carinhosa e dedicada nos momentos em que estivemos ausentes.

  - Finalmente, ao pessoal do Prédio da Rua Moura Brasil-47, onde moramos durante o doutorado.

Índice

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

FUNÇÕES DE MÉRITO PARA O PROBLEMA DE COMPLEMENTARIEDADE SEMIDEFINIDO

Rosely Maria Barbosa Goes

Outubro/ 2001

Orientador: Paulo Roberto Oliveira

Programa: Engenharia de Sistemas e Computação

Recentemente, Tseng estendeu uma classe de funções de mérito do problema de complementariedade não-linear (PCN) para o problema de complementariedade semidefinido $(PCSD)$ e mostrou algumas propriedades sob hipóteses adequadas. Yamashita e Fukushima também apresentaram outras propriedades. Neste trabalho, enfraquecemos algumas destas condições. Particularmente, damos uma certa condição de regularidade que é necessária e suficiente para um ponto estacionário ser solução do $PCSD$. Além disso, propomos uma nova classe de funções de mérito para o $PCSD$ e provamos algumas propriedades. Em particular, apresentamos condições sob as quais estas funções de mérito fornecem cota de erro global para o $PCSD$ e têm conjuntos de nível limitados.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

MERIT FUNCTIONS FOR THE SEMIDEFINITE COMPLEMENTARITY PROBLEM

Rosely Maria Barbosa Goes

October/2001

Advisor: Paulo Roberto Oliveira

Department: Systems and Computation Engineering

Recently, Tseng extended a class of merit functions for the nonlinear complementarity problem (NCP) to the semidefinite complementarity problem ($SDCP$) and showed some properties under suitable assumptions. Yamashita e Fukushima also presented other properties. In this work, we weaken some of those conditions. Particularly, we give a certain regularity condition that is necessary and sufficient for a stationary point to be a solution of the $SDCP$. Moreover, we propose a new class of merit functions for the $SDCP$ and prove some properties. In particular, we present conditions under which those merit functions provide a global error bound for the $SDCP$ and conditions under which they have bounded level sets.

Introdução

O problema de complementariedade não-linear (PCN) [18, 22] consiste em encontrar um $x \in \mathbb{R}^n$ tal que

$$[PCN] \qquad\qquad x \geq 0, \ F(x) \geq 0, \ x^T F(x) = 0,$$

onde F é uma aplicação de $\mathbb{R}^n$ em $\mathbb{R}^n$. No caso onde F é afim, o PCN reduz-se ao problema de complementariedade linear (PCL) [8]. Este problema tem várias aplicações, tais como programação matemática, economia, engenharia e mecânica [8, 18]. Por se tratar de uma classe de problemas práticos, várias extensões do PCL foram propostas, e a análise e métodos destes problemas incluem o PCN muito estudado nos últimos anos [9, 10, 11, 13, 14, 24, 25, 27, 29, 39]. Como uma extensão do PCN, temos o problema de complementariedade semidefinido ($PCSD$) [7, 15, 16, 30, 32, 38] definido como segue. Seja $\mathcal{S}$ o conjunto das matrizes simétricas reais $n \times n$, o $PCSD$ consiste em encontrar uma matriz $x \in \mathcal{S}$ tal que

$$[PCSD] \qquad\qquad F(x) \in \mathcal{K}, \ G(x) \in \mathcal{K}, \ \langle F(x), G(x) \rangle = 0,$$

onde $\mathcal{K} \subset \mathcal{S}$ denota o cone convexo fechado contendo os elementos de $\mathcal{S}$ que são semidefinidos positivos, F e G são aplicações de $\mathcal{S}$ em $\mathcal{S}$ e $\langle \cdot, \cdot \rangle$ é o produto interno definido por

$$\langle x, y \rangle := \mathrm{tr}[xy],$$

onde $x, y \in \mathcal{S}$ e $\mathrm{tr}[\cdot]$ denota o traço da matriz. Quando $\mathcal{S}$ é restrito ao conjunto das matrizes diagonais, este problema reduz-se ao PCN. O $PCSD$ está relacionado com as condições de otimalidade do problema de programação semidefinida (PSD) [1, 26] a qual tem recentemente atraído interesse na teoria do controle e otimização combinatória [34, 35].

Como o $PCSD$ é considerado uma extensão do PCN, é possível que o $PCSD$ possa ser resolvido aplicando métodos que foram desenvolvidos para o PCN. Recentemente, várias reformulações do $PCSD$ como um problema de minimização foram propostas [30, 32, 38]. Uma função que constitui um problema de minimização equivalente para o $PCSD$ é chamada de função de mérito. Um trabalho pioneiro nesta linha foi realizado por Tseng [32]; ele mostrou que algumas funções de mérito do PCN podem ser estendidas para o $PCSD$. Em particular, a seguinte classe de funções $f : \mathcal{S} \to I\!R$ foi estendida:

$$f(x) = \psi_0(\langle x, F(x)\rangle) + \psi(x, F(x)),$$

onde $\psi_0 : I\!R \longrightarrow [0, \infty)$ é tal que

$$\psi_0(t) = 0 \ \text{ se, e somente se, } \ t \leq 0,$$

e $\psi : \mathcal{S} \times \mathcal{S} \to [0, \infty)$ é tal que

$$\psi(a, b) = 0, \quad \langle a, b\rangle \leq 0 \Longleftrightarrow (a, b) \in \mathcal{K} \times \mathcal{K}, \ \langle a, b\rangle = 0.$$

Esta função, estudada por Luo e Tseng no contexto de PCN, é uma função de mérito sobre $\mathcal{S}$. Tseng [32] mostrou uma condição para que qualquer ponto estacionário de f seja solução do $PCSD$. No entanto, não respondeu em que condições f tem conjuntos de níveis limitados e fornece cota de erro global. Yamashita e Fukushima

[38] mostraram que os conjuntos de nível de f são limitados quando F é monótona e o $PCSD$ é estritamente viável, ou seja, existe uma matriz definida positiva $\widehat{x} \in \mathcal{S}$ tal que $F(\widehat{x})$ é definida positiva; e que f fornece cota de erro global desde que F seja fortemente monótona. Neste trabalho, descrevemos estes resultados, com $G = I$, que são válidos desde que ψ_0 e ψ satisfaçam certas condicões. Além disso, sob condições mais fracas do que as [32, 38], obtemos que os conjuntos de nível de f são limitados e qualquer ponto estacionário é solução do $PCSD$. Apresentamos dois métodos de descidas, um deles com a direção dada por Tseng, para resolver o $PCSD$ e mostramos suas convergências.

Para o PCN, Kanzow, Yamashita e Fukushima [25] propuseram uma função de mérito que fornece uma cota de erro global sob a condição de F ser $P-$uniforme. Naturalmente, esperamos que exista uma função de mérito para o $PCSD$ que tenha esta propriedade. Com este objetivo e motivados pelos resultados estendidos existentes para o $PCSD$, propomos a seguinte classe de funções para o $PCSD$:

$$g(x) = \phi(x, F(x)),$$

com $\phi : \mathcal{S} \times \mathcal{S} \longrightarrow I\!R$ definida por

$$\phi(a, b) = \psi_{0^*}(ab + ba) + \psi(a, b),$$

onde $\psi_{0^*} : \mathcal{S} \longrightarrow [0, \infty)$ é contínua tal que

$$\psi_{0^*}(t) = 0 \text{ se, e somente se, } t \preceq 0 \quad (-t \text{ é semidefinida positiva}),$$

e $\psi : \mathcal{S} \times \mathcal{S} \longrightarrow [0, \infty)$ é contínua tal que

$$\psi(a, b) = 0, \langle a, b \rangle \leq 0 \quad \text{se, e somente se,} \quad (a, b) \in \mathcal{K} \times \mathcal{K}, \langle a, b \rangle = 0.$$

Esta função, estudada por Kanzow, Yamahista e Fukushima [25] no contexto do PCN, é uma função de mérito sobre S. Mostramos que g fornece cota de erro global e tem conjuntos de nível limitados desde que F seja P-uniforme. Além disso, apresentamos outra condição para que g tenha conjuntos de nível limitados.

Este trabalho será organizado como segue: No Capítulo 1 fazemos um breve resumo dos conceitos de Teoria das Matrizes, Projeções e monotonia de aplicações necessários ao desenvolvimento dos capítulos posteriores. No Capítulo 2 apresentamos alguns resultados obtidos por Tseng [32], Shibata,Yamashita e Fukushima [30, 38] para o $PCSD$. Nos dois capítulos seguintes começa a nossa contribuição. No Capítulo 3 apresentamos as novas propriedades obtidas para a classe de funções Luo e Tseng. No Capítulo 4 apresentamos a nova classe de funções de mérito para o $PCSD$ e algumas propriedades. Finalmente, concluímos e damos sugestões de estudos posteriores.

Capítulo 1

Preliminares

1.1 Introdução

Neste capítulo apresentamos resultados sobre a Teoria das Matrizes, Projeções e Monotonia de aplicações, necessários ao desenvolvimento dos capítulos posteriores. O que faremos aqui, será reescrever o que já existe na literatura. As demonstrações serão, em princípio, todas elas referenciadas.

1.2 Teoria de Matrizes

Nesta seção apresentamos alguns resultados sobre matrizes simétricas, definidas positivas (dp), semidefinidas positivas (sdp) e traço. Estes resultados são clássicos, para mais detalhes, veja [17] e [20].

Seja $\mathcal{X}$ o conjunto das matrizes $n \times n$. Para $a \in \mathcal{X}$, denotamos por a_{ij} a $(i, j)-$ésima entrada de a, a^T a sua matriz transposta e, se a é inversível, por a^{-1} a sua inversa. Além disso, a é ortogonal se, e somente, se $a^T a = I = aa^T$, onde I é matriz identidade. Seja $\mathcal{O}$ o conjunto das matrizes ortogonais.

No decorrer desta tese trabalhamos com matrizes simétricas reais $n \times n$. Denotamos por $\mathcal{S}$ o conjunto das matrizes simétricas $n \times n$. A seguir relatamos algumas propriedades de matrizes simétricas.

Lema 1.2.1 *Seja $a \in \mathcal{S}$.*

(a) Então existe $p \in \mathcal{O}$ tal que $p^T ap$ é uma matriz diagonal cujos elementos diagonais são os autovalores de a.

(b) Seja $b \in \mathcal{S}$ tal que $ab = ba$, então existe $u \in \mathcal{O}$, e matrizes diagonais d e $\tilde{d}$ tal que $a = udu^T$ e $b = u\tilde{d}u^T$.

Prova. Ver [17] e [20]. ■

O traço de uma matriz a, denotado por tr$[a]$, é definido por

$$\mathrm{tr}[a] = \sum_{i=1}^{n} a_{ii}.$$

Para quaisquer matrizes a, $b \in \mathcal{X}$, e qualquer $p \in \mathcal{O}$ temos que $\mathrm{tr}[a]=\mathrm{tr}[pap^T]$, $\mathrm{tr}[a+b]=\mathrm{tr}[a]+\mathrm{tr}[b]$ e $\mathrm{tr}[ab]=\mathrm{tr}[ba]$.

O próximo resultado é uma relação entre os sinais dos elementos da diagonal de uma matriz pap^T, para algum $p \in \mathcal{O}$, e o traço de a.

Lema 1.2.2 *Seja $a \in \mathcal{X}$.*

(a) Então existe $p \in \mathcal{O}$ tal que os elementos não-nulos da diagonal de pap^T têm os mesmos sinais.

(b) Se tr$[a] \neq 0$, então existe $p \in \mathcal{O}$ tal que todos os elementos da diagonal de pap^T têm os mesmos sinais que o tr$[a]$.

Prova. Ver [6].

O seguinte resultado é uma relação entre o traço e o posto de uma matriz, denotado por $\mathrm{P}(\cdot)$.

Lema 1.2.3 *Seja a uma matriz $n \times n$ não-nula. Então*

$$[tr[a]]^2 \leq P(a) tr[a^2].$$

Prova. Ver [17]. ■

Seja $\|x\| := \sqrt{\langle x, x \rangle}$ a norma de Frobenius e $\langle x, y \rangle$ o produto interno definido por

$$\mathrm{tr}[xy] = \langle x, y \rangle.$$

Note que $(\mathcal{S}, \langle \cdot, \cdot \rangle, \|\cdot\|)$ forma um espaço de Hilbert. A matriz $x \in \mathcal{S}$ pode ser associada ao vetor $\widehat{x} := (\cdots, x_{ij}, \cdots)^T_{i \leq j} \in I\!R^v$, onde $v := n(n+1)/2$.

Considerando $\mathcal{K}$ o conjunto das matrizes simétricas sdp, temos que $\mathcal{K}$ é um cone convexo fechado em $\mathcal{S}$ [31] e seu interior é o conjunto das matrizes dp. Além disso apresenta a seguinte propriedade.

Lema 1.2.4 *$\mathcal{K}$ é auto-dual, ou seja, $\{y \in \mathcal{S} : \langle x, y \rangle \geq 0 \;\; \forall \; x \in \; \mathcal{K}\} = \mathcal{K}$.*

Prova. Ver [4] e [21]. ■

Apesar de $\mathcal{K}$ ter uma estrutura geométrica complicada comparada ao ortante não-negativo em $I\!R^v$, no entanto, pelo Lema 1.2.4, temos a seguinte propriedade $\mathcal{K}^0 = -\mathcal{K}$, onde $\mathcal{K}^0$ é o polar de $\mathcal{K}$:

$$\mathcal{K}^0 = \{y \in \mathcal{S} : \langle x, y \rangle \leq 0 \; \forall x \in \mathcal{K}\}.$$

Então, para qualquer $x \in \mathcal{S}$,

$$x \in \mathcal{K} \Longleftrightarrow \langle x, y \rangle \geq 0 \;\; \forall \;\; y \in \mathcal{K}. \tag{1.1}$$

Como $\mathcal{K}$ é um cone convexo fechado e $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$, então $\mathcal{K}$ induz uma ordem parcial "$\succeq_{\mathcal{K}}$" em $\mathcal{S}$: $x \succeq y \Longleftrightarrow x - y \in \mathcal{K}$. Deste modo, usaremos o símbolo

$$x \succeq (\succ) 0$$

para dizer que x é simétrica e sdp (dp); o símbolo $x \preceq 0$ significa que $-x \succeq 0$. Listaremos abaixo algumas propriedades bem conhecidas da teoria das matrizes que serão úteis neste trabalho.

Lema 1.2.5

(a) $x \succeq 0 \Longleftrightarrow pxp^T \succeq 0$ *para qualquer* $p \in \mathcal{O}$.

(b) Seja $x \succeq 0$ $(x \succ 0)$. *Então* $tr[x] \geq 0$ $(tr[x] > 0)$ *e* $tr[x] = 0 \Longleftrightarrow x = 0$.

(c) Sejam $x \succeq 0, y \succeq 0, \langle x, y \rangle = 0$, *então* $xy = yx = 0$.

Prova. Ver [17]. ■

Observe que diretamente dos Lemas 1.2.1 (b) e 1.2.5 (c) obtemos o seguinte corolário.

Corolário 1.2.1 *Sejam* $x, y \in \mathcal{K}$ *e* $\langle x, y \rangle = 0$, *então existe* $u \in \mathcal{O}$, *e matrizes diagonais* d *e* $\tilde{d}$ *tais que* $x = udu^T$ *e* $y = u\tilde{d}u^T$.

A seguir apresentamos uma relação entre os autovalores das matrizes a e $a + b$, para $a \in \mathcal{S}$ e $b \in \mathcal{K}$.

Lema 1.2.6 *Sejam* $a, b \in \mathcal{S}$. *Assuma que* $b \in \mathcal{K}$, *então*

$$\lambda_k(a) \leq \lambda_k(a + b), \quad \forall \ \ k = 1, \ldots, n,$$

onde $\lambda_1(\cdot) \leq \cdots \leq \lambda_n(\cdot)$ *são os autovalores.*

Prova. Ver [20]. ■

1.3 Projeção em $\mathcal{K}$

Nesta seção, apresentamos alguns resultados sobre projeções em $\mathcal{K}$ e suas propriedades. Para maiores detalhes ver [22].

A projeção ortogonal em $\mathcal{K}$, denotada por $[x]_+$, é definida por

$$[x]_+ = \arg\min_{y\in\mathcal{K}} \|y - x\|,$$

para qualquer $x \in \mathcal{S}$.

Como $(\mathcal{S}, \langle\cdot,\cdot\rangle, \|\cdot\|)$ é um espaço de Hilbert e $\mathcal{K}$ é um conjunto convexo fechado de $\mathcal{S}$, temos os seguintes resultados.

Lema 1.3.1 *Para qualquer $x \in \mathcal{S}$ existe um único elemento $[x]_+$ tal que*

$$\|x - [x]_+\| \leq \|x - y\|, \ \forall\, y \in \mathcal{K}.$$

Prova. Ver [22]. ■

Lema 1.3.2 *Para qualquer $x \in \mathcal{S}$, as seguintes afirmações são equivalentes:*

(a) $\|x - [x]_+\| \leq \|x - y\|, \ \forall\, y \in \mathcal{K}$,

(b) $\langle x - [x]_+, y - [x]_+\rangle \leq 0, \ \forall\, y \in \mathcal{K}$.

Prova. Ver [22]. ■

Lema 1.3.3 *Para qualquer $x \in \mathcal{S}$, temos*

$$x = [x]_+ + [x]_-, \quad [x]_+[x]_- = 0,$$

onde $[x]_-$ denota a projeção ortogonal em $\mathcal{K}^0$.

Prova. Ver [22]. ■

1.4 Monotonia de Aplicações

Relembramos algumas definições que serão úteis ao longo do texto.

Definição 1.4.1 *Seja $F : \mathcal{S} \longrightarrow \mathcal{S}$.*

(a) F é pseudomonótona se

$$\langle x - y, F(y) \rangle \geq 0 \Longrightarrow \langle x - y, F(x) \rangle \geq 0, \quad \forall x, y \quad \in \mathcal{S}.$$

(b) F é monótona se

$$\langle x - y, F(x) - F(y) \rangle \geq 0, \quad \forall \quad x, y \in \mathcal{S}.$$

(c) F é estritamente monótona se

$$\langle x - y, F(x) - F(y) \rangle > 0, \quad \forall x, y \quad \in \mathcal{S}, \quad x \neq y.$$

(d) F é fortemente monótona com módulo α, se existe $\alpha > 0$ tal que

$$\langle x - y, F(x) - F(y) \rangle \geq \alpha \left\| x - y \right\|^2, \quad \forall x, y \ \in \mathcal{S}.$$

(e) Se F é diferenciável, no sentido de Fréchet, sobre o conjunto aberto $\mathcal{D} \subset \mathcal{S}$, o Jacobiano de F em cada $x \in \mathcal{D}$, denotado por $\nabla F(x)$, é visto como uma aplicação $\nabla F(x) : \mathcal{S} \longrightarrow \mathcal{S}$.

(f) $\nabla F(x)$ é semidefinida positiva se

$$\langle y, \nabla F(x) y \rangle \geq 0 \quad \textit{para todo} \quad y \in \mathcal{S}.$$

(g) F é Lispchitz contínua em $\mathcal{S}$, com constante $L > 0$ se

$$\left\| F(x) - F(y) \right\| \leq L \left\| x - y \right\| \quad \forall \ x, y \in \mathcal{S}.$$

Note que se F é monótona e diferenciável, então $\nabla F(x)$ é sdp para todo $x \in \mathcal{S}$. É fácil verificar, a partir do Lema 1.3.2, que a projeção ortogonal $[x]_+$ é monótona e Lispchitz contínua com constante $L = 1$ (ou seja, não-expansiva).

Usando as definições, temos o seguinte resultado.

Proposição 1.4.1 *Uma aplicação diferenciável $F : \mathcal{S} \longrightarrow \mathcal{S}$ é dita ser*

(a) monótona se, e somente se, $\nabla F(x)$ é semidefinida positiva para todo x;

(b) estritamente monótona se $\nabla F(x)$ é definida positiva para todo x;

(c) fortemente monótona com módulo $\alpha > 0$ se, e somente se,

$$\langle x - y, \nabla F(x)(x - y) \rangle \geq \alpha \|x - y\|^2, \qquad \forall\ x, y \in \mathcal{S}.$$

Prova. Veja [28]. ■

A recíproca de (b) nem sempre vale. De fato, seja

$$F(x) = (x - 3I)^3 + I, \text{ para } x \in \mathcal{S}.$$

Note que F é estritamente monótona, mas $\nabla F(x)$ não é definida positiva em $x = 3I$.

Capítulo 2

Problema de Complementariedade Semidefinido

2.1 Introdução

Neste capítulo apresentaremos alguns resultados obtidos por Tseng [32], Shibata, Yamashita e Fukushima [30, 38] para o problema de complementariedade semidefinido $(PCSD)$. Por se tratar de um assunto recente e pouco conhecido, daremos as demonstrações dos resultados que serão usados nos capítulos posteriores. Para isso redefinimos o $PCSD$.

O $PCSD$ consiste em encontrar $x \in \mathcal{S}$ tal que

$$[PCSD] \qquad F(x) \in \mathcal{K},\ G(x) \in \mathcal{K},\ \ \langle F(x), G(x)\rangle = 0, \tag{2.1}$$

onde F e G são aplicações de $\mathcal{S}$ em $\mathcal{S}$ e $\langle \cdot, \cdot \rangle$ é o produto interno definido por $\langle x, y\rangle = \mathrm{tr}[xy]$, onde x, $y \in \mathcal{S}$.

O $PCSD$ é uma extensão do problema de complementariedade não-linear (PCN), onde o cone dos vetores reais não-negativos é substituído pelo cone das matrizes reais simétricas semidefinidas positivas. No entanto, nem sempre é possível estender os resultados, pois estamos trabalhando no espaço de matrizes.

As notações seguem a do capítulo 1. Além disso, para quaisquer $I, J \subset$

$\{1, \ldots, n\}$, denotamos por x_{IJ} a submatriz de x com linhas $i \notin I$ e colunas $j \notin J$ removidas da matriz x. Para quaisquer $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, denotamos por $\mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$ a matriz diagonal cujos elementos diagonais são $\lambda_1, \ldots, \lambda_n$.

2.2 Preliminares

Nesta seção, apresentamos algumas definições e resultados necessários para as seções subsequentes.

Definição 2.2.1 *(a) Para qualquer matriz $x \in \mathcal{S}$, $[x]_+$ e $[x]_-$ são as projeções ortogonais de x sobre $\mathcal{K}$ e $-\mathcal{K}$, respectivamente.*

(b) Para qualquer matriz $n \times n$, sym$[x]$ é definido por

$$sym[x] := x + x^T.$$

(c) Para qualquer subespaço $\mathcal{D} \subseteq \mathcal{S}$, uma aplicação $G : \mathcal{D} \longrightarrow \mathcal{D}$ é definida positiva em $\mathcal{D}$ se

$$\langle x, G[x] \rangle > 0 \ \forall \ x \in \mathcal{D} \quad (x \neq 0).$$

(d) Para qualquer $c \in \mathcal{K}$, seja

$$\mathcal{S}_c = \{x \in \mathcal{S} : pxp^T = \begin{bmatrix} \widetilde{x}_{II} & 0 \\ 0 & 0 \end{bmatrix} \text{ para alguma submatriz } \widetilde{x}_{II}\} \tag{2.2}$$

onde $p \in \mathcal{O} = \{x \in \mathcal{X} : x^T = x^{-1}\}$ e $I \subset \{1, \cdots, n\}$ são tais que

$$pcp^T = \begin{bmatrix} \widetilde{c}_{II} & 0 \\ 0 & 0 \end{bmatrix} \text{ para alguma submatriz } \widetilde{c}_{II} \ dp. \tag{2.3}$$

(e) Para qualquer $c \in \mathcal{K}$, a aplicação linear $L_c : \mathcal{S}_c \longrightarrow \mathcal{S}_c$ é definida por

$$L_c(x) := cx + xc.$$

Com estas definições, as seguintes propriedades valem:

Lema 2.2.1 *(a) Para qualquer matriz $x \in \mathcal{S}$, temos que*

$$x = [x]_+ + [x]_-, \quad [x]_+[x]_- = 0, \quad [x]_- = -[-x]_+, \quad e \quad x[x]_+ = [x]_+x = ([x]_+)^2 .$$

(b) Para quaisquer $a, b \in \mathcal{S}$, temos que $a, b \in \mathcal{S}_c$, onde $c := (a^2 + b^2)^{\frac{1}{2}}$.

(c) Para qualquer $c \in \mathcal{K}$, a aplicação linear L_c é definida positiva sobre $\mathcal{S}_c$, e portanto inversível. Especificamente, para qualquer $x \in \mathcal{S}_c$, $L_c^{-1}[x]$ é a única matriz $y \in \mathcal{S}_c$ tal que

$$cy + yc = x.$$

(d) Para $c \in \mathcal{K}$ e $x, y \in \mathcal{S}_c$ arbitrários, tem-se que

$$\langle y, L_c[x] \rangle = \langle L_c[y], x \rangle , \quad xL_c^{-1}[c] = L_c^{-1}[c]x = \frac{1}{2}x$$

$$e \qquad xL_c^{-1}[x] = 0 \Longrightarrow x = 0.$$

(e) Para $a \in \mathcal{K}$ e $b \in \mathcal{S}$ arbitrários, temos que

$$\langle a, b \rangle \leq \langle a, [b]_+ \rangle .$$

Prova.

(a) As identidades seguem dos Lemas 1.3.3 e 1.3.2 (b).

(b) Fixe a e $b \in \mathcal{S}$ arbitrários e tome $c := (a^2 + b^2)^{\frac{1}{2}}$. Como $c \in \mathcal{K}$, temos que a equação (2.3) vale para algum $p \in \mathcal{O}$ *e* $I \subset \{1, \cdots, n\}$ e alguma submatriz $\widetilde{c}_{II}$ dp. Então,

$$\begin{aligned} (pap^T)^2 + (pbp^T)^2 &= pap^Tpap^T + pbp^Tpbp^T \\ &= p(a^2 + b^2)p^T = \begin{bmatrix} \widetilde{c}_{II}^2 & 0 \\ 0 & 0 \end{bmatrix} . \end{aligned}$$

Como $(pap^T)^2$ e $(pbp^T)^2$ são sdp temos

$$pap^T = \begin{bmatrix} \widetilde{a}_{II} & 0 \\ 0 & 0 \end{bmatrix} \quad \text{e} \quad pbp^T = \begin{bmatrix} \widetilde{b}_{II} & 0 \\ 0 & 0 \end{bmatrix}, \tag{2.4}$$

para determinadas submatrizes $\widetilde{a}_{II}$ e $\widetilde{b}_{II}$. Por (2.2), $a, b \in \mathcal{S}_c$.

(c) Fixe $c \in \mathcal{K}$ arbitrário e $x \in \mathcal{S}_c$,

$$\langle x, L_c[x] \rangle = 2\text{tr}[x^T cx] = 2\text{tr}[px^T p^T pcp^T pxp^T].$$

Por (2.2), (2.3) e (1.1),

$$< x, L_c[x] >= 2\text{tr}[\widetilde{c}_{II}\widetilde{x}_{II}^2] > 0, \quad \widetilde{x}_{II} \neq 0, \widetilde{c}_{II} \ \text{ dp.}$$

Então, L_c é definida positiva, e portanto inversível.

(d) Fixe $c \in \mathcal{K}$ e $x, y \in \mathcal{S}_c$ arbitrários. As afirmações seguem da definição de L_c e do item (c).

(e) Temos que

$$\langle a, b \rangle = \langle a, [b]_+ \ + [b]_- \rangle = \langle a, [b]_+ \rangle + \langle a, [b]_- \rangle \leq \langle a, [b]_+ \rangle ,$$

onde a primeira igualdade segue da Definição 2.2.1 (a) e a desigualdade de (1.1) pois $a \in \mathcal{K}$ e $[b]_- \in -\mathcal{K}$. ■

A parte (a) do Lema abaixo apresenta uma maneira de calcular a projeção $[a]_+$ via decomposição espectral de a.

Lema 2.2.2 *(a) Para qualquer* $a \in \mathcal{S}$,

$$[a]_+ \ = p^T diag\{\max\{0, \lambda_1\}, \cdots, \max\{0, \lambda_n\}\}p,$$

onde $p \in \mathcal{O}$ *e* $\lambda_1, \cdots, \lambda_n \in I\!R$ *satisfazem* $a = p^T diag\{\lambda_1, \cdots, \lambda_n\}p$. *Além disso, para* $|a| = p^T diag\{|\lambda_1|, \cdots, |\lambda_n|\}p$ *temos que*

$$|a| - a = 2[-a]_+ \quad e \quad |a| + a = 2[a]_+,$$

onde $|x| = (x^2)^{1/2}, \quad \forall x \in \mathcal{S}$.

(b) Para todo $(a, b) \in \mathcal{S} \times \mathcal{S}$, *temos que*

$$a, b \in \mathcal{K} \text{ e } \langle a, b\rangle = 0 \text{ se, e somente se, } a = [a-b]_+.$$

(c) Para $a \in \mathcal{K}$ *e* $b \in \mathcal{S}$ *arbitrários, se* $a^2 - b^2 \in \mathcal{K}$, *então* $a - b \in \mathcal{K}$.

Prova.

(a) Para qualquer $c \in \mathcal{K}$, temos que

$$\begin{aligned}\|a-c\|^2 &= \operatorname{tr}[(a-c)(a-c)] = \operatorname{tr}[p(a-c)(a-c)p^T] \\ &= \left\|pap^T - pcp^T\right\|^2 = \left\|diag\{\lambda_1, \cdots, \lambda_n\} - pcp^T\right\|^2 \\ &= \sum_{i=1}^{n}\left[(\lambda_i - \left(pcp^T\right)_{ii})^2 + \sum_{j\neq i}\left(pcp^T\right)_{ij}^2\right],\end{aligned}$$

onde a última igualdade usa o fato que pcp^T é simétrica. Como $pcp^T \in \mathcal{K}$, então $\left(pcp^T\right)_{ii} \geq 0 \quad \forall i$, e o lado direito da expressão acima é minimizado por c tal que $\left(pcp^T\right)_{ii} = \max\{0, \lambda_i\}$ e $\left(pcp^T\right)_{ij} = 0$ para todo $i \neq j$, isto é,

$$c = p^T diag\{\max\{0, \lambda_1\}, \cdots, \max\{0, \lambda_n\}\}p.$$

Agora vamos provar a segunda parte. Para $|a| = p^T diag\{|\lambda_1|, \cdots, |\lambda_n|\}p$, onde $p \in \mathcal{O}$ e $\lambda_1, \cdots, \lambda_n \in I\!R$ satisfazem $a = p^T diag\{\lambda_1, \cdots, \lambda_n\}p$ temos que

$$\begin{aligned}|a| - a &= |a| + a - 2a \\ &= p^T diag\{|\lambda_1|, \cdots, |\lambda_n|\}p + p^T diag\{\lambda_1, \cdots, \lambda_n\}p - 2a\end{aligned}$$

$$
\begin{aligned}
&= p^T diag\{|\lambda_1| + \lambda_1, \cdots, |\lambda_n| + \lambda_n\}p - 2a \\
&= 2p^T diag\{\max\{0, \lambda_1\}, \cdots, \max\{0, \lambda_n\}\}p - 2a \\
&= 2([a]_+ - a) = 2[-a]_+,
\end{aligned}
$$

onde a penúltima igualdade segue da primeira parte de (a) e a última segue do Lema 2.2.1 (a). Analogamente, obtém-se que $|a| + a = 2[a]_+$.

(b) Considere $(a, b) \in \mathcal{K} \times \mathcal{K}$ arbitrário tal que $\langle a, b \rangle = 0$. Para qualquer $c \in \mathcal{K}$,

$$
\begin{aligned}
\|(a - b) - c\|^2 &= \|(a - c) - b\|^2 = \|a - c\|^2 - 2\langle a - c, b\rangle + \|b\|^2 \\
&= \|a - c\|^2 + 2\langle c, b\rangle + \|b\|^2 .
\end{aligned}
$$

Como $b \in \mathcal{K}$, por (1.1), $\langle c, b\rangle \geq 0$, então o lado direito da expressão acima atinge seu mínimo em $c = a$. Logo, $a = [a - b]_+$. Reciprocamente, considere $(a, b) \in \mathcal{S} \times \mathcal{S}$ arbitrário tal que $a = [a - b]_+$. Então $a \in \mathcal{K}$ e pelo Lema 1.3.1,

$$
0 \leq \|(a - c) - b\|^2 - \|b\|^2 = \|a - c\|^2 + 2\langle c - a, b\rangle \quad \forall c \in \mathcal{K}. \tag{2.5}
$$

Para qualquer $z \in \mathcal{K}$ e para qualquer $t \in (0, \infty)$, temos que $c := a + tz \in \mathcal{K}$. Assim, substituindo na equação (2.5), obtemos

$$
0 \leq t^2 \|z\|^2 + 2t\langle z, b\rangle .
$$

Dividindo ambos os lados por t e fazendo $t \longrightarrow 0$ segue que $0 \leq \langle z, b\rangle$ $\forall z \in \mathcal{K}$, implicando que, por (1.1), $b \in \mathcal{K}$. Similarmente, para qualquer $t \in (0, 1]$, temos $c := (1 - t)a \in \mathcal{K}$. Assim, substituindo na equação (2.5), obtemos

$$
0 \leq t^2 \|a\|^2 - 2t\langle a, b\rangle .
$$

Dividindo ambos os lados por t e fazendo $t \longrightarrow 0$ segue que $0 \leq -\langle a, b\rangle$. Como $a \in \mathcal{K}$, $b \in \mathcal{K}$, por (1.1) temos que $\langle a, b\rangle \geq 0$. Logo $\langle a, b\rangle = 0$.

(c) Fixe quaisquer $a \in \mathcal{K}$ e $b \in \mathcal{S}$ tais que $a^2 - b^2 \in \mathcal{K}$. Sabemos, por (a), que $|b| - b$ é sdp e como

$$a - b = a - |b| + |b| - b,$$

basta mostrar que $a - |b|$ é sdp. Suponha que $a - |b|$ não é sdp, então existe $0 \neq v \in I\!R^n$ e $\lambda \in (-\infty, 0)$ com $(a - |b|)v = \lambda v$. Como $a \in \mathcal{K}$, então

$$pap^T = \begin{bmatrix} \tilde{a}_{II} & 0 \\ 0 & 0 \end{bmatrix}$$

para algum $p \in \mathcal{O}$, $I \subset \{1, \cdots, n\}$ e para alguma submatriz $\tilde{a}_{II}$ dp. Como $a^2 - b^2 \in \mathcal{K}$, temos

$$p\,|b|\,p^T = \begin{bmatrix} \tilde{b}_{II} & 0 \\ 0 & 0 \end{bmatrix}$$

para alguma submatriz $\tilde{b}_{II}$ sdp. Logo,

$$\begin{bmatrix} (\tilde{a}_{II} - \tilde{b}_{II})(pv)_I \\ 0 \end{bmatrix} = p(a - |b|)p^T(pv) = p(a - |b|)v = \lambda pv$$

implicando que $[pv]_I \neq 0$. Como

$$(a + |b|)(a - |b|) + (a - |b|)(a + |b|) = 2a^2 - 2b^2 \in \mathcal{K},$$

temos que

$$\begin{aligned} 0 &\leq v^T((a + |b|)(a - |b|) + (a - |b|)(a + |b|))v = 2\lambda v^T(a + |b|)v \\ &= 2\lambda(pv)^T p(a + |b|)p^T pv = 2\lambda [pv]_I^T (\tilde{a}_{II} + \tilde{b}_{II})\,[pv]_I < 0, \end{aligned}$$

onde a última desigualdade segue do fato que $\tilde{a}_{II} + \tilde{b}_{II}$ é dp, $[pv]_I \neq 0$ e $\lambda < 0$, gerando uma contradição. ∎

No que segue, usaremos "O(t)" (respectivamente, "o(t)") para denotar um elemento de $\mathcal{S}$ que depende de t e tal que

$$\limsup_{t \longrightarrow 0} \frac{\|\mathrm{O}(t)\|}{t} < \infty \quad \left(\text{respectivamente, } \limsup_{t \longrightarrow 0} \frac{\|\mathrm{o}(t)\|}{t} = 0\right).$$

Lema 2.2.3 *Fixe $c \in \mathcal{K}$ e $p \in \mathcal{O}$ arbitrários tais que a equação (2.3) vale para algum $I \subset \{1, \cdots, n\}$ e para alguma submatriz $\widetilde{c}_{II}$ dp. Para cada $w \in \mathcal{S}$ tal que $c^2 + w \in \mathcal{K}$, fazendo $z = (c^2 + w)^{1/2} - c$ e*

$$\widetilde{w} = \begin{bmatrix} \widetilde{w}_{II} & \widetilde{w}_{IJ} \\ \widetilde{w}_{IJ}^T & \widetilde{w}_{JJ} \end{bmatrix} := pwp^T, \quad \widetilde{z} = \begin{bmatrix} \widetilde{z}_{II} & \widetilde{z}_{IJ} \\ \widetilde{z}_{IJ}^T & \widetilde{z}_{JJ} \end{bmatrix} := pzp^T, \tag{2.6}$$

onde $J = \{1, \cdots, n\} \backslash I$, então temos

$$\|\widetilde{z}_{JJ}\| \leq n^{1/4} \|\widetilde{w}_{JJ}\|^{1/2},$$

$$\widetilde{z}_{IJ} = \widetilde{c}_{II}^{-1} \widetilde{w}_{IJ} + o(\|w\|) \quad e \quad \widetilde{z}_{II} = L_{\widetilde{c}_{II}}^{-1}[\widetilde{w}_{II}] + o(\|w\|).$$

Prova. Elevando ambos os lados de $(c^2 + w)^{1/2} = c + z$ ao quadrado, multiplicando a esquerda por p e a direita por p^T e usando as equações (2.3) e (2.6), temos

$$\begin{bmatrix} \widetilde{c}_{II}^2 + \widetilde{w}_{II} & \widetilde{w}_{IJ} \\ \widetilde{w}_{IJ}^T & \widetilde{w}_{JJ} \end{bmatrix} = p(c^2 + w)p^T = p(c+z)^2 p^T = \begin{bmatrix} \widetilde{c}_{II} + \widetilde{z}_{II} & \widetilde{z}_{IJ} \\ \widetilde{z}_{IJ}^T & \widetilde{z}_{JJ} \end{bmatrix}^2$$

ou, equivalentemente:

$$\begin{aligned} \widetilde{w}_{II} &= \widetilde{c}_{II}\widetilde{z}_{II} + \widetilde{z}_{II}\widetilde{c}_{II} + \widetilde{z}_{II}^2 + \widetilde{z}_{IJ}\widetilde{z}_{IJ}^T \\ \widetilde{w}_{IJ} &= \widetilde{c}_{II}\widetilde{z}_{IJ} + \widetilde{z}_{II}\widetilde{z}_{IJ} + \widetilde{z}_{IJ}\widetilde{z}_{JJ}, \\ \widetilde{w}_{JJ} &= \widetilde{z}_{IJ}^T\widetilde{z}_{IJ} + \widetilde{z}_{JJ}^2. \end{aligned} \tag{2.7}$$

Da última equação de (2.7), temos

$$\|\widetilde{z}_{IJ}\|^2 + \|\widetilde{z}_{JJ}\|^2 = \mathrm{tr}[\widetilde{w}_{JJ}] \leq \sqrt{n}\, \|\widetilde{w}_{JJ}\|,$$

onde a desigualdade segue do Lema 1.2.3. Portanto, obtemos que

$$\|\widetilde{z}_{JJ}\| \leq n^{1/4} \|\widetilde{w}_{JJ}\|^{1/2}.$$

Podemos exigir que quando $\|\widetilde{w}_{II}\| \to 0$ e $\|\widetilde{w}_{IJ}\| \to 0$, tenhamos $\|\widetilde{z}_{II}\| \to 0$. Caso contrário, a primeira equação de (2.7) junto com $\|\widetilde{z}_{IJ}\| \to 0$ implica no limite (usando também a continuidade da multiplicação de matriz) que $0 = \widetilde{c}_{II}\widetilde{z}_{II} + \widetilde{z}_{II}\widetilde{c}_{II} + \widetilde{z}_{II}^2$ tem uma solução não-nula $\widetilde{z}_{II}$. Adicionando $\widetilde{c}_{II}^2$ em ambos os lados temos $\widetilde{c}_{II}^2 = (\widetilde{c}_{II} + \widetilde{z}_{II})^2$ e, como $\widetilde{c}_{II}$ e $\widetilde{c}_{II} + \widetilde{z}_{II}$ são ambos sdp implicando que $\widetilde{z}_{II} = 0$, chega-se a uma contradição. Assim, a segunda equação de (2.7) torna-se

$$\begin{aligned} \widetilde{w}_{IJ} &= (\widetilde{c}_{II} + \widetilde{z}_{II})\widetilde{z}_{IJ} + \widetilde{z}_{IJ}\mathrm{O}(\|\widetilde{w}_{JJ}\|^{1/2}) \\ &= \widetilde{c}_{II}\widetilde{z}_{IJ} + \mathrm{o}(\|w\|). \end{aligned}$$

Como $\widetilde{c}_{II}$ é dp, obtemos do Teorema da função implícita que

$$\widetilde{z}_{IJ} = \widetilde{c}_{II}^{-1}\widetilde{w}_{IJ} + \mathrm{o}(\|w\|).$$

Finalmente, da primeira equação (2.7) temos

$$\widetilde{w}_{II} = \widetilde{c}_{II}\widetilde{z}_{II} + \widetilde{z}_{II}\widetilde{c}_{II} + \widetilde{z}_{II}^2 + \widetilde{z}_{IJ}\widetilde{z}_{IJ}^T = L_{\widetilde{c}_{II}}[\widetilde{z}_{II}] + \widetilde{z}_{II}^2 + \mathrm{O}(\|w\|^2).$$

Como $\widetilde{c}_{II}$ é dp, de modo que $L_{\widetilde{c}_{II}}$ é uma aplicação linear inversível e $\|\widetilde{z}_{II}\| \to 0$ quando $\|w\| \to 0$, obtemos do Teorema da função implícita que

$$\widetilde{z}_{II} = L_{\widetilde{c}_{II}}^{-1}[\widetilde{w}_{II}] + \mathrm{o}(\|w\|) + \mathrm{O}(\|w\|^2) = L_{\widetilde{c}_{II}}^{-1}[\widetilde{w}_{II}] + \mathrm{o}(\|w\|).$$

■

Finalmente, lembraremos alguns conceitos sobre as aplicações $F : \mathcal{S} \longrightarrow \mathcal{S}$.

Definição 2.2.2 *(a) F e G são relativamente pseudomonótonas sobre $\mathcal{D} \subseteq \mathcal{S}$ se,*

$$\langle F(x), G(x) - G(y) \rangle \leq 0 \Longrightarrow \langle F(y), G(x) - G(y) \rangle \leq 0, \quad \forall\ x, y \in \mathcal{D}.$$

(b) F e G são relativamente monótonas se,

$$\langle F(x) - F(y), G(x) - G(y) \rangle \geq 0, \quad \forall\ x, y \in \mathcal{D}.$$

(c) F e G são relativamente estritamente monótonas se,

$$\langle F(x) - F(y), G(x) - G(y) \rangle > 0, \quad \forall\, x, y \in \mathcal{D},\ x \neq y.$$

(d) F e G são relativamente fortemente monótonas com módulo α se, existe $\alpha > 0$ tal que

$$\langle F(x) - F(y), G(x) - G(y) \rangle \geq \alpha \|x - y\|^2, \quad \forall\, x, y \in \mathcal{D}.$$

No caso onde $G = I$, as três condições acima se reduzem a F, sendo, respectivamente, pseudomonótona, monótona, estritamente monótona e fortemente monótona. Quando $f : \mathcal{D} \longrightarrow I\!R$ é diferenciável, no sentido de Fréchet, sobre $\mathcal{D}$, o gradiente de f é denotado por $\nabla f : \mathcal{D} \longrightarrow \mathcal{S}$.

2.3 Funções de Mérito

Uma função que define um problema de minimização equivalente ao $PCSD$ é chamada função de mérito. Nesta seção, apresentamos algumas funções de mérito estendidas do PCN [25, 27] para o $PCSD$ por Tseng [32] e uma nova função de mérito tanto para o PCN como para o $PCSD$ obtida por Yamashita e Fukushima [38]. Além disso, citamos algumas propriedades.

Definição 2.3.1 *Uma função $f : \mathcal{D} \longrightarrow [0, \infty)$ é uma função de mérito para o $PCSD$ sobre $\mathcal{D} \subseteq \mathcal{S}$ (tipicamente, $\mathcal{D} = \mathcal{S}$ ou $\mathcal{D} = G^{-1}(\mathcal{K})$) desde que*

$$f(x) = 0 \quad \text{se, e somente se,} \quad x \quad \text{é solução do} \quad PCSD.$$

Podemos reformular o $PCSD$ como um problema de minimização

$$\min_{x \in \mathcal{D}} f(x)$$

e aplicar um método adequado para resolver este problema. Para obter este método é desejável que a função de mérito tenha propriedades tais como:

- f seja diferenciável (pelos menos duas vezes),
- f seja convexa,
- qualquer ponto estacionário (x é tal que $\nabla f(x) = 0$) ou de mínimo local de f sobre $\mathcal{D}$ seja também um minimizador global de f,
- f forneça uma cota de erro global para o $PCSD$, ou seja, existe uma constante $\eta > 0$ tal que

$$dist\{x, X\} \leq \eta f(x), \quad \forall\, x \in \mathcal{S},$$

onde X é o conjunto solução do $PCSD$. Apresentamos abaixo as funções de mérito para $PCSD$ obtidas por Tseng [32] e Yamashita e Fukushima [38].

1. **Funções gap**

Seja $f : G^{-1}(\mathcal{K}) \longrightarrow I\!R \cup \{\infty\}$. A função gap é definida por

$$f(x) = \max_{y \in \mathcal{K}} \langle F(x), G(x) - y \rangle,$$

a qual é uma função de mérito sobre $G^{-1}(\mathcal{K})$. A função gap "dual" é dada por

$$f(x) = \max_{y \in \mathcal{K}} \langle F(G^{-1}(y)), G(x) - y \rangle,$$

a qual é uma função de mérito sobre $G^{-1}(\mathcal{K})$ desde que F e G sejam relativamente pseudomonótonas sobre $G^{-1}(\mathcal{K})$ e G^{-1} seja definida e contínua sobre $\mathcal{K}$. Devido à relação entre $(\mathcal{S}, < \cdot, \cdot >)$ e $(I\!R^v, < \cdot, \cdot >_{I\!R^v})$ com $v := n(n+1)/2$, a maioria dos resultados são extensões de resultados conhecidos em otimização [3, 19]. A função gap é convexa sobre $G^{-1}(\mathcal{K})$ se F e G são afins e relativamente monótonas sobre

$G^{-1}(\mathcal{K})$. Por outro lado, a função gap "dual" é convexa sobre $G^{-1}(\mathcal{K})$ se F e G são relativamente pseudomonótonas sobre $G^{-1}(\mathcal{K})$, G^{-1} é definida e contínua sobre $\mathcal{K}$ e G afim.

2. Função gap regularizada

A função gap regularizada, parametrizada por $\alpha > 0$, definida por

$$f_\alpha(x) = \max_{y \in \mathcal{K}} \left\{ \langle F(x), G(x) - y \rangle - \frac{1}{2\alpha} \|G(x) - y\|^2 \right\}, \tag{2.8}$$

é uma função de mérito sobre $G^{-1}(\mathcal{K})$. Como as funções de gap dadas anteriormente, a maioria dos resultados são extensões de resultados conhecidos [2, 12, 33]. Para qualquer $\alpha \in (0, \infty)$, a função gap regularizada $f_\alpha : G^{-1}(\mathcal{K}) \longrightarrow I\!R$ é diferenciável se F e G são diferenciáveis sobre $G^{-1}(\mathcal{K})$.

3. Lagrangeano Implícito

O Lagrangeano Implícito, parametrizado por $\alpha > 1$, definida por

$$\begin{aligned} f_\alpha(x) &= \max_{y,z \in \mathcal{K}} \{ \langle F(x), G(x) - z \rangle - \\ &- \langle y, G(x) \rangle - \frac{1}{2\alpha} \left(\|F(x) - y\|^2 + \|G(x) - z\|^2 \right) \}, \end{aligned} \tag{2.9}$$

é uma função de mérito sobre $\mathcal{S}$. Como as funções de mérito f dadas anteriormente, a maioria dos resultados são extensões de resultados conhecidos [33, 37]. Para qualquer $\alpha \in (1, \infty)$, seja $f_\alpha : \mathcal{S} \longrightarrow I\!R$ dada pela equação acima, então f_α é diferenciável se F é diferenciável. Além disso, qualquer ponto estacionário da função gap regularizada e do Lagrangeano Implícito é solução do $PCSD$ se F e G são diferenciáveis sobre $G^{-1}(\mathcal{K})$, $\nabla G(x)$ é inversível e $\nabla G^{-1}(x) \nabla F(x)$ é dp.

4. **Função projeção residual**

A função $f : \mathcal{D} \longrightarrow I\!R$ definida por

$$f(x) = \|G(x) - [G(x) - F(x)]_+\|^2,$$

é uma função de mérito sobre $\mathcal{S}$. A função projeção residual fornece uma cota de erro para o $PCSD$ sob as hipóteses de F e G serem Lipschitz contínuas e relativamente fortemente monótonas sobre $\mathcal{S}$, e com isso obtém-se também cotas de erros para as funções dadas por (2.8) e (2.9).

5. **A norma quadrada da função Fischer-Burmeister**

A função $f : \mathcal{S} \longrightarrow I\!R$ definida por

$$f(x) = \frac{1}{2}\|\phi(F(x), G(x))\|^2, \tag{2.10}$$

onde $\phi : \mathcal{S} \times \mathcal{S} \longrightarrow \mathcal{S}$ é a função

$$\phi(a, b) = \sqrt{a^2 + b^2} - a - b, \tag{2.11}$$

é uma função de mérito sobre $\mathcal{S}$. Não é fácil estender a análise desta função de mérito do caso do PCN. Em particular, ϕ envolve a raiz quadrada de duas matrizes simétricas sdp, o que complica significatimente a análise e necessita do desenvolvimento de novos argumentos, veja [32]. Se F e G são diferenciáveis sobre $\mathcal{S}$, então f é diferenciável.

6. **Função Luo e Tseng**

A função $f : \mathcal{S} \longrightarrow I\!R$ definida por

$$f(x) = \psi_0(\langle F(x), G(x)\rangle) + \psi(F(x), G(x)), \tag{2.12}$$

onde $\psi_0 : I\!\!R \longrightarrow [0, \infty)$ é tal que

$$\psi_0(t) = 0 \ \text{ se, e somente se, } \ t \leq 0, \tag{2.13}$$

e $\psi : \mathcal{S} \times \mathcal{S} \rightarrow [0, \infty)$ é tal que

$$\psi(a, b) = 0, \ \ \langle a, b \rangle \leq 0 \Longleftrightarrow \ \ (a, b) \in \mathcal{K} \times \mathcal{K}, \ \langle a, b \rangle = 0, \tag{2.14}$$

é uma função de mérito sobre $\mathcal{S}$. A função $f : \mathcal{S} \longrightarrow I\!\!R$ é diferenciável se ψ_0, ψ, F e G são diferenciáveis para qualquer $x \in \mathcal{S}$. A função Luo e Tseng é convexa se ψ_0, ψ são convexas e F e G são afins e relativamente monótonas. Além disso, qualquer ponto estacionário, desta função e da função dada por (2.10), é solução do $PCSD$ se F e G são diferenciáveis sobre $G^{-1}(\mathcal{K})$, $\nabla G(x)$ é inversível e $\nabla G^{-1}(x) \nabla F(x)$ é sdp.

7. Função proposta por Yamashita e Fukushima

A função proposta por Yamashita e Fukushima [38], dada por

$$f(x) = \psi_0(\langle x, F(x) \rangle) + \psi(x, F(x)), \tag{2.15}$$

onde $\psi_0 : I\!\!R \longrightarrow I\!\!R$ é tal que

$$\psi_0(t) = \frac{1}{4}(\max\{0, t\})^4, \tag{2.16}$$

e $\psi : \mathcal{S} \times \mathcal{S} \rightarrow I\!\!R$ é tal que

$$\psi(a, b) = \frac{1}{2} \left\| \sqrt{a^2 + b^2} - a - b \right\|^2, \tag{2.17}$$

é uma função de mérito sobre $\mathcal{S}$. Observe que eles estudaram o $PCSD$ com $G = I$. Esta função não é reduzida para uma função de mérito existente para PCN quando o conjunto $\mathcal{S}$ é restrito ao espaço das matrizes diagonais. Portanto é considerada uma nova função de mérito para o PCN. A função f nem sempre é convexa e é diferenciável quando F é diferenciável. A função f fornece uma cota de erro para o $PCSD$ desde que F seja fortemente monótona, e qualquer ponto estacionário de f é uma solução do $PCSD$ se F é monótona. Além disso, se $F : \mathcal{S} \longrightarrow \mathcal{S}$ é diferenciável e satisfaz uma das condições: ou F é monótona e o $PCSD$ é estritamente viável, ou F é fortemente monótona, então os conjuntos de nível

$$\mathcal{L}(\gamma) = \{x \in \mathcal{S} \,|\, f(x) \leq \gamma\}$$

são não-vazios e limitados para qualquer $\gamma \geq 0$. Yamashita e Fukushima também propuseram um método de descida e provaram sua convergência sob a condição de F ser continuamente diferenciável e uma das condições acima ocorra.

Observe que a função proposta por Yamashita e Fukushima (2.15)-(2.17) é uma modificação da função proposta por Tseng (2.12)-(2.14), com $G = I$, onde ψ_0 dada por (2.16), satisfaz (2.13) e ψ, dada por (2.17) satisfaz

$$\psi(a, b) = 0 \Longleftrightarrow \quad (a, b) \in \mathcal{K} \times \mathcal{K}, \; \langle a, b \rangle = 0. \tag{2.18}$$

2.4 Existência de solução do $PCSD$

Nesta seção, apresentamos alguns resultados que garantem a existência de solução para o $PCSD$. Para o $PCSD$, Tseng [32], Yamashita e Fukushima [38] mostraram que qualquer ponto estacionário, das funções definidas por (2.10), (2.12)-(2.14) e (2.15)-(2.17), é uma solução do $PCSD$ sob a condição de $\nabla F(x)$ ser sdp. Para o PCN, resultados similares foram provados sob condições mais fracas. De fato, Fachinei e Soares [10] mostraram que qualquer ponto estacionário da função Fischer-Burmeister é uma solução do PCN quando $\nabla F(x) : I\!R^n \rightarrow I\!R^n$ é uma

P_0-matriz, isto é, para qualquer $x \in \mathbb{R}^n$ com $x \neq 0$, existe um índice i tal que

$$x_i \neq 0 \;\text{ e }\; x_i[\nabla F(x)x]_i \geq 0.$$

Tseng [32] levantou a questão da extensão desta propriedade para o espaço das matrizes simétricas, $\mathcal{S}$. Respondendo esta questão, Qi e Chen [38] obtiveram esta propriedade para o $PCSD$, considerando a seguinte definição para P_0-matriz:

Definição 2.4.1 *Uma aplicação $M : \mathcal{S} \longrightarrow \mathcal{S}$ é uma P_0-matriz (respectivamente, $P-$matriz) se para quaisquer $x \in \mathcal{S}$, com $x \neq 0$, e $p \in \mathcal{O}$, existe um i tal que $(px)_i \neq 0$ e $\left[(px)M(px)^T\right]_{ii} \geq 0$ (respectivamente, $\left[(px)M(px)^T\right]_{ii} > 0$), onde $(px)_i$ é a i-ésima linha da matriz px.*

Observe que, pela definição acima, uma aplicação sdp (dp) é uma aplicação P_0-matriz (P-matriz). E, uma aplicação P-matriz é uma aplicação P_0-matriz.

A existência de solução para o PCN é garantida sob a condição de $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ser $P-$uniforme, ou seja, para quaisquer $x, y \in \mathbb{R}^n$, existe um índice i e $\mu > 0$ tais que

$$[(x-y)(F(x)-F(y))]_{ii} \geq \mu \|x-y\|^2. \tag{2.19}$$

É possível que esta propriedade seja preservada para o $PCSD$.

Para o $PCSD$, Shibata, Yamashita e Fukushima [30] mostraram a existência de solução sob a hipótese de F ser P-uniforme e $F(x) = \varpi(\lambda_{max}(x))$, onde $\varpi(t)$ denota um elemento de $\mathcal{S}$ que depende de t e cuja norma tende para o infinito mais lentamente que t^2, i.e., $\lim\limits_{t \longrightarrow \infty} \|\varpi(t)\| / t^2 = 0$ e $\lambda_{max}(x)$ é o maior autovalor de $x \in \mathcal{S}$. Agora, daremos os conceitos de função-P e função-P uniforme de $\mathcal{S}$ em $\mathcal{S}$.

Definição 2.4.2 *A função $F : \mathcal{S} \longrightarrow \mathcal{S}$ é uma* *(a) função-P se, para quaisquer $x, y \in \mathcal{S}$ com $x \neq y$ e qualquer $p \in \mathcal{O}$, existe um i tal que*

$$\left[p(x-y)(F(x)-F(y))p^T\right]_{ii} > 0,$$

(b) função-P uniforme se, para quaisquer $x, y \in \mathcal{S}$ e qualquer $p \in \mathcal{O}$, existe um i tal que

$$\left[p(x-y)(F(x)-F(y))p^T\right]_{ii} \geq \mu \|x-y\|^2,$$

onde μ é uma constante positiva.

O Lema a seguir mostra uma relação entre função-P e P_0-matriz que é importante para estudar condições para a existência de uma solução para o $PCSD$.

Lema 2.4.1 *Suponha que $F : \mathcal{S} \longrightarrow \mathcal{S}$ é diferenciável. Se F é uma função-P, então $\nabla F(x)$ é uma P_0-matriz. Além disso, se F é uma função-P uniforme, então para quaisquer $d \in \mathcal{S}$, com $d \neq 0$, e $p \in \mathcal{O}$ existe um i tal que*

$$\left[(pd)(\nabla F(x))(pd)^T\right]_{ii} \geq \mu \|d\|^2,$$

onde μ é uma constante positiva.

Prova. Seja $d \in \mathcal{S}$ com $d \neq 0$. Da diferenciabilidade de F temos que para todo $x \in \mathcal{S}$,

$$F(y) - F(x) = \nabla F(x)(td) + o(t)\|d\|, \tag{2.20}$$

onde $y = x + td$ e $t > 0$. Como F é uma função-P, existe um i tal que, para qualquer $p \in \mathcal{O}$, temos que

$$\left[p(y-x)(F(y)-F(x))p^T\right]_{ii} > 0, \tag{2.21}$$

e $0\neq [p(y-x)]_i = t(pd)_i$. Por (2.20) e (2.21),

$$\begin{aligned} & \left[p(td)\left\{t\nabla F(x)d + o(t)\left\|d\right\|\right\}p^T\right]_{ii} \\ = & \left[t^2(pd)\nabla F(x)(pd)^T + t(pd)o(t)\left\|d\right\|p^T\right]_{ii} > 0, \end{aligned} \tag{2.22}$$

onde a igualdade segue da simetria de d. Dividindo a desigualdade (2.22) por t^2 e fazendo $t \to 0$, temos que

$$\left[(pd)(\nabla F(x))(pd)^T\right]_{ii} \geq 0.$$

Logo $\nabla F(x)$ é uma P_0-matriz. Analogamente, mostra-se a segunda afirmação do Lema. ■

Toda função-P uniforme é uma função-P, então se F é uma função-P uniforme diferenciável, temos, pelo Lema 2.4.1, que $\nabla F(x)$ é uma P_0-matriz para qualquer x.

Suponha que $F(x) = Mx + q$. Se F é uma função-P, então, pela Definição 2.4.2 (a), para quaisquer $x, y \in \mathcal{S}$, com $x \neq y$, e qualquer matriz ortogonal p existe um i tal que $\left[p(x-y)M(p(x-y))^T\right]_{ii} > 0$. Isto implica que M é uma $P-$matriz.

2.5 Funções -$PCSD^*$

Nesta seção apresentamos algumas funções $\psi : \mathcal{S} \times \mathcal{S} \longrightarrow [0, \infty)$ que satisfazem (2.18) e algumas propriedades interessantes que elas apresentam. Denominaremos estas funções de funções-$PCSD^*$.

O Lema abaixo apresenta uma função ψ cujas propriedades garantem que a função dada pela equação (2.8) é uma função de mérito sobre $G^{-1}(\mathcal{K})$. A maioria dos resultados são extensões de [2, 12, 33].

Lema 2.5.1 *Para $\alpha \in (0,\infty)$ arbitrário, defina a função $\psi_\alpha : \mathcal{S} \times \mathcal{S} \longrightarrow I\!R$ por*

$$\psi_\alpha(a,b) = \max_{y\in\mathcal{K}} \left\{ \langle a, b-y\rangle - \frac{1}{2\alpha}\|b-y\|^2 \right\}.$$

Então

(a) Para $(a,b) \in \mathcal{S}\times\mathcal{K}$ arbitrário, temos

$$\psi_\alpha(a,b) \geq \frac{1}{2\alpha}\|b-[b-\alpha a]_+\|^2, \tag{2.23}$$

$$\psi_\alpha(a,b) = 0 \quad \textit{se, e somente se,} \quad a \in \mathcal{K}, \ \ \langle a,b\rangle = 0.$$

(b) ψ_α é diferenciável para todo $(a,b)\in\mathcal{S}\times\mathcal{S}$, com

$$\nabla_a\psi_\alpha(a,b) = b-[b-\alpha a]_+, \qquad \nabla_b\psi_\alpha(a,b) = a - \frac{1}{\alpha}(b-[b-\alpha a]_+).$$

Prova.

(a) Como $b \in \mathcal{K}$, por (1.1),

$$\left\langle b, [-a+\frac{1}{\alpha}b]_- \right\rangle \leq 0.$$

Pelo Lema 2.2.1 (a), podemos escrever a equação acima assim

$$0 \geq \left\langle b, [-a+\frac{1}{\alpha}b]_- \right\rangle = \frac{1}{\alpha}\langle b, -\alpha a + b - [-\alpha a + b]_+\rangle .$$

Logo,

$$\langle a,b\rangle \geq \left\langle \frac{b}{\alpha}, b-[b-\alpha a]_+ \right\rangle . \tag{2.24}$$

Mas

$$\begin{aligned}
\psi_\alpha(a,b) &= \langle a,b\rangle - \langle a, [b-\alpha a]_+\rangle - \frac{1}{2\alpha}\|b-[b-\alpha a]_+\|^2, \\
&\geq \left\langle \frac{b}{\alpha}, b-[b-\alpha a]_+ \right\rangle - \langle a,[b-\alpha a]_+\rangle - \frac{1}{2\alpha}\|b-[b-\alpha a]_+\|^2 \\
&= \frac{1}{\alpha}\|b-[b-\alpha a]_+\|^2 + \frac{1}{\alpha}\langle [b-\alpha a]_+, b-\alpha a - [b-\alpha a]_+\rangle
\end{aligned}$$

$$
\begin{aligned}
&- \frac{1}{2\alpha}\|b-[b-\alpha a]_+\|^2 \\
&= \frac{1}{2\alpha}\|b-[b-\alpha a]_+\|^2 + \frac{1}{\alpha}\langle [b-\alpha a]_+, [b-\alpha a]_- \rangle \\
&= \frac{1}{2\alpha}\|b-[b-\alpha a]_+\|^2 .
\end{aligned}
$$

A desigualdade segue de (2.24), na segunda expressão somou-se e subtraiu-se $\frac{1}{\alpha}[b-\alpha a]_+$ ao termo b/α para obter a segunda igualdade, e pelo Lema 2.2.1 (a) obtém-se a penútima e última igualdades. Logo, obtemos (2.23). Assim $\psi_\alpha(a,b) \geq 0$. Se $\psi_\alpha(a,b) = 0$ então por (2.23) temos $b = [b-\alpha a]_+$. Portanto pelo Lema 2.2.2 (b), $a \in \mathcal{K}$ e $\langle a, b\rangle = 0$. Reciprocamente, se $a \in \mathcal{K}$, e $\langle a, b\rangle = 0$, então

$$\psi_\alpha(a,b) = -\langle a, [b-\alpha a]_+\rangle - \frac{1}{2\alpha}\|b-[b-\alpha a]_+\|^2 \leq 0.$$

Por (2.23) $\psi_\alpha(a,b) \geq 0$. Portanto $\psi_\alpha(a,b) = 0$.

(b) Devido a relação entre $(\mathcal{S}, \langle\cdot,\cdot\rangle)$ e $(I\!R^v, \langle\cdot,\cdot\rangle_{I\!R^v})$ com

$$v := n(n+1)/2,$$

e [[3], cap. 4, Teorema 1.7], ψ_α é diferenciável e seu gradiente é dado por

$$\nabla_a \psi_\alpha(a,b) = b - [b-\alpha a]_+, \qquad \nabla_b \psi_\alpha(a,b) = a - \frac{1}{\alpha}\left(b - [b-\alpha a]_+\right).$$

■

Os resultados abaixo apresentam, para ψ definida por (2.17), a garantia que as funções dadas pelas equações (2.10) e (2.15) são funções de mérito sobre $\mathcal{S}$. Para isso, primeiro precisamos mostrar uma propriedade para ϕ dada por (2.11).

Lema 2.5.2 *(a) Para todo $(a,b) \in \mathcal{K} \times \mathcal{K}$, temos*

$$\langle a, b\rangle = 0 \textit{ se, e somente se, } [pbp^T]_{ij} = 0 \textit{ para todo } i \geq k \textit{ ou } j \geq k,$$

onde $p \in \mathcal{O}$ *e* $\lambda_k, \cdots, \lambda_n \in (0, \infty)$ *tais que* $pap^T = diag\{0, \cdots, 0, \lambda_k, \cdots, \lambda_n\}$.

(b) Seja ϕ *dada por* (2.11). *Então,*

$$\phi(a,b) = 0 \text{ se, e somente se, } a \in \mathcal{K}, \quad b \in \mathcal{K}, \quad \langle a, b\rangle = 0.$$

Prova.

(a) Fixe $(a,b) \in \mathcal{K} \times \mathcal{K}$ arbitrário. Sejam $p \in \mathcal{O}$ e $\lambda_k, \cdots, \lambda_n \in (0, \infty)$ tais que $pap^T = diag\{0, \cdots, 0, \lambda_k, \cdots, \lambda_n\}$. Então

$$\begin{aligned} \langle a, b\rangle = \mathrm{tr}[ab] &= \mathrm{tr}[pap^T pbp^T] \\ &= \mathrm{tr}[diag\{0, \cdots, 0, \lambda_k, \cdots, \lambda_n\} pbp^T] = \sum_{i=k}^{n} \lambda_i [pbp^T]_{ii}. \end{aligned}$$

Logo $\langle a, b\rangle = 0$ se, e somente se, $[pbp^T]_{ii} = 0$ para todo $i \geq k$. Como pbp^T é simétrica e sdp, os dois últimos valem se, e somente se, $[pbp^T]_{ij} = 0$ para todo $i \geq k$ ou $j \geq k$.

(b) Fixe $(a,b) \in \mathcal{S} \times \mathcal{S}$ arbitrário. Se $a, b \in \mathcal{K}$ e $\langle a, b\rangle = 0$, então, por (a), para $p \in \mathcal{O}$ e $\lambda_k, \cdots, \lambda_n \in (0, \infty)$ tais que $pap^T = diag\{0, \cdots, 0, \lambda_k, \cdots, \lambda_n\}$, temos que $[pbp^T]_{ij} = 0$ para todo $i \geq k$ ou $j \geq k$. Logo,

$$pabp^T = pap^T pbp^T = diag\{0, \cdots, 0, \lambda_k, \cdots, \lambda_n\} pbp^T = 0.$$

Assim, segue que $ab = 0$. Similarmente, $ba = 0$. Portanto $(a+b)^2 = a^2 + b^2$ ou, equivalentemente, $a + b = (a^2 + b^2)^{1/2}$, isto é, $\phi(a,b) = 0$. Reciprocamente, se $\phi(a,b) = 0$, então $a + b = (a^2 + b^2)^{1/2} \in \mathcal{K}$ e $(a+b)^2 = a^2 + b^2$ ou, equivalentemente, $ab + ba = 0$. Suponha que $p \in \mathcal{O}$ e $\lambda_1, \cdots, \lambda_n \in I\!R$ (onde $\lambda_1, \cdots, \lambda_{k-1}$ são nulos, $\lambda_k, \cdots, \lambda_{l-1}$ são positivos, e $\lambda_l, \cdots, \lambda_n$ são negativos para alguns valores de k e l que verificam $1 \leq k \leq l \leq n+1$) satisfazendo $pap^T = diag\{\lambda_1, \cdots, \lambda_n\}$. Logo

$$\begin{aligned} 0 &= p(ab + ba)p^T = pabp^T + pbap^T = pap^T pbp^T + pbp^T pap^T \\ &= diag\{\lambda_1, \cdots, \lambda_n\} pbp^T + pbp^T diag\{\lambda_1, \cdots, \lambda_n\}, \end{aligned}$$

implicando em $(\lambda_i + \lambda_j)(pbp^T)_{ij} = 0$ para todo i e j. Assim, $(pbp^T)_{ij} = 0$, exceto possivelmente quando $\lambda_i = \lambda_j = 0$ ou $\lambda_i\lambda_j < 0$. Portanto,

$$
\begin{aligned}
p(a+b)p^T &= pap^T + pbp^T \\
&= \begin{bmatrix}
\left[pbp^T\right]_{\substack{j<k \\ i<k}} & 0 & 0 \\
0 & diag\{\lambda_k, \cdots, \lambda_{l-1}\} & \left[pbp^T\right]_{\substack{j \geq l \\ k \leq i < l}} \\
0 & \left[pbp^T\right]_{\substack{k \leq j < l \\ i \geq l}} & diag\{\lambda_l, \cdots, \lambda_n\}
\end{bmatrix}.
\end{aligned}
$$

Como $p(a+b)p^T \in \mathcal{K}$, seus elementos diagonais são não-negativos; devemos ter $l = n+1$ e a submatriz $\left[pbp^T\right]_{\substack{j<k \\ i<k,}}$ tem que ser sdp. Logo, pap^T e pbp^T estão em $\mathcal{K}$, implicando, pelo Lema 1.2.5, que a e b estão em $\mathcal{K}$. Então pelo item (a), $\langle a, b\rangle = 0$. ∎

Agora, usando os Lemas 2.2.3 e 2.5.2 mostraremos as propriedades da função ψ dada por (2.17), cujos resultados são extensões de [24] e [14] para o caso do PCN.

Lema 2.5.3 *Seja ϕ dada pela equação* (2.11) *e defina $\psi : \mathcal{S} \times \mathcal{S} \longrightarrow I\!R$ como*

$$\psi(a,b) = \frac{1}{2}\left\|\phi(a,b)\right\|^2.$$

Então

(a) Para todo $(a,b) \in \mathcal{S} \times \mathcal{S}$, temos $\psi(a,b) \geq 0$ e

$$\psi(a,b) = 0 \ \text{ se, e somente se, } \ a, b \in \mathcal{K}, \ \langle a, b\rangle = 0.$$

(b) ψ é diferenciável em todo $(a,b) \in \mathcal{S} \times \mathcal{S}$. Além disso,

$$\nabla_a\psi(a,b) = sym[L_c^{-1}[c-a-b][a-c]],$$

$$\nabla_b\psi(a,b) = sym[L_c^{-1}[c-a-b][b-c]],$$

onde $c = (a^2 + b^2)^{1/2}$.

(c) Para todo $(a, b) \in \mathcal{S} \times \mathcal{S}$, *temos*

$$\langle \nabla_a \psi(a, b), \nabla_b \psi(a, b) \rangle \geq \|(c - a - b) g\|^2 ,$$

onde $c = (a^2 + b^2)^{1/2}$ *e* $g = L_c^{-1}[c - a - b]$.

(d) Para todo $(a, b) \in \mathcal{S} \times \mathcal{S}$, *temos*

$$\langle a, \nabla_a \psi(a, b) \rangle + \langle b, \nabla_b \psi(a, b) \rangle = \|c - a - b\|^2 ,$$

onde $c = (a^2 + b^2)^{1/2} \in \mathcal{K}$.

Prova.

(a) A não-negatividade segue da definição de ψ. Agora, seja $\psi(a, b) = 0$, o que implica que $\phi(a, b) = 0$, então pelo Lema 2.5.2 (b), $a, b \in \mathcal{K}$ e $\langle a, b \rangle = 0$. Reciprocamente, se $a, b \in \mathcal{K}$ e $\langle a, b \rangle = 0$, então, pelo Lema 2.5.2 (b), $\phi(a, b) = 0$ o que implica que $\psi(a, b) = 0$.

(b) Note que

$$\begin{aligned} \psi(a, b) &= \frac{1}{2} \left\| \left(a^2 + b^2\right)^{1/2} - a - b \right\|^2 = \frac{1}{2} \mathrm{tr} \left[\left(\left(a^2 + b^2\right)^{1/2} - a - b \right)^2 \right] \\ &= \mathrm{tr} \left[a^2 + b^2 + ab - \left(a^2 + b^2\right)^{1/2} (a + b) \right] . \end{aligned} \tag{2.25}$$

Fixe qualquer $(a, b) \in \mathcal{S} \times \mathcal{S}$ e seja $c = (a^2 + b^2)^{1/2}$. Assim, pelo Lema 2.2.1 (b), temos que $a, b \in \mathcal{S}_c$. Fixe $u \in \mathcal{S}$ arbitrário e defina $w = au + ua + u^2$ e $z = \left(c^2 + w\right)^{1/2} - c$. Pelo Lema 2.2.3, temos $\tilde{z}_{II} = L_{\tilde{c}_{II}}^{-1}[\tilde{w}_{II}] + o(\|w\|)$, onde $\tilde{w}_{II}$ e $\tilde{z}_{II}$ (como também $\tilde{z}_{IJ}$ e $\tilde{z}_{JJ}$) são dadas pela equação (2.6). Assim,

$$\mathrm{tr}[z(a + b)] = \mathrm{tr} \left[p z p^T (p a p^T + p b p^T) \right]$$

$$
\begin{aligned}
&= \operatorname{tr}\left[\begin{bmatrix} \widetilde{z}_{II} & \widetilde{z}_{IJ} \\ \widetilde{z}_{IJ}^T & \widetilde{z}_{JJ} \end{bmatrix}\begin{bmatrix} \widetilde{a}_{II}+\widetilde{b}_{II} & 0 \\ 0 & 0 \end{bmatrix}\right] \\
&= \operatorname{tr}\left[\widetilde{z}_{II}(\widetilde{a}_{II}+\widetilde{b}_{II})\right] = \operatorname{tr}\left[\left(L^{-1}_{\widetilde{c}_{II}}[\widetilde{w}_{II}] + \mathrm{o}(\|w\|)\right)(\widetilde{a}_{II}+\widetilde{b}_{II})\right] \\
&= \operatorname{tr}\left[L^{-1}_{\widetilde{c}_{II}}[\left[p(au+ua)p^T\right]_{II}](\widetilde{a}_{II}+\widetilde{b}_{II})\right] + o(\|u\|) \\
&= \operatorname{tr}\left[L^{-1}_{\widetilde{c}_{II}}[\widetilde{a}_{II}+\widetilde{b}_{II}]\left[p(au+ua)p^T\right]_{II}\right] + o(\|u\|),
\end{aligned}
$$

onde a última igualdade segue da primeira identidade Lema 2.2.1 (d). Sejam

$$
\widetilde{d}_{II} = L^{-1}_{\widetilde{c}_{II}}[\widetilde{a}_{II}+\widetilde{b}_{II}] \text{ e } d = p^T \begin{bmatrix} \widetilde{b}_{II} & 0 \\ 0 & 0 \end{bmatrix} p \in \mathcal{S}_c.
$$

Então,

$$
\begin{aligned}
\operatorname{tr}[z(a+b)] &= \operatorname{tr}\left[\widetilde{d}_{II}\left[p(au+ua)p^T\right]_{II}\right] + o(\|u\|) \\
&= \operatorname{tr}\left[\begin{bmatrix} \widetilde{d}_{II} & 0 \\ 0 & 0 \end{bmatrix} p(au+ua)p^T\right] + o(\|u\|) \\
&= \operatorname{tr}[d(au+ua)] + o(\|u\|) \\
&= \operatorname{tr}\left[L_c^{-1}[a+b](au+ua)\right] + o(\|u\|),
\end{aligned} \tag{2.26}
$$

onde a última igualdade segue do fato que $\widetilde{c}_{II}\widetilde{d}_{II} + \widetilde{d}_{II}\widetilde{c}_{II} = \widetilde{a}_{II} + \widetilde{b}_{II}$ ou, equivalentemente,

$$
\begin{aligned}
cd + dc &= p^T \begin{bmatrix} \widetilde{c}_{II} & 0 \\ 0 & 0 \end{bmatrix} pp^T \begin{bmatrix} \widetilde{d}_{II} & 0 \\ 0 & 0 \end{bmatrix} p + p^T \begin{bmatrix} \widetilde{d}_{II} & 0 \\ 0 & 0 \end{bmatrix} pp^T \begin{bmatrix} \widetilde{c}_{II} & 0 \\ 0 & 0 \end{bmatrix} p \\
&= p^T \begin{bmatrix} \widetilde{a}_{II}+\widetilde{b}_{II} & 0 \\ 0 & 0 \end{bmatrix} p = a+b.
\end{aligned}
$$

Logo, $d = L_c^{-1}[a+b]$. Assim, por (2.25), temos

$$
\begin{aligned}
\psi(a+u,b) &- \psi(a,b) = \\
&= \operatorname{tr}[(a+u)^2 + (a+u)b + b^2 - \left((a+u)^2+b^2\right)^{1/2}(a+b+u) - \\
&- a^2 - ab - b^2 + c(a+b)] \\
&= \operatorname{tr}\left[2au + ub + u^2 - \left(c^2+w\right)^{1/2}(a+b+u) + c(a+b)\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \operatorname{tr}\left[2au + ub + u^2 - cu - \left(\left(c^2 + w\right)^{1/2} - c\right)(a + b + u)\right] \\
&= \operatorname{tr}\left[2au + ub + u^2 - cu - z(a + b + u)\right] \\
&= \operatorname{tr}\left[2au + bu - cu - L_c^{-1}[a + b](au + ua)\right] + o(\|u\|) \\
&= \left\langle 2a + b - c - L_c^{-1}[a + b]a - aL_c^{-1}[a + b], u\right\rangle + o(\|u\|),
\end{aligned}
$$

onde a segunda igualdade usa $c^2 = a^2 + b^2$ e $w = au + ua + u^2$, a quarta usa-se $z = (c^2 + w)^{1/2} - c$ e a quinta segue da equação (2.26). Logo,

$$
\begin{aligned}
\nabla_a\psi(a, b) &= 2a + b - c - L_c^{-1}[a + b]a - aL_c^{-1}[a + b] \\
&= L_c^{-1}[a + b](c - a) + (c - a)L_c^{-1}[a + b] + a - c \\
&= L_c^{-1}[a + b](c - a) + (c - a)L_c^{-1}[a + b] + 2\left[L_c^{-1}(c)(a - c)\right] \\
&= \left(L_c^{-1}[c] - L_c^{-1}[a + b]\right)(a - c) + (a - c)\left(L_c^{-1}[c] - L_c^{-1}[a + b]\right) \\
&= L_c^{-1}[c - a - b](a - c) + (a - c)L_c^{-1}[c - a - b] \\
&= sym[L_c^{-1}[c - a - b][a - c]],
\end{aligned}
$$

onde a segunda igualdade usa o fato que $x = L_c^{-1}[x]c + cL_c^{-1}[x]$ com $x = a + b$, a terceira usa que $\frac{1}{2}x = L_c^{-1}[c]x = xL_c^{-1}[c]$ (veja Lema 2.2.1 (d)) com $x = a - c$, a penúltima igualdade segue da linearidade de L_c^{-1} e a última segue da Definição 2.2.1 (b).

Um argumento análogo dá a fórmula para $\nabla_b\psi(a, b)$.

Para quaisquer $u, v \in \mathcal{S}$, de um argumento como o acima segue que

$$
\psi(a + u, b + v) - \psi(a, b) = \langle\nabla_a\psi(a, b), u\rangle + \langle\nabla_b\psi(a, b), v\rangle + o(\|u\|) + o(\|v\|).
$$

Assim, ψ é diferenciável em (a, b).

(c) Fixe $(a, b) \in \mathcal{S}\times\mathcal{S}$ arbitrário. Fazendo $x = c - a$, $y = c - b$ e $g = L_c^{-1}[c - a - b]$, temos

$$
\begin{aligned}
\langle \nabla_a \psi(a,b), \nabla_b \psi(a,b) \rangle &= \mathrm{tr}\left[(gx+xg)(gy+yg)\right] = 2\mathrm{tr}\left[xgyg+gxyg\right] \\
&= 2\mathrm{tr}\left[x^{1/2}gy^{1/2}y^{1/2}gx^{1/2} + g(c-a)(c-b)g\right] \\
&= 2\left\|y^{1/2}gx^{1/2}\right\|^2 + 2\mathrm{tr}\left[g(c^2-ac+ab-cb)g\right] \\
&= 2\left\|y^{1/2}gx^{1/2}\right\|^2 + \mathrm{tr}\left[g(2ab-2ac-2cb+c^2+a^2+b^2)g\right] \\
&= 2\left\|y^{1/2}gx^{1/2}\right\|^2 \\
&+ \mathrm{tr}\left[g(ab+ba-ac-ca-cb-bc+c^2+a^2+b^2)g\right] \\
&= 2\left\|y^{1/2}gx^{1/2}\right\|^2 + \mathrm{tr}\left[g(c-a-b)^2g\right] \\
&= 2\left\|y^{1/2}gx^{1/2}\right\|^2 + \|(c-a-b)g\|^2,
\end{aligned}
$$

onde a terceira igualdade usa o fato de $c^2 - a^2$, $c^2 - b^2$ e c pertencerem a $\mathcal{K}$, logo, pelo Lema 2.2.2 (c), $x = c - a$ e $y = c - b$ estão em $\mathcal{K}$; a quinta igualdade segue do fato que $c^2 = a^2 + b^2$; e a segunda e sexta igualdades seguem da propriedade de traço tr$[w]$ =tr$[w^T]$ para qualquer matriz $w \in \mathcal{X}$. Portanto,

$$
\langle \nabla_a \psi(a,b), \nabla_b \psi(a,b) \rangle \geq \|(c-a-b)g\|^2 .
$$

(d) Por (b) temos

$$
\begin{aligned}
\langle a, \nabla_a \psi(a,b) \rangle &+ \langle b, \nabla_b \psi(a,b) \rangle = \left\langle a, L_c^{-1}[c-a-b](a-c) + (a-c)L_c^{-1}[c-a-b] \right\rangle + \\
&+ \left\langle b, L_c^{-1}[c-a-b](b-c) + (b-c)L_c^{-1}[c-a-b] \right\rangle \\
&= \mathrm{tr}\left[L_c^{-1}[c-a-b]\left((a-c)a + a(a-c) + (b-c)b + b(b-c)\right)\right] \\
&= \mathrm{tr}\left[L_c^{-1}[c-a-b]L_c[c-a-b]\right] \\
&= \|c-a-b\|^2 ,
\end{aligned}
$$

onde a segunda igualdade segue da propriedade de traço, tr$[xy]$=tr$[yx]$, $\forall\ x, y \in \mathcal{X}$, a terceira segue da definição de c e de L_c e a última igualdade segue do Lema 2.2.1 (d). ■

A seguir, daremos exemplos de funções ψ que satisfazem (2.14). Para isso, definimos $h : \mathcal{S} \to I\!R$ por

$$h(x) = \|[x]_+\|^2. \tag{2.27}$$

Para mostrar sua convexidade precisamos do seguinte resultado.

Lema 2.5.4 *Para quaisquer $x, y \in \mathcal{S}$, temos*

$$\|[x+y]_+\| \leq \|[x]_+\| + \|[y]_+\|.$$

Prova. Consideremos o seguinte problema de otimização

$$\min_{z \in \mathcal{K}^0} \|x + y - z\|. \tag{2.28}$$

Como $[x]_- + [y]_- \in \mathcal{K}^0$, temos

$$\min_{z \in \mathcal{K}^0} \|x + y - z\| \leq \|x + y - [x]_- - [y]_-\| = \|[x]_+ + [y]_+\| \leq \|[x]_+\| + \|[y]_+\|,$$

onde a igualdade segue do Lema 2.2.1 (a) e a última desigualdade segue da desigualdade triangular. Pelo Lema 1.3.1, a solução ótima do problema (2.28) é atingida em $z = [x+y]_-$. Como, pelo Lema 2.2.1 (a),

$$\|[x+y]_+\| = \|x + y - [x+y]_-\|$$

para quaisquer $x, y \in \mathcal{S}$, temos

$$\|[x+y]_+\| = \min_{z \in \mathcal{K}^0} \|x + y - z\| \leq \|[x]_+\| + \|[y]_+\|.$$

■

Usando o Lema acima, mostraremos que h é convexa e diferenciável em $\mathcal{S}$.

Lema 2.5.5 *Seja h definida por* (2.27). *Então*

(a) h é convexa em $\mathcal{S}$;

(b) h é diferenciável em $\mathcal{S}$ com $\nabla h(x) = 2[x]_+$.

Prova.

(a) Sejam as funções $h_1 : \mathcal{S} \to I\!R$ e $h_2 : I\!R \to I\!R$ definidas por $h_1(u) := \|[u]_+\|$ e $h_2(t) := t^2$. Note que h_1 é uma função convexa em $\mathcal{S}$ pelo Lema 2.5.4. Então, para quaisquer $x, y \in \mathcal{S}$ e $\alpha \in [0,1]$, temos

$$\begin{aligned} h((1-\alpha)x + \alpha y)) &= h_2\left(h_1\left((1-\alpha)x + \alpha y\right)\right) \\ &\leq h_2\left((1-\alpha)h_1(x) + \alpha h_1(y)\right) \\ &\leq (1-\alpha)h_2(h_1(x)) + \alpha h_2(h_1(y)) \\ &= (1-\alpha)h(x) + \alpha h(y), \end{aligned}$$

onde a primeira desigualdade segue da convexidade e não-negatividade da função h_1, e a monotonicidade da função h_2 em $[0, \infty)$; e a segunda desigualdade segue da convexidade da função h_2.

(b) Para qualquer $x \in \mathcal{S}$, pelo Lema 2.2.1 (a), temos que $h(x) = \|x - [x]_-\|^2 = \min_{y \in \mathcal{K}^0} \|x - y\|^2$. Defina a função $\bar{h} : \mathcal{S} \times \mathcal{S} \to I\!R$ por $\bar{h}(x, y) = \|x - y\|^2$. Note que h é diferenciável. Como, por definição,

$$h(x) = \min\{\bar{h}(x, y) | y \in \mathcal{K}^0\} \tag{2.29}$$

e, pelo Lema 1.3.1, o mínimo do lado direito de (2.29) é unicamente atingido em $y = [x]_-$, segue de [[3], Capítulo 4, Teorema 1.7] que h é diferenciável e $\nabla h(x)$ é dada por $\nabla h(x) = \nabla_x \bar{h}(x, [x]_-)$. Assim temos $\nabla h(x) = 2(x - [x]_)) = 2[x]_+$. ■

Seja Ψ_+ a coleção de funções $\psi : \mathcal{S} \times \mathcal{S} \to [0, \infty)$, satisfazendo (2.14), que são diferenciáveis e para todo $(a, b) \in \mathcal{S} \times \mathcal{S}$ satisfazem as seguintes condições:

$$\langle \nabla_a \psi(a, b), \nabla_b \psi(a, b) \rangle \geq 0, \quad \langle a, \nabla_a \psi(a, b) \rangle + \langle b, \nabla_b \psi(a, b) \rangle \geq 0.$$

O Lema abaixo fornece uma ψ pertencendo ao Ψ_+. Esta também é uma escolha de ψ para a função de mérito dada por (2.12)-(2.14).

Lema 2.5.6 *Seja* $\psi : \mathcal{S} \times \mathcal{S} \longrightarrow [0, \infty)$ *dada por*

$$\psi(a, b) = \frac{1}{2} \left(\|[-a]_+\|^2 + \|[-b]_+\|^2 \right). \tag{2.30}$$

Então

(a) ψ *satisfaz a equação* (2.14).

(b) ψ *é convexa e diferenciável em todo* $(a, b) \in \mathcal{S} \times \mathcal{S}$ *com* $\nabla_a \psi(a, b) = -[-a]_+$ *e* $\nabla_b \psi(a, b) = -[-b]_+$.

(c) Para todo $(a, b) \in \mathcal{S} \times \mathcal{S}$*, temos* $\langle \nabla_a \psi(a, b), \nabla_b \psi(a, b) \rangle \geq 0$.

(d) Para todo $(a, b) \in \mathcal{S} \times \mathcal{S}$*, temos*

$$\langle a, \nabla_a \psi(a, b) \rangle + \langle b, \nabla_b \psi(a, b) \rangle = \|[-a]_+\|^2 + \|[-b]_+\|^2.$$

Prova.

((a) e (b)) Pelo Lema 2.5.5, a função $\|[-a]_+\|^2$ é convexa e diferenciável em $a \in \mathcal{S}$ com $\nabla_a \psi(a, b) = -2[-a]_+$, e igual a zero se, e somente se, $a \in \mathcal{K}$. Similarmente, para $\|[-b]_+\|^2$. Assim, $\psi(a, b)$ é convexa e diferenciável em (a, b) e igual a zero se, e somente se, $a, b \in \mathcal{K}$. Seja agora $\langle a, b \rangle \leq 0$. Como $a, b \in \mathcal{K}$, por (1.1), $\langle a, b \rangle \geq 0$, segue que $\langle a, b \rangle = 0$. Logo, ψ satisfaz (2.14).

((c) e (d)) Por (b) temos $\langle \nabla_a \psi(a, b), \nabla_b \psi(a, b) \rangle = \langle [-a]_+, [-b]_+ \rangle \geq 0$, onde a desigualdade segue de (1.1). E, temos

$$\begin{aligned}
\langle a, \nabla_a \psi(a, b) \rangle + \langle b, \nabla_b \psi(a, b) \rangle &= \langle a, [a]_- \rangle + \langle b, [b]_- \rangle \\
&= \|[a]_-\|^2 + \|[b]_-\|^2,
\end{aligned}$$

onde a primeira e segunda igualdades seguem do Lema 2.2.1 (a). ■

A seguir, consideramos uma restrição adicional sobre ψ. Seja Ψ_{++} a coleção de $\psi \in \Psi_+$ tal que

$$\psi(a,b) = 0 \;\; \forall \;\; (a,b) \in \mathcal{S} \times \mathcal{S} \;\; \text{com} \;\; \langle \nabla_a \psi(a,b), \nabla_b \psi(a,b) \rangle = 0. \tag{2.31}$$

Uma escolha de ψ, que serve também para a função de mérito (2.12)-(2.14), baseada na função Fischer-Burmeister dada por (2.11), pertencendo a Ψ_{++} é dada pelo Lema abaixo. Observe que a função dada por (2.30) não pertence a Ψ_{++}.

Lema 2.5.7 *Seja ψ dada por*

$$\psi(a,b) = \frac{1}{2} \left\| [\phi(a,b)]_+ \right\|^2 \tag{2.32}$$

onde ϕ é dada por (2.11). *Então*

(a) ψ satisfaz a equação (2.14).

(b) ψ é diferenciável em todo $(a,b) \in \mathcal{S} \times \mathcal{S}$. Além disso,

$$\nabla_a \psi(a,b) = sym[L_c^{-1}[[c-a-b]_+][a-c]],$$

$$\nabla_b \psi(a,b) = sym[L_c^{-1}[[c-a-b]_+][b-c]],$$

onde $c = (a^2+b^2)^{1/2}$.

(c) Para qualquer $(a,b) \in \mathcal{S} \times \mathcal{S}$, temos

$$\langle \nabla_a \psi(a,b), \nabla_b \psi(a,b) \rangle \geq \left\| (c-a-b)g \right\|^2,$$

onde $c = (a^2+b^2)^{1/2}$ e $g = L_c^{-1}[[c-a-b]_+]$. Consequentemente, ψ satisfaz a equação (2.31).

(d) Para todo $(a,b) \in \mathcal{S} \times \mathcal{S}$, temos

$$\langle a, \nabla_a \psi(a,b) \rangle + \langle b, \nabla_b \psi(a,b) \rangle = \left\| [c-a-b]_+ \right\|^2.$$

Prova.

(a) Fixe $(a,b) \in \mathcal{S}\times\mathcal{S}$ arbitrário tal que $\psi(a,b) = 0$ e $\langle a,b\rangle \leq 0$. Seja $z = -\phi(a,b)$. Então, $[-z]_+ = [\phi(a,b)]_+ = 0$, e $z \in \mathcal{K}$. Como $z = -\left(a^2+b^2\right)^{1/2} + a + b$, logo $a+b = \left(a^2+b^2\right)^{1/2} + z$, elevando ambos os lados ao quadrado e simplificando, segue que

$$ab + ba = \left(a^2+b^2\right)^{1/2} z + z\left(a^2+b^2\right)^{1/2} + z^2.$$

Tomando o traço de ambos os lados temos que

$$2\langle a,b\rangle = 2\mathrm{tr}\left[z^{1/2}\left(a^2+b^2\right)^{1/2} z^{1/2}\right] + \|z\|^2.$$

Como $z^{1/2}\left(a^2+b^2\right)^{1/2} z^{1/2} \in \mathcal{K}$, então seu traço é não-negativo e portanto $\langle a,b\rangle \geq 0$, que juntamente com $\langle a,b\rangle \leq 0$ implica que $z = 0$. Então $\phi(a,b) = 0$ e pelo Lema 2.5.2 (b) segue que $a,b \in \mathcal{K}$ e $\langle a,b\rangle = 0$. Reciprocamente, se $a,b\in\mathcal{K}$ e $\langle a,b\rangle = 0$, então, pelo Lema 2.5.2 (b), segue que $\phi(a,b) = 0$. Logo, $\psi(a,b) = 0$ e $\langle a,b\rangle \leq 0$.

(b) Fixe qualquer $(a,b)\in\mathcal{S}\times\mathcal{S}$ e defina $c = (a^2+b^2)^{1/2}$ e $r = c - a - b$. Para qualquer $u \in \mathcal{S}$, fazendo $w = au + ua + u^2$ e $r' = \left(c^2+w\right)^{1/2} - a - b - u$ temos que

$$\begin{aligned}
&|\psi(a+u,b) - \psi(a,b) - \langle [r]_+, r'-r\rangle| = \\
= \quad & \left|\frac{1}{2}\left\|\left[\sqrt{c^2+au+ua+u^2} - a - u - b\right]_+\right\|^2 - \frac{1}{2}\left\|[c-a-b]_+\right\|^2 - \langle [r]_+, r'-r\rangle\right| \\
= \quad & \left|\frac{1}{2}\left\|\left[\sqrt{c^2+w} - a - b - u\right]_+\right\|^2 - \frac{1}{2}\left\|[r]_+\right\|^2 - \langle [r]_+, r'-r\rangle\right| \\
= \quad & \left|\frac{1}{2}\left\|[r']_+\right\|^2 - \frac{1}{2}\left\|[r]_+\right\|^2 - \langle [r]_+, r'-r\rangle\right| = |\langle [r'']_+ - [r]_+, r'-r\rangle| \\
\leq \quad & \|[r'']_+ - [r]_+\|\,\|r'-r\| \leq \|r'-r\|^2,
\end{aligned} \tag{2.33}$$

onde $r'' = (1-t)r + tr'$ para algum $t\in[0,1]$; a quarta igualdade segue da diferenciabilidade de $\left\|[\cdot]_+\right\|^2$ e o teorema do valor médio, a primeira desigualdade usa a desigualdade de Cauchy e a última desigualdade usa a não-expansividade de $[\cdot]_+$.

Agora, limitaremos o lado direito. Sabemos que a equação (2.3) vale para algum $p \in \mathcal{O}$, algum $I \subset \{1, \cdots, n\}$ e alguma submatriz $\widetilde{c}_{II}$ dp, implicando que a equação (2.4) vale para determinadas submatrizes $\widetilde{a}_{II}$ e $\widetilde{b}_{II}$. Então,

$$prp^T = \begin{bmatrix} \widetilde{c}_{II} - \widetilde{a}_{II} - \widetilde{b}_{II} & 0 \\ 0 & 0 \end{bmatrix},$$

e assim,

$$p\,[r]_+\,p^T = \begin{bmatrix} \left[\widetilde{c}_{II} - \widetilde{a}_{II} - \widetilde{b}_{II}\right]_+ & 0 \\ 0 & 0 \end{bmatrix}.$$

Logo, $[r]_+ \in \mathcal{S}_c$ (veja Definição 2.2.1 (d)). Também, ao definirmos $\widetilde{w}$ e $\widetilde{z}$ pela equações (2.6), com $z = (c^2 + w)^{1/2} - c$ segue que

$$\begin{aligned}
\langle [r]_+, r' - r\rangle &= \langle [r]_+, z - u\rangle = \mathrm{tr}[p[r]_+p^Tpzp^T - [r]_+u] \\
&= \mathrm{tr}[\left[\widetilde{c}_{II} - \widetilde{a}_{II} - \widetilde{b}_{II}\right]_+ \widetilde{z}_{II} - [r]_+u] \\
&= \mathrm{tr}[\left[\widetilde{c}_{II} - \widetilde{a}_{II} - \widetilde{b}_{II}\right]_+ L^{-1}_{\widetilde{c}_{II}}\left[[p(au + ua + u^2)p^T]_{II}\right] - [r]_+u] + o\left(\|u\|\right) \\
&= \mathrm{tr}[L^{-1}_{\widetilde{c}_{II}}\left[\left[\widetilde{c}_{II} - \widetilde{a}_{II} - \widetilde{b}_{II}\right]_+\right][p(au + ua)p^T]_{II} - [r]_+u] + o\left(\|u\|\right) \\
&= \mathrm{tr}[L^{-1}_c\left[[r]_+\right](au + ua) - [r]_+u] + o\left(\|u\|\right),
\end{aligned} \tag{2.34}$$

onde a terceira igualdade usa $\widetilde{w} = pwp^T = p\,(au + ua + u^2)\,p^T$ e o Lema 2.2.3 e as duas últimas igualdades seguem do mesmo argumento usado na prova da equação (2.26). Portanto, fazendo $\widetilde{u} = pup^T$, temos das equações (2.6) e (2.4) que

$$\begin{aligned}
\widetilde{w} = \begin{bmatrix} \widetilde{w}_{II} & \widetilde{w}_{IJ} \\ \widetilde{w}_{IJ}^T & \widetilde{w}_{JJ} \end{bmatrix} &= pwp^T = pap^Tpup^T + pup^Tpap^T + pu^2p^T \\
&= \begin{bmatrix} \widetilde{a}_{II} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \widetilde{u}_{II} & \widetilde{u}_{IJ} \\ \widetilde{u}_{IJ}^T & \widetilde{u}_{JJ} \end{bmatrix} + \begin{bmatrix} \widetilde{u}_{II} & \widetilde{u}_{IJ} \\ \widetilde{u}_{IJ}^T & \widetilde{u}_{JJ} \end{bmatrix}\begin{bmatrix} \widetilde{a}_{II} & 0 \\ 0 & 0 \end{bmatrix} + pu^2p^T \\
&= \begin{bmatrix} \widetilde{a}_{II}\widetilde{u}_{II} + \widetilde{u}_{II}\widetilde{a}_{II} & \widetilde{a}_{II}\widetilde{u}_{IJ} \\ \widetilde{u}_{IJ}^T\widetilde{a}_{II} & 0 \end{bmatrix} + pu^2p^T,
\end{aligned}$$

onde $J := \{1, \cdots, n\}\backslash I$. Assim, $\widetilde{w}_{II} = \mathrm{O}(\|u\|)\,, \widetilde{w}_{IJ} = \mathrm{O}(\|u\|)\,, \widetilde{w}_{JJ} = \mathrm{O}\left(\|u\|^2\right)$ e pelo Lema 2.2.3 segue que $\widetilde{z}_{II}, \widetilde{z}_{IJ}, \widetilde{z}_{JJ}$ são todos de $\mathrm{O}(\|u\|)$ ou, equivalentemente,

$z = \mathrm{O}(\|u\|)$. Isto implica que

$$r' - r = z - u = O\left(\|u\|\right). \tag{2.35}$$

Usando as equações (2.34) e (2.35) para limitar o lado direito da equação (2.33) segue que

$$\begin{aligned}
\psi(a+u,b) - \psi(a,b) &= \operatorname{tr}\left[L_c^{-1}[[r]_+]\right](au + ua) - [r]_+u] + o\left(\|u\|\right) \\
&= \left\langle L_c^{-1}[[r]_+]a + aL_c^{-1}[[r]_+] - [r]_+, u\right\rangle + o\left(\|u\|\right),
\end{aligned}$$

implicando que

$$\begin{aligned}
\nabla_a\psi(a,b) &= L_c^{-1}[[r]_+]a + aL_c^{-1}[[r]_+] - [r]_+ \\
&= L_c^{-1}[[r]_+]a + aL_c^{-1}[[r]_+] - \left(L_c^{-1}[[r]_+]c + cL_c^{-1}[[r]_+]\right) \\
&= L_c^{-1}[[r]_+](a-c) + (a-c)L_c^{-1}[[r]_+],
\end{aligned}$$

onde a segunda igualdade usa a Definição 2.2.1 (e) e a última usa a linearidade de L_c^{-1}.

Um argumento análogo obtém-se $\nabla_b\psi(a,b)$. Logo, para quaisquer $u, v \in \mathcal{S}$,

$$\psi(a+u, b+v) - \psi(a,b) = \langle \nabla_a\psi(a,b), u\rangle + \langle \nabla_b\psi(a,b), v\rangle + o(\|u\|) + o(\|v\|).$$

Assim, ψ é diferenciável em (a,b).

(c) A prova da primeira parte é idêntica ao do Lema 2.5.3 (c) mas com $g = L_c^{-1}\left[[c-a-b]_+\right]$. Vamos mostrar que ψ satisfaz (2.31). Suponha que $\langle \nabla_a\psi(a,b), \nabla_b\psi(a,b)\rangle = 0$. Então pela primeira parte temos que $(c-a-b)g = 0$ ou equivalentemente, $rL_c^{-1}\left[[r]_+\right] = 0$, onde $r = c-a-b$. Assim,

$$\begin{aligned}
0 &= \operatorname{tr}\left[L_c^{-1}[[r]_+][r]_+ rL_c^{-1}[[r]_+]\right] = \operatorname{tr}\left[L_c^{-1}[[r]_+][r]_+^2 L_c^{-1}[[r]_+]\right] \\
&= \|[r]_+ L_c^{-1}[[r]_+]\|
\end{aligned}$$

onde a segunda igualdade segue do Lema 2.2.1 (a). Assim, $[r]_+ L_c^{-1}[[r]_+] = 0$, implicando, pelo Lema 2.2.1 (d), que $[r]_+ = 0$ e portanto segue que $\psi(a,b) = \frac{1}{2}\left\|[r]_+\right\|^2 = 0$.

(d) Por (b), temos

$$
\begin{aligned}
&\langle a, \nabla_a \psi(a,b)\rangle + \langle b, \nabla_b \psi(a,b)\rangle = \\
= &\ \operatorname{tr}\left[L_c^{-1}\left[[c-a-b]_+\right](a(a-c)+(a-c)a+b(b-c)+(b-c)b)\right] \\
= &\ \operatorname{tr}\left[L_c^{-1}\left[[c-a-b]_+\right](c(c-a-b)+(c-a-b)c)\right] \\
= &\ \operatorname{tr}\left[L_c^{-1}\left[[c-a-b]_+\right]L_c[c-a-b]\right] \\
= &\ \operatorname{tr}[[c-a-b]_+[c-a-b]] \\
= &\ \left\|[c-a-b]_+\right\|^2,
\end{aligned}
$$

onde a segunda igualdade usa que $a^2+b^2 = c^2$, a terceira igualdade segue da definição de c e L_c e as duas últimas igualdades seguem do Lema 2.2.1 (d) e (a), respectivamente. ■

A seguir apresentaremos um resultado que relaciona as funções dadas pelas equações (2.17) e (2.30) com a dada por (2.32). Para demonstrar precisamos do seguinte Lema.

Lema 2.5.8 *Para qualquer $a \in \mathcal{S}$, temos*

$$\|[a+b]_+\| \geq \|[a]_+\| \quad \forall \ \ b \in \mathcal{K}.$$

Prova. Para $a \in \mathcal{S}$ arbitrário, sejam $\lambda_1 \leq \cdots \leq \lambda_n$ seus autovalores. Para $b \in \mathcal{K}$ arbitrário, sejam $\widehat{\lambda}_1 \leq \cdots \leq \widehat{\lambda}_n$ os autovalores de $a+b$. Então, pelo Lema 1.2.6, a desigualdade $\widehat{\lambda}_i \geq \lambda_i$ vale para todo i. Logo,

$$\|[a+b]_+\|^2 = \operatorname{tr}[[a+b]_+^2] = \sum_{i:\widehat{\lambda}_i>0} \widehat{\lambda}_i^2 \geq \sum_{i:\lambda_i>0} \lambda_i^2 = \operatorname{tr}[[a]_+^2] = \|[a]_+\|^{\ 2},$$

e assim chegamos ao resultado. ■

Lema 2.5.9 *Para qualquer* $(a,b) \in \mathcal{S} \times \mathcal{S}$, *temos que*

$$\sqrt{\frac{1}{2}\|\phi(a,b)\|^2} \geq \sqrt{\frac{1}{2}\left\|[\phi(a,b)]_+\right\|^2} \geq \frac{1}{2\sqrt{2}}\left(\|[-a]_+\| + \|[-b]_+\|\right),$$

onde ϕ *é dada por* (2.11).

Prova. Primeiro vamos mostrar que a seguinte igualdade vale

$$\|x\|^2 = \|[x]_+\|^2 + \|[x]_-\|^2. \tag{2.36}$$

De fato,

$$\begin{aligned}
\|x\|^2 &= \|[x]_+ + [x]_-\|^2 \\
&= \|[x]_+\|^2 + 2\langle [x]_+, [x]_-\rangle + \|[x]_-\|^2 \\
&= \|[x]_+\|^2 + \|[x]_-\|^2,
\end{aligned}$$

onde a primeira igualdade e última igualdades seguem do Lema 2.2.1 (a).

Agora, mostraremos a primeira desigualdade. Por (2.36),

$$\begin{aligned}
\frac{1}{2}\left\|\sqrt{a^2+b^2} - a - b\right\|^2 &= \frac{1}{2}\left\|\left[\sqrt{a^2+b^2} - a - b\right]_+\right\|^2 + \frac{1}{2}\left\|\left[\sqrt{a^2+b^2} - a - b\right]_-\right\|^2 \\
&\geq \frac{1}{2}\left\|\left[\sqrt{a^2+b^2} - a - b\right]_+\right\|^2.
\end{aligned}$$

Logo,

$$\sqrt{\frac{1}{2}\|\phi(a,b)\|^2} \geq \sqrt{\frac{1}{2}\left\|[\phi(a,b)]_+\right\|^2}.$$

Provaremos a segunda desigualdade. Como $\left(\sqrt{a^2+b^2}\right)^2 - a^2 = b^2 \in \mathcal{K}$, pelo Lema 2.2.2 (c), $\sqrt{a^2+b^2} - a \in \mathcal{K}$. Logo, pelo Lema 2.5.8,

$$\left\|\left[\sqrt{a^2+b^2} - a - b\right]_+\right\|^2 \geq \|[-b]_+\|^2.$$

E assim,

$$\sqrt{\frac{1}{2}\left\|[\phi(a,b)]_+\right\|^2} \geq \frac{1}{\sqrt{2}}\|[-b]_+\|.$$

De maneira análoga, obtemos a desigualdade

$$\sqrt{\frac{1}{2}\left\|[\phi(a,b)]_+\right\|^2} \geq \frac{1}{\sqrt{2}}\|[-a]_+\|.$$

Adicionando as duas desigualdades acima, segue a segunda desigualdade. ∎

Capítulo 3

Ponto estacionário da classe de funções Luo e Tseng para o $PCSD$

3.1 Introdução

A partir deste capítulo começa a nossa contribuição. Trabalhamos com o $PCSD$ dado por (2.1), com $G = I$. Recentemente, Tseng [32] estendeu, do PCN para $PCSD$, uma classe de funções de mérito $f : \mathcal{S} \to I\!R$ definida por

$$f(x) = \psi_0(\langle x, F(x)\rangle) + \psi(x, F(x)), \tag{3.1}$$

onde $\psi_0 : I\!R \longrightarrow [0, \infty)$ é tal que

$$\psi_0(t) = 0 \ \text{ se, e somente se, } \ t \leq 0, \tag{3.2}$$

e $\psi : \mathcal{S} \times \mathcal{S} \to [0, \infty)$ é tal que

$$\psi(a, b) = 0, \ \ \langle a, b\rangle \leq 0 \ \text{ se, e somente se, } \ (a, b) \in \mathcal{K} \times \mathcal{K}, \ \langle a, b\rangle = 0. \tag{3.3}$$

Além disso, Tseng [32] mostrou que qualquer ponto estacionário de f é solução do $PCSD$. No entanto, não respondeu em que condições f tem conjuntos de nível limitados e fornece cota de erro global. Yamashita e Fukushima [38] mostraram que os conjuntos de nível de f são limitados quando F é monótona e o $PCSD$ é estritamente viável, e que f fornece cota de erro global desde que F seja fortemente monótona. Neste capítulo, descreveremos estes resultados, que são válidos desde que ψ_0 e ψ satisfaçam certas condicões. Além disso, obtemos que, sob condições mais fracas do que as de [32, 38], os conjuntos de nível de f são limitados e a propriedade de que qualquer ponto estacionário é solução do $PCSD$; apresentaremos também dois métodos de descida, um deles com a direção dada por Tseng, para resolver o $PCSD$, e mostraremos suas convergências.

3.2 Problema de otimização equivalente

O seguinte Teorema mostra que a função f definida por (3.1) é uma função de mérito sobre $\mathcal{S}$.

Teorema 3.2.1 *[32] Seja $f : \mathcal{S} \longrightarrow I\!R$ dada por* (3.1) *com ψ_0 e ψ verificando* (3.2) *e* (3.3)*, respectivamente. Então, f é não-negativa. Além disso, $f(x) = 0$ se, e somente se, x é solução do $PCSD$.*

Prova. A não-negatividade segue de (3.1) e da hipótese de ψ_0 e ψ.

Suponha que $f(x) = 0$. Então, pela não-negatividade de ψ_0 e ψ, $\psi_0(\langle x, F(x)\rangle) = 0$ e $\psi(x, F(x)) = 0$. Assim, por definição de ψ_0, temos que $\langle x, F(x)\rangle \leq 0$, e por definição de ψ temos que x é solução do $PCSD$. A recíproca é direta. ■

A seguir daremos alguns exemplos de ψ_0 e ψ.

Exemplo 3.2.0.1 *Para $\alpha > 0$, seja $\psi_0^\alpha : I\!R \longrightarrow [0, \infty)$ definida por*

$$\psi_0^\alpha(t) = \frac{1}{\alpha}\left(\max\{0, t\}\right)^\alpha,$$

e sejam ψ_1 e $\psi_2 : \mathcal{S} \times \mathcal{S} \longrightarrow [0, \infty)$ definidas por

$$\psi_1(a, b) = \frac{1}{2}\left(\|[-a]_+\|^2 + \|[-b]_+\|^2\right)$$

$$\psi_2(a,b) \;=\; \frac{1}{2}\|[\sqrt{a^2+b^2}-a-b]_+\|^2.$$

É óbvio que ψ_0^α satisfaz (3.2). Pelos Lemas 2.5.6 (a) e 2.5.7 (a) temos que ψ_i, $i=1,2$ satisfazem (3.3). Além disso, ψ_0^α é diferenciável para $\alpha > 1$ e pelos Lemas 2.5.6 (b) e 2.5.7 (b) as funções $\psi_i, i=1,2$ são diferenciáveis, respectivamente.

3.3 Cota de erro

Nesta seção daremos condições sob as quais f fornece uma cota de erro para o *PCSD*. O termo $\psi_0(\langle x, F(x)\rangle)$ é a chave para determinar a cota de erro. Mostraremos que se F é fortemente monótona e ψ_0 e ψ satisfazem certas condições, a função f fornece uma cota de erro global. Para isso definimos a seguinte função

$$f_1(x) = \max\{0, \langle x, F(x)\rangle\} + \|[-x]_+\| + \|[-F(x)]_+\|\,.$$

Yamashita e Fukushima [38] mostraram que f_1 fornece uma cota de erro global para o *PCSD* sob a hipótese de F ser fortemente monótona. A seguir demonstraremos este resultado.

Lema 3.3.1 [*38*] *Seja F fortemente monótona com módulo μ. Então existe uma constante $\alpha > 0$ tal que*

$$\alpha\left\|x-\overline{x}\right\|^2 \le f_1(x), \qquad \forall\ x \in \mathcal{S}, \tag{3.4}$$

onde $\overline{x}$ é a única solução do PCSD.

Prova. Como F é fortemente monótona temos que

$$\begin{aligned}\mu\left\|x-\overline{x}\right\|^2 &\le \langle x-\overline{x}, F(x)-F(\overline{x})\rangle \\ &= \langle x, F(x)\rangle + \langle \overline{x}, -F(x)\rangle + \langle -x, F(\overline{x})\rangle\end{aligned}$$

$$
\begin{aligned}
&\leq \max\{0, \langle x, F(x)\rangle\} + \left\langle \overline{x}, [-F(x)]_+ \right\rangle + \left\langle [-x]_+, F(\overline{x}) \right\rangle \\
&\leq \max\{0, \langle x, F(x)\rangle\} + \|\overline{x}\| \, \|[-F(x)]_+\| + \|F(\overline{x})\| \, \|[-x]_+\| \\
&\leq \max\{1, \|\overline{x}\|, \|F(\overline{x})\|\} \left[\max\{0, \langle x, F(x)\rangle\} + \|[-x]_+\| + \|[-F(x)]_+\|\right],
\end{aligned}
$$

onde a primeira igualdade segue do fato de $\overline{x}$ ser solução do $PCSD$ e a segunda desigualdade segue do Lema 2.2.1 (e). Fazendo

$$\alpha := \mu / \max\{1, \|\overline{x}\|, \|F(\overline{x})\|\}$$

obtemos (3.4). ■

A seguir mostraremos uma relação entre $f_1(x)$ e f, tal que, sob certas condições para ψ_0 e ψ, e considerando F fortemente monótona, permite que f forneça cota de erro global para o $PCSD$.

Teorema 3.3.1 *Seja F fortemente monótona com módulo μ e $\alpha > 0$. Suponha que*

$$\psi_0(t) \geq \beta_1 (\max\{0, t\})^\alpha \ \forall\ t, \tag{3.5}$$

e

$$\psi(a, b) \geq \beta_2 \left(\|[-a]_+\| + \|[-b]_+\|\right)^2 \ \forall\ (a, b) \in \mathcal{S} \times \mathcal{S}, \tag{3.6}$$

onde β_1 e β_2 são constantes positivas. Então existe um $\beta > 0$ tal que

$$f_1(x) \leq \beta \left(f(x)^{1/\alpha} + f(x)^{1/2}\right) \ \forall \quad x \in \mathcal{S}.$$

Além disso, a função f dada por (3.1) *fornece uma cota de erro global para o $PCSD$.*

Prova. Pela hipótese sobre ψ_0 e ψ temos que existem constantes positivas k_1 e k_2 tais que

$$\max\{0, \langle x, F(x)\rangle\} \leq k_1 \psi_0^{1/\alpha}(\langle x, F(x)\rangle)$$

e

$$\|[-x]_+\| + \|[-F(x)]_+\| \leq k_2\psi^{1/2}(x, F(x)).$$

Seja $\beta = \max\{k_1, k_2\}$, então

$$\begin{aligned} f_1(x) &\leq \beta\left(\psi_0^{1/\alpha}(\langle x, F(x)\rangle) + \psi_1^{1/2}(x, F(x))\right) \\ &\leq \beta\left(f(x)^{1/\alpha} + f(x)^{1/2}\right) \quad \text{para qualquer } \ x \in \mathcal{S}. \end{aligned}$$

Portanto, pelo Lema 3.3.1, f fornece uma cota de erro global para o $PCSD$. ■

Pela definição das funções dadas no Exemplo 3.2.0.1 temos que ψ_0^α, com $\alpha > 0$, satisfaz (3.5) e ψ_1 satisfaz (3.6). Além disso, temos, pelo Lema 2.5.9, que ψ_2 satisfaz (3.6).

3.4 Conjunto de nível limitado

Nesta seção, mostraremos que os conjuntos de nível de f são limitados sob certas condições. Do Teorema 3.3.1 obtemos o seguinte resultado sobre a limitação dos conjuntos de nível de f sob a condição de F ser fortemente monótona.

Corolário 3.4.1 *Seja f definida por* (3.1). *Se F é diferenciável e fortemente monótona, então os conjuntos de nível, $\mathcal{L}(\gamma) = \{x \in \mathcal{S}/f(x) \leq \gamma\}$, são não-vazios e limitados para todo $\gamma \geq 0$.*

Para o $PCSD$, Yamashita e Fukushima [38] mostraram que a função definida por (2.15)-(2.17) tem conjuntos de nível limitados se F é monótona e o $PCSD$ é estritamente viável, ou seja, existe uma matriz definida positiva $\widehat{x} \in \mathcal{S}$ tal que $F(\widehat{x})$ é definida positiva. Sob as mesmas condições acima e desde que ψ_0 e ψ satisfaçam (3.5) e (3.6), respectivamente, podemos mostrar que $\mathcal{L}(\gamma)$ é limitado. Para isso precisamos do seguinte resultado, demonstrado em [38].

Lema 3.4.1 *Para qualquer sequência* $\left\{\left(a^k, b^k\right)\right\} \subseteq \mathcal{S} \times \mathcal{S}$, *considere* $\lambda_1^k \leq \lambda_2^k \leq \cdots \leq \lambda_n^k$ *e* $\mu_1^k \leq \mu_2^k \leq \cdots \leq \mu_n^k$ *os autovalores de* a^k *e* b^k, *respectivamente. Suponha que* $\left\{\lambda_1^k\right\}$ *e* $\left\{\mu_1^k\right\}$ *são limitados inferiormente. Se* $\lambda_n^k \to \infty$, *então, para matrizes definidas positivas* $g, h \in \mathcal{S}$, *temos que*

$$\left\langle g, a^k \right\rangle + \left\langle h, b^k \right\rangle \to \infty.$$

Prova. Seja p^k uma matriz ortogonal tal que $p^k a^k \left(p^k\right)^T = diag\left[\lambda_1^k, \cdots, \lambda_n^k\right]$ e seja $g^k := p^k g \left(p^k\right)^T$. Então

$$\begin{aligned} \left\langle g, a^k \right\rangle &= \operatorname{tr}[g a^k] \\ &= \operatorname{tr}\left[p^k g \left(p^k\right)^T p^k a^k \left(p^k\right)^T\right] \\ &= \operatorname{tr}\left[g^k diag\{\lambda_1^k, \cdots, \lambda_n^k\}\right] \\ &= \sum_{i=1}^{n} \lambda_i^k g_{ii}^k. \end{aligned} \tag{3.7}$$

Note que, pelo Lema 1.2.5 (a), g^k é definida positiva, e portanto os elementos da diagonal g_{ii}^k são positivos. Além disso, temos que para cada i, $\liminf_{k\to\infty} g_{ii}^k > 0$. De fato, suponha que existe um índice i e uma subsequência $\left\{g_{ii}^k\right\}_{k\in\kappa}$ convergindo a zero. Como a sequência $\left\{p^k\right\}$ de matrizes ortogonais é limitada, segue que, tomando uma subsequência, se necessário, a sequência $\left\{p^k\right\}_{k\in\kappa}$ converge para um limite, digamos p^*, logo $\left\{g^k\right\}_{k\in\kappa}$ converge a $g^* = p^* g \left(p^*\right)^T$. Como p^* é também uma matriz ortogonal, então, pelo Lema 1.2.5 (a), g^* é definida positiva. No entanto isto contradiz a hipótese de g_{ii}^* ser zero. Como $\left\{\lambda_i^k\right\}$ está limitada inferiormente para cada i e $\lambda_n^k \to \infty$, segue de (3.7) que

$$\left\langle g, a^k \right\rangle \to \infty.$$

De forma análoga, podemos mostrar que $\left\{\left\langle h, b^k\right\rangle\right\}$ está limitada inferiormente. ■

Agora, podemos mostrar que os conjuntos de nível da função f definida por (3.1) são limitados.

Teorema 3.4.1 [38] *Suponha que F é monótona e o PCSD é estritamente viável. Seja f definida por* (3.1) *com ψ_0 e ψ satisfazendo* (3.5) *e* (3.6)*, respectivamente. Então os conjuntos de nível de f, $\mathcal{L}(\gamma)$, são limitados para todo $\gamma \geq 0$.*

Prova. Assuma que existe uma sequência ilimitada $\{x^k\} \subseteq \mathcal{L}(\gamma)$ para algum $\gamma \geq 0$. Temos que as sequências dos autovalores mínimos de $\{x^k\}$ e $\{F(x^k)\}$ são limitadas inferiormente. De fato, caso contrário, segue de (3.6) que $\psi\left(x^k, F(x^k)\right) \to \infty$ e assim $f(x^k) \to \infty$, o que contradiz o fato que $\{x^k\} \subseteq \mathcal{L}(\gamma)$. Portanto, a ilimitação de $\{x^k\}$ implica que a sequência do máximo dos autovalores de $\{x^k\}$ tende a infinito. Como o *PCSD* é estritamente viável, existe $\widehat{x} \succ 0$ tal que $F(\widehat{x}) \succ 0$, assim pela monotonicidade de F temos que

$$\left\langle x^k, F(\widehat{x})\right\rangle + \left\langle \widehat{x}, F(x^k)\right\rangle \leq \left\langle x^k, F(x^k)\right\rangle + \langle \widehat{x}, F(\widehat{x})\rangle . \tag{3.8}$$

Logo, pelo Lema 3.4.1, o lado esquerdo de (3.8) tende para infinito. Portanto, $\langle x^k, F(x^k)\rangle \to \infty$ e por (3.5) temos que $\psi_0\left(\langle x^k, F(x^k)\rangle\right) \to \infty$. Assim, $f(x^k) \to \infty$, o que contradiz o fato que $\{x^k\} \subseteq \mathcal{L}(\gamma)$. ■

A seguir, apresentaremos outra condição que garante a limitação dos conjuntos de nível de f. Para isso introduzimos uma classe de função tipo-R_0.

Definição 3.4.1 *Dizemos que a função $F : \mathcal{S} \to \mathcal{S}$ é uma função-R_{01}, se para qualquer sequência $\{x^k\}$ satisfazendo*

$$\|x^k\| \longrightarrow \infty, \frac{[-x^k]_+}{\|x^k\|} \longrightarrow 0, \frac{[-F(x^k)]_+}{\|x^k\|} \longrightarrow 0 \tag{3.9}$$

temos que

$$\liminf_{k \longrightarrow \infty} \frac{\langle x^k, F(x^k) \rangle}{\|x^k\|} > 0. \tag{3.10}$$

Esta classe de funções é uma extensão de classe utilizada no *PCN* [5, 8, 27]. Para o *PCSD*, esta condição foi usada por Chen e Tseng [7]. A seguir mostraremos que se F é monótona e o *PCSD* é estritamente viável então F é uma função-R_{01}.

Proposição 3.4.1 *Suponha que F é monótona e o PCSD é estritamente viável. Então F é uma função-R_{01}.*

Prova. Considere $\{x^k\}$ uma sequência satisfazendo (3.9) e tome $\{x^k\}_{k \in \kappa}$ uma subsequência tal que

$$\lim_{k \in \kappa} \frac{x^k}{\|x^k\|} = w.$$

Claramente, por (3.9) e pelo Lema 2.2.2 (a), temos que

$$w \neq 0 \text{ e } w \succeq 0. \tag{3.11}$$

Como o *PCSD* é estritamente viável, existe uma matriz definida positiva $\overline{x}$ tal que $F(\overline{x}) \succ 0$. Além disso, pela monotonicidade de F, temos que

$$\langle x^k, F(x^k) \rangle \geq \langle x^k, F(\overline{x}) \rangle + \langle \overline{x}, F(x^k) \rangle - \langle \overline{x}, F(\overline{x}) \rangle.$$

Dividindo ambos os lados da desigualdade acima por $\|x^k\|$ e tomando o limite sobre a subsequência κ, temos que

$$\begin{aligned} \liminf_{k \in \kappa} \frac{\langle x^k, F(x^k) \rangle}{\|x^k\|} &\geq \langle w, F(\overline{x}) \rangle + \langle \overline{x}, \liminf_{k \in \kappa} \frac{F(x^k)}{\|x^k\|} \rangle \\ &\geq \langle w, F(\overline{x}) \rangle > 0, \end{aligned}$$

onde a segunda desigualdade segue da condição (3.9), Lema 2.2.2 (a) e $\overline{x} \succ 0$, e a terceira segue de (3.11) e o fato que $F(\overline{x}) \succ 0$. Portanto, por definição, F é uma função-R_{01}. ■

O Teorema a seguir é uma extensão do resultado obtido por Luo e Tseng para o PCN [[27], Teorema 4.1].

Teorema 3.4.2 *Suponha que F é uma função-R_{01} e considere a função f definida por* (3.1) *tal que ψ_0 e ψ satisfazem* (3.5) *e* (3.6), *respectivamente. Então os conjuntos de nível de f, $\mathcal{L}(\gamma)$, são limitados para todo $\gamma \geq 0$.*

Prova. Suponha que para algum $\gamma \geq 0$, o conjunto de nível $\mathcal{L}(\gamma)$ é ilimitado. Então existe uma sequência ilimitada $\{x^k\} \subseteq \mathcal{L}(\gamma)$ tal que $\left\|x^k\right\| \longrightarrow \infty$. Pela condição (3.6) temos que as sequências dos autovalores mínimos de $\{x^k\}$ e $\{F(x^k)\}$ são limitadas inferiormente. Portanto,

$$\frac{[-x^k]_+}{\|x^k\|} \longrightarrow 0, \frac{[-F(x^k)]_+}{\|x^k\|} \longrightarrow 0. \tag{3.12}$$

Agora, como $f(x^k) \leq \gamma$ e ψ_0 satisfaz (3.5) temos que existe $\beta > 0$ tal que

$$\langle x^k, F(x^k)\rangle \leq \beta.$$

Então, dividindo ambos os lados da desigualdade acima por $\left\|x^k\right\|$ e fazendo $k \longrightarrow \infty$, obtemos

$$\liminf_{k \longrightarrow \infty} \frac{\langle x^k, F(x^k)\rangle}{\|x^k\|} \leq 0. \tag{3.13}$$

Entretanto, (3.12) e (3.13) contradizem a hipótese de F ser uma função-R_{01}. Logo $\mathcal{L}(\gamma)$ são limitados para todo $\gamma \geq 0$.. ■

3.5 Ponto estacionário

Como a função de mérito f não é convexa, exceto quando ψ_0 e ψ são convexas e F é afim e monótona [32], é importante saber quando um ponto estacionário de f é um mínimo global de f. Para isso definimos as seguintes classes de funções diferenciáveis.

Definição 3.5.1 *Seja* Ψ_0 *a classe das funções diferenciáveis* $\psi_0 : I\!R \longrightarrow [0, \infty)$ *satisfazendo*

$$\psi_0(t) = 0 \iff t \leq 0,$$

tais que

(C1) $t\nabla\psi_0(t) \geq 0$ *para todo* $t \in I\!R$;

(C2) $\nabla\psi_0(t) \geq 0$ *para todo* $t \in I\!R$;

(C3) $\nabla\psi_0(t) = 0 \iff t \leq 0.$

Definição 3.5.2 *Seja* Ψ *a classe das funções continuamente diferenciáveis* $\psi : \mathcal{S} \times \mathcal{S} \longrightarrow [0, \infty)$ *satisfazendo*

$$\psi(a, b) = 0, \langle a, b\rangle \leq 0 \iff (a, b) \in \mathcal{K} \times \mathcal{K},\ \langle a, b\rangle = 0.$$

tais que

$(\widetilde{C}1)$ $\langle a, \nabla_a\psi(a, b)\rangle + \langle b, \nabla_b\psi(a, b)\rangle \geq 0$ *para todo* $(a, b) \in \mathcal{S} \times \mathcal{S}$;

$(\widetilde{C}2)$ $\nabla_a\psi(a, b) \preceq 0, \quad \nabla_b\psi(a, b)) \preceq 0,$ *para todo* $(a, b) \in \mathcal{S} \times \mathcal{S}$;

Seja Ψ_+ *a classe das funções* $\psi \in \Psi$ *tal que*

$(\widetilde{C}3)$ $\nabla_a\psi(a, b) = 0, \quad \nabla_b\psi(a, b) = 0 \iff (a, b) \in \mathcal{K} \times \mathcal{K}.$

Além disso, seja Ψ_{++} *a classe das funções* $\psi \in \Psi$ *tais que*

$(\widetilde{C}4)$ $\langle\nabla_a\psi(a, b), \nabla_b\psi(a, b)\rangle = 0 \Longrightarrow (a, b) \in \mathcal{K} \times \mathcal{K}.$

$(\widetilde{C}5)$ $(a, b) \in \mathcal{K} \times \mathcal{K}, \langle a, b\rangle = 0 \Rightarrow \nabla_a\psi(a, b) = 0, \nabla_b\psi(a, b) = 0.$

Do Exemplo 3.2.0.1, a função ψ_0^α para $\alpha > 1$ pertence a Ψ_0, e pelo Lema 2.5.6 (b) e (d), $\psi_1 \in \Psi_+$. A seguir mostraremos que $\psi_2 \in \Psi_{++}$.

Lema 3.5.1 *A função ψ_2, dada no Exemplo 3.2.0.1, pertence a Ψ_{++}.*

Prova. $(\tilde{C}1)$ segue do Lema 2.5.7 (d).

$(\tilde{C}2)$ Sejam $c = \sqrt{a^2 + b^2}$, $g = [c - a - b]_+$ e $x = c - a$. Pelo Lema 2.2.2 (b), x é sdp, implicando, pelo Lema 1.2.5, que $(L_c^{-1}[g])^{1/2} x (L_c^{-1}[g])^{1/2}$ é sdp. Mas, $L_c^{-1}[g]x$ é similar a esta matriz, então é sdp, e portanto $xL_c^{-1}[g]$ também é. Assim, pelo Lema 2.5.7 (b), $\nabla_a \psi_2(a, b) \preceq 0$. De forma análoga, provamos que $\nabla_b \psi_2(a, b) \preceq 0$.

$(\tilde{C}4)$ Suponha que $\langle \nabla_a \psi_2(a, b), \nabla_b \psi_2(a, b) \rangle = 0$, implicando, pelo Lema 2.5.7 (c), que $\psi_2(a, b) = 0$. Portanto, pelo Lema 2.5.9, $a \in \mathcal{K}$ e $b \in \mathcal{K}$.

$(\tilde{C}5)$ Suponha que $(a, b) \in \mathcal{K} \times \mathcal{K}$ e $\langle a, b \rangle = 0$. Então, pelo Lema 2.5.2 (c), segue que $\sqrt{a^2 + b^2} - a - b = 0$. Logo, pelo Lema 2.5.7 (b), $\nabla_a \psi_2(a, b) = 0$ e $\nabla_b \psi_2(a, b) = 0$.

■

Seja $\phi : \mathcal{S} \times \mathcal{S} \to I\!R$ definida por

$$\phi(a, b) = \psi_0(\langle a, b \rangle) + \psi(a, b) \tag{3.14}$$

Note que, pela definição da ϕ e as condições sobre ψ_0 e ψ, temos o seguinte resultado.

$$\phi(a, b) = 0 \;\text{ se, e somente, se }\; (a, b) \in \mathcal{K} \times \mathcal{K}, \langle a, b \rangle = 0. \tag{3.15}$$

Além disso, como ϕ é a soma das funções continuamente diferenciáveis $\psi_0 \in \Psi_0$ e $\psi \in \Psi_+$ ou Ψ_{++}, então ϕ é diferenciável e

$$\nabla_a \phi(a, b) = \nabla \psi_0(\langle a, b \rangle) b + \nabla_a \psi(a, b), \tag{3.16}$$

$$\nabla_b \phi(a, b) = \nabla \psi_0(\langle a, b \rangle) a + \nabla_b \psi(a, b). \tag{3.17}$$

O Teorema seguinte mostra algumas propriedades para ϕ.

Lema 3.5.2 *Seja ϕ definida por (3.14) com $\psi_0 \in \Psi_0$ e $\psi \in \Psi$. Então, ϕ tem as seguintes propriedades:*

($\widetilde{P}$1) $\langle \nabla_a \phi(a,b), \nabla_b \phi(a,b) \rangle \geq 0$, *para todo $a \in \mathcal{S}$ e $b \in \mathcal{S}$.*

Além disso, se $\psi \in \Psi_+$ ($\psi \in \Psi_{++}$) temos que

($\widetilde{P}$2) $\nabla_a \phi(a,b) = 0, \quad \nabla_b \phi(a,b) = 0 \Longleftrightarrow \phi(a,b) = 0.$

E se $\psi \in \Psi_{++}$, temos

($\widetilde{P}$3) $\nabla_a \phi(a,b) = 0 \Longleftrightarrow \nabla_b \phi(a,b) = 0.$

Prova.

($\widetilde{P}$1) Por (C1) e (C2), temos

$$\nabla \psi_0^2(\langle a,b \rangle)\langle a,b \rangle = \nabla \psi_0(\langle a,b \rangle)\left(\nabla \psi_0(\langle a,b \rangle)\langle a,b \rangle\right) \geq 0. \tag{3.18}$$

Temos também por ($\widetilde{C}$2) e por (1.1)

$$\langle \nabla_a \psi(a,b), \nabla_b \psi(a,b) \rangle \geq 0. \tag{3.19}$$

Além disso, por (C2) e ($\widetilde{C}$1),

$$\nabla \psi_0(\langle a,b \rangle)\left(\langle a, \nabla_a \psi(a,b) \rangle + \langle b, \nabla_b \psi(a,b) \rangle\right) \geq 0. \tag{3.20}$$

De (3.16) e (3.17), temos

$$\begin{aligned} \langle \nabla_a \phi(a,b), \nabla_b \phi(a,b) \rangle &= \nabla \psi_0^2(\langle a,b \rangle)\langle a,b \rangle + \langle \nabla_a \psi(a,b), \nabla_b \psi(a,b) \rangle \\ &+ \nabla \psi_0(\langle a,b \rangle)\left(\langle a, \nabla_a \psi(a,b) \rangle + \langle b, \nabla_b \psi(a,b) \rangle\right). \end{aligned} \tag{3.21}$$

Logo, de (3.18) a (3.20) provamos ($\widetilde{P}$1).

($\widetilde{P}$2) Primeiro, analisamos para o caso onde $\psi \in \Psi_+$. Suponha que $\nabla_a \phi(a,b) = 0, \nabla_b \phi(a,b) = 0$. Então, por (3.16) e (3.17),

$$\nabla \psi_0(\langle a,b \rangle)b = -\nabla_a \psi(a,b), \tag{3.22}$$

$$\nabla\psi_0(\langle a,b\rangle)a = -\nabla_b\psi(a,b). \tag{3.23}$$

Suponha que ambos os lados de (3.22) e (3.23) são não-nulos. Então, por (C2) $\nabla\psi_0(\langle a,b\rangle) > 0$ e por $(\tilde{C}2)$ $-\nabla_a\psi(a,b) \succeq 0, -\nabla_b\psi(a,b) \succeq 0$ e não-nulos. Portanto, segue das equações (3.22) e (3.23) que $a \succeq 0$ e $b \succeq 0$, implicando por $(\tilde{C}3)$ que $\nabla_a\psi(a,b) = 0$ e $\nabla_b\psi(a,b) = 0$, o que é uma contradição. Portanto, ambos os lados das equações (3.22) e (3.23) são nulos. Então do lado esquerdo das igualdades temos que $\nabla\psi_0(\langle a,b\rangle) = 0$ ou $b = 0$ ou $a = 0$, de onde segue de qualquer maneira que $\langle a,b\rangle \leq 0$ (usando (C3) se necessário). Agora, do lado direito das igualdades temos que $\nabla_a\psi(a,b) = 0$ e $\nabla_b\psi(a,b) = 0$, de onde segue por $(\tilde{C}3)$ que $(a,b) \in \mathcal{K}\times\mathcal{K}$, implicando por (1.1) que $\langle a,b\rangle \geq 0$. Portanto, as condições acima implicam que $(a,b) \in \mathcal{K}\times\mathcal{K}, \langle a,b\rangle = 0$, e assim, por (3.15), $\phi(a,b) = 0$. Reciprocamente, suponha que $\phi(a,b) = 0$, então, por (3.15)

$$a,b \in \mathcal{K} \quad \text{e} \quad \langle a,b\rangle = 0.$$

Da igualdade temos que $\langle a,b\rangle \leq 0$, e por (C3)

$$\nabla\psi_0(\langle a,b\rangle) = 0. \tag{3.24}$$

Além disso, de $a,b \in \mathcal{K}$ temos que, por $(\widetilde{C}3)$

$$\nabla_a\psi(a,b) = 0, \quad \nabla_b\psi(a,b) = 0. \tag{3.25}$$

Logo, de (3.16),(3.17), (3.24) e (3.25) temos que

$$\nabla_a\phi(a,b) = 0, \quad \nabla_b\phi(a,b) = 0.$$

Agora, analisamos para o caso onde $\psi \in \Psi_{++}$. Suponha que $\nabla_a\phi(a,b) = 0, \nabla_b\phi(a,b) = 0$. Logo, $\langle\nabla_a\phi(a,b), \nabla_b\phi(a,b)\rangle = 0$. Portanto por (3.18)-(3.20), cada termo do lado direito de (3.21) é nulo. Em particular, o primeiro termo dado por (3.18) sendo zero implica, por (C3), que $\langle a, b\rangle \leq 0$. Além disso, temos que o segundo termo de (3.21) é nulo, implicando, por $(\tilde{C}4)$, que $a, b \in \mathcal{K}$, e de (1.1) segue que $\langle a, b\rangle \geq 0$. Assim, $a, b \in \mathcal{K}$ e $\langle a, b\rangle = 0$. Logo, por (3.15), $\phi(a, b) = 0$. Portanto, obtemos o seguinte resultado

$$\langle \nabla_a \phi(a, b), \nabla_b \phi(a, b)\rangle = 0 \Longrightarrow \phi(a, b) = 0. \tag{3.26}$$

A recíproca é análoga ao caso anterior, mas neste caso usamos $(\tilde{C}5)$ para provar (3.25).

$(\tilde{P}3)$ Suponha que $\nabla_a \phi(a, b) = 0$. Logo, $\langle \nabla_a \phi(a, b), \nabla_b \phi(a, b)\rangle = 0$. Portanto, por (3.26) e $(\tilde{P}2)$, segue que $\nabla_b \phi(a, b) = 0$. A recíproca é idêntica. ∎

A seguir daremos uma condição necessária e suficiente para que um ponto estacionário de f seja solução do $PCSD$. Tseng [32] mostrou que, sob a condição da matriz jacobiana $\nabla F(x)$ ser sdp, qualquer ponto estacionário da função de Fischer-Burmeister dada por (2.10)-(2.11) e da função f é uma solução do $PCSD$. Da mesma forma Yamashita e Fukushima [38] mostraram para a função dada por (2.15)-(2.17). Aqui, daremos uma condição de regularidade que enfraquece a hipótese acima. Com este propósito, definimos os seguintes conjuntos de índices

$$\begin{aligned}
\mathcal{P}(x) &:= \{i | [p \nabla_b \phi(x, F(x)) p^T]_{ii} > 0\}, \\
\mathcal{C}(x) &:= \{i | [p \nabla_b \phi(x, F(x)) p^T]_{ii} = 0\}, \\
\mathcal{N}(x) &:= \{i | [p \nabla_b \phi(x, F(x)) p^T]_{ii} < 0\}, \\
\mathcal{R}(x) &:= \mathcal{P}(x) \cup \mathcal{N}(x),
\end{aligned}$$

onde $x \in \mathcal{S}$ e $p \in \mathcal{O}$.

Para simplificar denotaremos estes conjuntos por $\mathcal{P}, \mathcal{C}, \mathcal{N}$.

Usando estes conjuntos de índices, definimos condições de regularidade, que são extensões do PCN [25].

Definição 3.5.3 *Uma matriz $x \in \mathcal{S}$ é dita ser regular se, para quaisquer $z \in \mathcal{S}$, com $z \neq 0$, e $p \in \mathcal{O}$, tal que*

$$[(pzp^T)_{ii}]_{\mathcal{C}} = 0, \quad [(pzp^T)_{ii}]_{\mathcal{P}} > 0, \quad [(pzp^T)_{ii}]_{\mathcal{N}} < 0, \tag{3.27}$$

existe uma matriz $y \in \mathcal{S}$ tal que

$$[(pyp^T)_{ii}]_{\mathcal{P}} \geq 0, \quad [(pyp^T)_{ii}]_{\mathcal{N}} \leq 0, \quad [(pyp^T)_{ii}]_{\mathcal{R}} \neq 0, \tag{3.28}$$

$$\langle y, \nabla F(x) z \rangle \geq 0. \tag{3.29}$$

Além disso, uma matriz $x \in \mathcal{S}$ é dita ser estritamente regular se, para quaisquer $z \in \mathcal{S}$, com $z \neq 0$, e $p \in \mathcal{O}$, tal que

$$[(pzp^T)_{ii}]_{\mathcal{C}} = 0, \quad [(pzp^T)_{ii}]_{\mathcal{P}} > 0, \quad [(pzp^T)_{ii}]_{\mathcal{N}} < 0,$$

existe uma matriz $y \in \mathcal{S}$ tal que

$$[(pyp^T)_{ii}]_{\mathcal{C}} = 0, \quad [(pyp^T)_{ii}]_{\mathcal{P}} \geq 0, \quad [(pyp^T)_{ii}]_{\mathcal{N}} \leq 0,$$

$$\langle y, \nabla F(x) z \rangle > 0.$$

Observe que se $\nabla F(x)$ é sdp (dp), então podemos escolher $y = z$ e obtemos da Definição 3.5.3 que x é regular (estritamente regular). O seguinte Teorema mostra que a (estrita) regularidade fornece uma condição necessária e suficiente para que um ponto estacionário de f seja solução do $PCSD$.

Teorema 3.5.1 *Seja $f : \mathcal{S} \longrightarrow I\!\!R$ dada por* (3.1) *com ϕ dada por* (3.14). *Então* *(a) Se ψ_0 e ψ são continuamente diferenciáveis e F é continuamente diferenciável, então f é continuamente diferenciável e*

$$\nabla f(x) = \nabla F(x)\nabla_b\phi(x, F(x)) + \nabla_a\phi(x, F(x)), \tag{3.30}$$

para todo $x \in \mathcal{S}$.

(b) Suponha que F é continuamente diferenciável sobre $\mathcal{S}$ e $\psi_0 \in \Psi_0$ e $\psi \in \Psi_{++}$ (respectivamente, $\psi \in \Psi_+$). Suponha que x é um ponto estacionário de f. Então, x é solução do $PCSD$, se e somente se, x é regular (respectivamente, estritamente regular).

Prova. (a) Se F é continuamente diferenciável, então da definição de f, da hipótese de ψ_0 e ψ, e da regra da cadeia, temos que f é continuamente diferenciável, e de (3.16) e (3.17)

$$\begin{aligned}\nabla f(x) &= \nabla\psi_0\left(\langle x, F(x)\rangle\right)\left(\nabla F(x)x + F(x)\right)\\ &- \nabla F(x)\nabla_b\psi(x, F(x)) - \nabla_a\psi(x, F(x))\\ &= \nabla F(x)\nabla_b\phi(x, F(x)) + \nabla_a\phi(x, F(x),\end{aligned}$$

para todo $x \in \mathcal{S}$.

(b) Suponha que $x \in \mathcal{S}$ é solução do $PCSD$, então, por $(\tilde{P}2)$ temos que $\nabla_b\phi(x, F(x)) = 0$, logo $\mathcal{P} = \mathcal{N} = \emptyset$, e portanto não existe uma matriz não-nula z satisfazendo (3.27). Assim, x é regular. Suponha que x é regular e $\nabla f(x) = 0$. Então, por (3.30), temos que

$$\nabla F(x)\nabla_b\phi(x, F(x)) + \nabla_a\phi(x, F(x)) = 0.$$

Consequentemente, temos

$$\langle y, \nabla_a\phi(a, b)\rangle + \langle y, \nabla F(x)\nabla_b\phi(a, b)\rangle = 0, \tag{3.31}$$

para todo $y \in \mathcal{S}$.

Suponha que x não é solução do $PCSD$. Então, pelas propriedades $(\tilde{P}2)$ e $(\tilde{P}3)$ do Lema 3.5.2, $z := \nabla_b\phi(x, F(x))$ é uma matriz não-nula e satisfaz (3.27). Como x é regular, existe $y \in \mathcal{S}$ que satisfaz (3.28). Além disso, seja $q \in \mathcal{O}$ tal que $q\nabla_a\phi(x, F(x))q^T = diag\{\lambda_1, \ldots, \lambda_n\}$, onde λ_i, $i = 1, \ldots, n$ são os autovalores de $\nabla_a\phi(x, F(x))$. Logo, por $(\tilde{P}1)$, (3.27) e pelo Lema 1.2.2, para todo i, temos que $\lambda_i(\nabla_a\phi(x, F(x)))$ e $[pzp^T]_{ii}$ têm os mesmos sinais. Portanto, por (3.27) e (3.28), $\lambda_i(\nabla_a\phi(x, F(x)))$ e $[pyp^T]_{ii}$ também têm os mesmos sinais. E asssim, usando o fato que $[(pyp^T)_{ii}]_{\mathcal{R}} \neq 0$, obtemos que

$$\langle y, \nabla_a\phi(x, F(x))\rangle > 0, \tag{3.32}$$

e de (3.29), segue que

$$\langle y, \nabla F(x)\nabla_b\phi(x, F(x))\rangle = \langle y, \nabla F(x)z\rangle \geq 0. \tag{3.33}$$

Portanto, as desigualdades (3.32) e (3.33) juntas, contradizem a condição (3.31). Logo, x é solução do $PCSD$.

De forma análoga, mostra-se para o caso onde x é estritamente regular. ∎

Agora apresentaremos uma condição suficiente para que uma matriz $x \in \mathcal{S}$ seja (estritamente) regular.

Lema 3.5.3 *Seja $x \in \mathcal{S}$. Se $\nabla F(x)$ é uma P_0- matriz ($P-$matriz), então x é uma matriz (estritamente) regular.*

Prova. Somente consideremos o caso onde $\nabla F(x)$ é uma P_0-matriz. A prova é análoga para o caso onde $\nabla F(x)$ é $P-$matriz. Suponha que x não é uma matriz regular, logo, por definição $\nabla F(x)$ não é sdp. Então existe $y \neq 0$ tal que

$\langle y, \nabla F(x)y\rangle < 0$. Logo, pelo Lema 1.2.2, existe $p \in \mathcal{O}$ tal que todos os elementos da diagonal de $py\nabla F(x)yp^T$ são não-positivos, contradizendo assim a hipótese. ∎

Do Lema 3.5.3, obtemos o seguinte corolário, que garante condição suficiente para que um ponto estacionário de f seja solução do $PCSD$.

Corolário 3.5.1 *Suponha que F é diferenciável sobre $\mathcal{S}$. Seja ϕ definida por* (3.14) *com $\psi_0 \in \Psi_0$ e $\psi \in \Psi_{++}$ (respectivamente, $\psi \in \Psi_{+}$), e f definida por* (3.1). *Se F é uma função-P (respectivamente, $\nabla F(x)$ é uma P-matriz para qualquer x) então qualquer ponto estacionário de f é uma solução do $PCSD$.*

Prova. Suponha que F é uma função-P, então, pelo Lema 2.4.1, temos que $\nabla F(x)$ é uma P_0- matriz e portanto, pelo Lema 3.5.3, Teorema 3.5.1 e usando o fato que x é ponto estacionário de f, temos que x é solução do $PCSD$. A segunda afirmação segue de forma análoga. ∎

Como uma aplicação dos resultados anteriores, obtemos o seguinte corolário que garante a existência de solução do $PCSD$.

Corolário 3.5.2 *Suponha que F é diferenciável sobre $\mathcal{S}$. Seja ϕ definida por* (3.14) *com ψ_0 e ψ verificando* (3.2) *e* (3.3)*, respectivamente, e f definida por* (3.1). *Suponha que F é monótona e que o $PCSD$ é estritamente viável. Então, o conjunto solução do $PCSD$ é não-vazio e limitado.*

Prova. Escolha um $\overline{x} \in \mathcal{S}$ arbitrário e seja $\overline{\gamma} := f(\overline{x})$. Então $L(\overline{\gamma})$ é não-vazio e, como mostrado pelo Teorema 3.4.1, limitado. Portanto $\mathcal{L}(\overline{\gamma})$ contém um minimizador local $\tilde{x}$ de f. Em particular, $\tilde{x}$ é um ponto estacionário de f. Como F é monótona, $\nabla F(\tilde{x})$ é sdp e portanto é uma P_0-matriz. Logo, pelo Corolário 3.5.1, $\tilde{x}$ é solução do $PCSD$.

A limitação do conjunto de solução segue do fato que o conjunto de nível $\mathcal{L}(0)$ é o conjunto solução do $PCSD$, e é limitado pelo Teorema 3.4.1. ■

3.6 Método de Descida

Nesta seção, apresentamos dois métodos de descida para resolver o $PCSD$ e mostraremos suas convergências.

O método proposto a seguir é baseado na direção de busca dada por Tseng[32]:

$$d(x) = -\nabla_b \phi\left(x, F(x)\right). \tag{3.34}$$

Como $d(x)$ não envolve a matriz jacobiana $\nabla F(x)$, o método é considerado uma extensão dos métodos de derivada livre desenvolvidos para o PCN [23, 37].

O próximo lema mostra que, quando $\nabla F(x)$ é semidefinida positiva, a matriz $d(x)$ é uma direção de descida de f desde que x não seja solução do problema.

Lema 3.6.1 *Suponha que F é diferenciável sobre $\mathcal{S}$. Seja ϕ definida por* (3.14) *com $\psi_0 \in \Psi_0$ e $\psi \in \Psi_{++}$, e f definida por* (3.1). *Suponha que x não é solução do $PCSD$. Se $\nabla F(x)$ é semidefinida positiva, então a matriz $d(x) \in \mathcal{S}$ definida por* (3.34) *satisfaz a condição de descida $\langle \nabla f(x), d(x) \rangle < 0$.*

Prova. Por (3.30) e (3.34) temos que

$$\begin{aligned} \langle \nabla f(x), d(x) \rangle &= -\left\langle \nabla_a \phi(x, F(x)), \nabla_b \phi(x, F(x)) \right\rangle \\ &- \left\langle \nabla_b \phi(x, F(x)), \nabla F(x) \nabla_b \phi(x, F(x)) \right\rangle. \end{aligned} \tag{3.35}$$

Mas por $(\tilde{P}1)$ mostrada no Lema 3.5.2, o primeiro termo é não-positivo. Como $\nabla F(x)$ é sdp, o segundo termo é também não-positivo. Logo,

$$\langle \nabla f(x), d(x) \rangle \leq 0.$$

Assuma que $\langle \nabla f(x), d(x) \rangle = 0$. Isto implica que ambos os termos do lado direito de (3.35) são-nulos. Em particular, temos que

$$\langle \nabla_a \phi(x, F(x)), \nabla_b \phi(x, F(x)) \rangle = 0.$$

E, por (3.26) e (3.15), temos que x é solução do $PCSD$, o que contradiz a hipótese. Assim,

$$\langle \nabla f(x), d(x) \rangle < 0,$$

ou seja, $d(x)$ é uma direção de descida. ■

Agora apresentamos o método de descida que usa a direção de busca $d(x)$.

Algoritmo 3.6.1 *(P0) Escolha* $x^0 \in \mathcal{S}$, $\epsilon \geq 0$, $\sigma \in (0, 1/2)$, $\beta \in (0,1)$ *e seja* $k := 0$.

(P1) Se $f(x^k) \leq \epsilon$, *então pare.*

(P2) Faça $d(x^k) := -\nabla_b \phi\left(x^k, F(x^k)\right)$.

(P3) Calcule um comprimento de passo $t_k := \beta^{m_k}$, *onde* m_k *é o menor inteiro não-negativo* m *satisfazendo a condição tipo Armijo*

$$f(x^k + \beta^m d(x^k)) \leq \left(1 - \sigma\beta^{2m}\right) f(x^k). \tag{3.36}$$

(P4) Defina $x^{k+1} := x^k + t_k d(x^k)$, $k := k+1$, *e volte para o (P1).*

Agora mostraremos a convergência do Algoritmo. Supomos que $\epsilon = 0$ tal que o algoritmo gere uma sequência infinita, isto é, x^k não é uma solução do $PCSD$ para cada k. A demonstração é análoga a de [[38], Teorema 5.1]. Para completar daremos a demonstração.

Teorema 3.6.1 *Seja ϕ definida por* (3.14) *com $\psi_0 \in \Psi_0$ e $\psi \in \Psi_{++}$ e f definida por* (3.1). *Suponha que F é diferenciável e que uma das seguintes condições ocorre: ou F é fortemente monótona ou F é monótona e o $PCSD$ é estritamente viável. Então a sequência gerada pelo Algoritmo tem pelo menos um ponto de acumulação e qualquer ponto de acumulação é solução do $PCSD$.*

Prova. Supondo que x^k não é uma solução do $PCSD$ mostraremos que existe um inteiro não-negativo m_k satisfazendo (3.36). Assuma o contrário. Então para qualquer inteiro positivo m temos

$$f(x^k + \beta^m d(x^k)) - f(x^k) > -\sigma\beta^{2m} f(x^k). \tag{3.37}$$

Dividindo ambos os lados por β^m e fazendo $m \longrightarrow \infty$ segue que

$$\left\langle \nabla f(x^k), d(x^k) \right\rangle \geq 0. \tag{3.38}$$

Como $\nabla F(x)$ é sdp sob as hipóteses dadas temos que (3.38) contradiz o Lema 3.6.1. Logo podemos encontrar um inteiro m_k satisfazendo (3.36).

A seguir mostraremos que a sequência gerada pelo Algoritmo tem no mínimo um ponto de acumulação. Pela propriedade de descida do Algoritmo, a sequência $\left\{ f(x^k) \right\}$ é monótona decrescente. Portanto, pelo Corolário 3.4.1 ou Teorema 3.4.1, a sequência gerada é limitada, e logo tem pelo menos um ponto de acumulação.

Provaremos agora a última parte do Teorema. Seja x^* um ponto de acumulação arbitrário da sequência gerada $\left\{ x^k \right\}$, e $\left\{ x^k \right\}_{k \in \kappa}$ uma subsequência convergindo a x^*. Como $d(\cdot)$ é contínua, $\left\{ d(x^k) \right\}_{k \in \kappa}$ converge para $d(x^*)$. Primeiro considere o caso onde existe uma constante $\overline{\beta}$ tal que $\beta^{m_k} \geq \overline{\beta} > 0$ para todo $k \in \kappa$. Então

$$f(x^{k+1}) \leq (1 - \sigma\overline{\beta}^2) f(x^k) \tag{3.39}$$

para todo $k \in \kappa$ e a sequência $\left\{f(x^k)\right\}$ é decrescente, e temos assim $f\left(x^*\right) = 0$. Isto implica que, pelo Teorema 3.2.1, x^* é uma solução do *PCSD*. Agora, considere o caso onde existe uma subsequência tal que $\beta^{m_k} \longrightarrow 0$. Note que da regra de Armijo em (3.36), temos que

$$f(x^k + \beta^{m_k - 1} d(x^k)) - f(x^k) > -\sigma\beta^{2\left(m_k - 1\right)} f(x^k). \tag{3.40}$$

Dividindo ambos os lados por $\beta^{m_k - 1}$ e fazendo $m_k \longrightarrow \infty$ segue que

$$\left\langle \nabla f(x^k), d(x^k) \right\rangle \geq 0. \tag{3.41}$$

Em vista do Lema 3.6.1, deduzimos que x^* é solução do *PCSD*. ∎

O Teorema acima demonstrou em particular que o algoritmo proposto é válido para a função de mérito f com $\psi_0 \in \Psi_0$ e $\psi \in \Psi_{++}$. Para a função de mérito f com $\psi_0 \in \Psi_0$ e $\psi \in \Psi_+$, é possível mostrar que $d(x^k) \to 0$ quando $k \to \infty$, tal que $\nabla_b \phi(x^*, F(x^*)) = 0$ vale para qualquer ponto de acumulação da sequência gerada. No entanto, isto não necessariamente implica que x^* é solução do *PCSD* (conforme Lema 3.5.1). Para suprir esta dificuldade, podemos regularizar a direção de busca para

$$d(x) = -\nabla_b \phi\left(x, F(x)\right) - \rho \nabla_a \phi\left(x, F(x)\right), \tag{3.42}$$

onde ρ é uma constante positiva suficientemente pequena. Observe que $d(x)$ dada acima é uma variante da direção dada por (3.34).

O próximo resultado mostra que, sob a condição de $F(x)$ ser fortemente monótona, a matriz $d(x)$, definida por (3.42), é uma direção de descida.

Lema 3.6.2 *Seja ϕ definida por* (3.14) *com $\psi_0 \in \Psi_0$ e $\psi \in \Psi_+$ e f definida por* (3.1). *Suponha que F é continuamente diferenciável e fortemente monótona com* *módulo μ. Seja $x^0 \in S$. Então, existe uma constante positiva η tal que, para cada x no conjunto de nível*

$$\mathcal{L}(x^0) = \{x | f(x) \leq f(x^0)\},$$

$d(x)$, definida por (3.42), satisfaz

$$\langle \nabla f(x), d(x) \rangle \leq -\eta \left(\|\nabla_b \phi(x, F(x)\| + \rho \|\nabla_a \phi(x, F(x)\| \right)^2,$$

onde ρ é uma constante positiva suficientemente pequena. Além disso, suponha que x não é solução do PCSD, então $d(x)$ é uma direção de descida.

Prova. Por (3.30) e (3.42) temos que

$$\begin{aligned} \langle \nabla f(x), d(x) \rangle &= - \langle \nabla_a \phi(x, F(x)), \nabla_b \phi(x, F(x)) \rangle \\ &- \langle \nabla_b \phi(x, F(x)), \nabla F(x) \nabla_b \phi(x, F(x)) \rangle \\ &- \rho \langle \nabla_a \phi(x, F(x)), \nabla F(x) \nabla_b \phi(x, F(x)) \rangle - \rho \|\nabla_a \phi(x, F(x))\|^2. \end{aligned} \tag{3.43}$$

Como F é fortemente monótona com módulo μ, pela Proposição 1.4.1 (c),

$$\langle \nabla_b \phi(x, F(x)), \nabla F(x) \nabla_b \phi(x, F(x)) \rangle \geq \mu \|\nabla_b \phi(x, F(x))\|^2. \tag{3.44}$$

Observe que, pelo Teorema 3.4.1, o conjunto de nível $\mathcal{L}(x^0)$ é limitado. Como $\nabla F(x)$ é contínua, existe uma constante $\tau > 0$ tal que

$$\|\nabla F(x)\| \leq \tau, \quad \text{para todo} \quad x \in \mathcal{L}(x^0).$$

Logo,

$$\begin{aligned} &- \langle \nabla_a \phi(x, F(x)), \nabla F(x) \nabla_b \phi(x, F(x)) \rangle \\ &\leq \tau \|\nabla_a \phi(x, F(x))\| \|\nabla_b \phi(x, F(x))\| \end{aligned} \tag{3.45}$$

para todo $x \in \mathcal{L}(x^0)$.

Portanto, de (3.43)-(3.45) e da propriedade $(\widetilde{P}1)$ do Lema 3.5.2 que

$$\begin{aligned} \langle \nabla f(x), d(x) \rangle &\leq \tau\rho \|\nabla_a \phi(x, F(x))\| \|\nabla_b \phi(x, F(x))\| \\ &- \mu \|\nabla_b \phi(x, F(x))\|^2 - \rho \|\nabla_a \phi(x, F(x))\|^2. \end{aligned} \tag{3.46}$$

Mas o lado direito de (3.46) pode ser reescrito como

$$\begin{aligned} &\tau\rho \|\nabla_a \phi(x, F(x))\| \|\nabla_b \phi(x, F(x))\| - \mu \|\nabla_b \phi(x, F(x))\|^2 - \rho \|\nabla_a \phi(x, F(x))\|^2 = \\ &= -\eta \left(\|\nabla_b \phi(x, F(x))\| + \rho \|\nabla_a \phi(x, F(x))\| \right)^2 \\ &- 1/2 \left(\sqrt{\mu} \|\nabla_b \phi(x, F(x))\| - \sqrt{\rho} \|\nabla_a \phi(x, F(x))\| \right)^2 \\ &+ (\eta - \mu/2) \|\nabla_b \phi(x, F(x))\|^2 + (\eta\rho - 1/2)\rho \|\nabla_a \phi(x, F(x))\|^2 \\ &+ (2\eta\rho + \tau\rho - \sqrt{\rho\mu}) \|\nabla_a \phi(x, F(x))\| \|\nabla_b \phi(x, F(x))\|. \end{aligned} \tag{3.47}$$

Tomando ρ suficientemente pequeno tal que

$$\rho \leq \mu/4\tau^2.$$

Logo,

$$\tau\rho - \sqrt{\rho\mu} \leq -\sqrt{\rho\mu}/2.$$

E assim, o último termo de (3.47) satisfaz a desigualdade

$$\begin{aligned} &(2\eta\rho + \tau\rho - \sqrt{\rho\mu}) \|\nabla_a \psi(x, F(x))\| \|\nabla_b \psi(x, F(x))\| \\ &\leq (2\eta\rho - \sqrt{\rho\mu}/2) \|\nabla_a \psi(x, F(x))\| \|\nabla_b \psi(x, F(x))\|. \end{aligned} \tag{3.48}$$

Além disso, se η satisfaz

$$\eta \leq \min\{\mu/2, 1/2\rho, \sqrt{\mu}/4\sqrt{\rho}\},$$

então temos

$$\eta - \mu/2 \leq 0, \quad \eta\rho - 1/2 \leq 0, \quad 2\eta\rho - \sqrt{\rho\mu}/2 \leq 0. \tag{3.49}$$

Portanto, segue de (3.46), (3.47), (3.48), e (3.49) que

$$\langle \nabla f(x), d(x) \rangle \leq -\eta \left(\|\nabla_b \phi(x, F(x))\| + \rho \|\nabla_a \phi(x, F(x))\| \right)^2.$$

Para provar a segunda parte, asssuma que $\langle \nabla f(x), d(x) \rangle = 0$. Isto implica que $\nabla_a \phi(x, F(x)) = 0, \nabla_b \phi(x, F(x)) = 0$ e, pelo Lema 3.5.2 ($\tilde{P}2$), x é solução do $PCSD$, contradizendo a hipótese. Logo $d(x)$ é uma direção de descida. ■

Agora apresentaremos o método de descida que usa $d(x)$ definida por (3.42).

Algoritmo 3.6.2

(P0) Escolha $x^0 \in \mathcal{S}$, $\epsilon \geq 0$, $\rho > 0{,}0 < \sigma < \eta$, $\beta \in (0, 1)$ *e seja* $k := 0$.

(P1) Se $f(x^k) \leq \epsilon$, *então pare.*

(P2) Faça $d(x^k) := -\nabla_b \phi \left(x^k, F(x^k) \right) - \rho \nabla_a \phi \left(x^k, F(x^k) \right)$.

(P3) Calcule um comprimento de passo $t_k := \beta^{m_k}$, *onde* m_k *é o menor inteiro não-negativo* m *satisfazendo a condição tipo Armijo*

$$f(x^k + \beta^m d(x^k)) \leq f(x^k) - \sigma \beta^m \left(\|\nabla_b \phi(x^k, F(x^k))\| + \rho \|\nabla_a \phi(x^k, F(x^k))\| \right)^2. \quad (3.50)$$

(P4) Defina $x^{k+1} := x^k + t_k d(x^k)$, $k := k + 1$, *e volte para o (P1).*

O próximo resultado segue do Corolário 3.4.1 e do Lema 3.6.2.

Teorema 3.6.2 *Seja* ϕ *definida por* (3.14) *com* $\psi_0 \in \Psi_0$ *e* $\psi \in \Psi_+$ *e* f *definida por* (3.1). *Suponha que* F *é continuamente diferenciável e fortemente monótona com módulo* μ. *Suponha também que* $F(x)$ *e* $\nabla F(x)$ *são Lipschitz contínuas. Então a sequência gerada pelo Algoritmo 3.6.2converge para a solução do* $PCSD$.

Prova. Como $F(x)$ e $\nabla F(x)$ são Lipschitz contínuas, $\nabla f(x)$ é Lipschitz contínua sobre qualquer conjunto limitado S, i. e., existe uma constante $\Gamma > 0$ tal que

$$\|\nabla f(x) - \nabla f(y)\| \leq \Gamma \|x - y\|, \quad \text{para todo } \; x, y \in S. \tag{3.51}$$

Por (3.50) temos que

$$f(x^{k+1}) \leq f(x^k) - \sigma t_k \left(\|\nabla_b \phi(x^k, F(x^k))\| + \rho \|\nabla_a \phi(x^k, F(x^k))\| \right)^2, \tag{3.52}$$

implicando que a sequência $\{f(x^k)\}$ é monótona decrescente. Logo, pelo Corolário 3.4.1, $\{x^k\}$ é limitada. Além disso, pela continuidade de F, a limitação de $\{x^k\}$ implica a limitação de $\{d(x^k)\}$. Assim, podemos supor que $\{x^k\}$ e $\{x^k + d(x^k)\}$ estão contidos num conjunto limitado S. Então, para $0 \leq t \leq 1$,

$$\begin{aligned}
& f(x^k + td(x^k)) - f(x^k) = \int_0^t \langle \nabla f(x^k + sd(x^k)), d(x^k) \rangle ds \\
= \; & t \langle \nabla f(x^k), d(x^k) \rangle + \int_0^t \langle \nabla f(x^k + sd(x^k)) - \nabla f(x^k), d(x^k) \rangle ds \\
\leq \; & -\eta t \left(\|\nabla_b \phi(x^k, F(x^k))\| + \rho \|\nabla_a \phi(x^k, F(x^k))\| \right)^2 + \int_0^t \Gamma s \|d(x^k)\|^2 ds \\
\leq \; & \left(-\eta t + \frac{1}{2} \Gamma t^2 \right) \left(\|\nabla_b \phi(x^k, F(x^k))\| + \rho \|\nabla_a \phi(x^k, F(x^k))\| \right)^2,
\end{aligned} \tag{3.53}$$

onde a primeira desigualdade segue do Lema 3.6.2 e (3.51), e a última desigualdade segue de

$$\|d(x^k)\| \leq \|\nabla_b \phi(x^k, F(x^k))\| + \rho \|\nabla_a \phi(x^k, F(x^k))\|.$$

De (3.53), podemos deduzir que a desigualdade

$$f(x^k + td(x^k)) \leq f(x^k) - \sigma t \left(\|\nabla_b \phi(x^k, F(x^k))\| + \rho |\nabla_a \phi(x^k, F(x^k))\| \right)^2$$

vale para qualquer t satisfazendo

$$0 \leq t \leq min\{1, 2(\eta - \sigma)/\Gamma\}.$$

Consequentemente, da forma como o comprimento do passo t_k é determinado, segue que $t_k = \beta^{m_k}$ satisfaz

$$t_k > min\{\beta, 2\beta(\eta - \sigma)/\Gamma\}, \tag{3.54}$$

para todo k, isto é, $\{t_k\}$ é limitado inferiormente. Como $\{f(x^k)\}$ é limitada inferiormente, temos que $\{f(x^k)\}$ converge. Logo, de (3.52) e (3.54) temos que

$$\lim_{k\to\infty} \{\|\nabla_b\phi(x^k, F(x^k))\| + \rho\|\nabla_a\phi(x^k, F(x^k))\|\} = 0.$$

Além disso, como $\{x^k\}$ é limitada, tem pelo menos um ponto de acumulação. Seja x^* um ponto de acumulação da sequência $\{x^k\}$. Então,

$$\|\nabla_a\phi(x^*, F(x^*))\| = 0 \quad \text{e} \quad \|\nabla_b\phi(x^*, F(x^*))\| = 0.$$

E, portanto, pelo Lema 3.5.2 ($\tilde{P}2$), $\{x^*\}$ é solução do $PCSD$. Como F é fortemente monótona temos que a sequência $\{x^k\}$ converge para a solução do $PCSD$. ■

Capítulo 4

Uma nova classe de funções de mérito para o $PCSD$ e algumas propriedades

4.1 Introdução

Neste capítulo, também definimos o $PCSD$ dado por (2.1), com $G = I$. Estamos interessados em estabelecer uma equivalência entre o $PCSD$ e um problema de minimização

$$\min_{x \in \mathcal{S}} f(x). \tag{4.1}$$

Vimos, no Capítulo 2 (seção 2.3), que uma função que constitui um problema de minimização equivalente ao $PCSD$ é chamada de função de mérito. Em particular, a classe de funções definida abaixo serve para construir uma função de mérito.

Definição 4.1.1 *Dizemos que a aplicação $\phi : \mathcal{S} \times \mathcal{S} \longrightarrow \mathcal{S}$ é uma função-$PCSD$ se satisfaz a seguinte condição*

$$\phi(a, b) = 0 \text{ se, e somente se, } a, b \in \mathcal{K} \text{ e } \langle a, b \rangle = 0.$$

Observe que, na Definição acima, a imagem de ϕ é uma matriz nula. Usando uma função-$PCSD$, ϕ, podemos construir uma função de mérito $f : \mathcal{S} \to I\!R$ para o $PCSD$ como segue

$$f(x) = \frac{1}{2}\|\phi(x, F(x))\|^2.$$

A seguir daremos dois exemplos de função-$PCSD$ que foram estendidas do PCN por Tseng [32], veja Lemas 2.2.2 (b) e 2.5.2 (b).

Exemplo 4.1.0.2

(i) $\phi_{NR}(a, b) = a - [a - b]_+$.

(ii) $\phi_{FB}(a, b) = \sqrt{a^2 + b^2} - a - b$.

A seguir daremos uma demonstração de que a função ϕ_{FB} é uma função-$PCSD$ de uma forma mais simples do que a dada por Tseng, veja Lema (2.5.2) (b).

Lema 4.1.1 *A função ϕ_{FB} é uma função-$PCSD$.*

Prova. Suponha que $\phi_{FB}(a, b) = 0$, logo $\sqrt{a^2 + b^2} = a + b$. Elevando ao quadrado ambos os lados temos que $ab + ba = 0$. Portanto,

$$\langle a, b \rangle = 0. \tag{4.2}$$

Além disso, também implica que

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

e portanto,

$$0 = \sqrt{a^2 + b^2} - a - b = |a - b| - a - b = 2[b - a]_+ - 2b,$$

onde a última igualdade segue do Lema 2.2.2 (a). Então, $[b - a]_+ - b = 0$. Mas, pela primeira expressão do Lema 2.2.1 (a), $[a - b]_+ - a = 0$. Assim, $a, b \in \mathcal{K}$. E, por (4.2), $\langle a, b \rangle = 0$. Reciprocamente, suponha que $a, b \in \mathcal{K}$ e $\langle a, b \rangle = 0$. Pelo

Lema 1.2.5 (c) temos que $ab = 0$ e $ba = 0$. Portanto $ab + ba = 0$, implicando que $\sqrt{(a+b)^2} = \sqrt{a^2+b^2}$. Como $a, b \in \mathcal{K}$ temos que $\sqrt{(a+b)^2} = |a+b| = a+b$, e assim $\phi_{FB}(a,b) = 0$. ■

A seguir, consideramos a classe de funções vista no Capítulo 2, seção 2.5, que serve também para construir uma função de mérito.

Definição 4.1.2 *Dizemos que a aplicação $\psi : \mathcal{S} \times \mathcal{S} \longrightarrow [0, \infty)$ é uma função-$PCSD^*$ se satisfaz a seguinte condição*

$$\psi(a,b) = 0 \text{ se, e somente se, } a, b \in \mathcal{K} \quad e \quad \langle a, b \rangle = 0.$$

Note, agora, que na Definição acima a imagem de ψ é um número real nulo, e não uma matriz nula como na Definição 4.1.1. Usando uma função-$PCSD^*$, ψ, temos uma função de mérito $f : \mathcal{S} \to [0, \infty)$ para o $PCSD$ como segue

$$f(x) = \psi(x, F(x)).$$

A seguir daremos alguns exemplos de funções-$PCSD^*$ que foram estendidas por Tseng [32], veja Lemas 2.2.2 (b), 2.5.3 (a) e 2.5.1 (a), respectivamente.

Exemplo 4.1.0.3

(i) $\psi_{NR}(a,b) = ||\phi_{NR}(a,b)||^2$.

(ii) $\psi_{FB}(a,b) = ||\phi_{FB}(a,b)||^2$.

(iii) $\psi_{MS}(a,b) = \langle a, b \rangle - \langle b, [a - \alpha b]_+ \rangle - \frac{1}{2\alpha}||a - [a - \alpha b]_+||^2$, *onde* $\alpha > 0$.

Observe que, no Exemplo acima (i) e (ii), as funções obtidas a partir da função-$PCSD$ são funções-$PCSD^*$.

Para o PCN, Kanzow et al [25] propuseram uma função de mérito que fornece uma cota de erro global sob a condição de F ser $P-$uniforme, veja (2.19). Naturalmente, esperamos que exista uma função de mérito para o $PCSD$ que tenha esta propriedade. Com este objetivo, propomos a seguinte classe de funções para o $PCSD$:

$$g(x) = \phi(x, F(x)), \tag{4.3}$$

com $\phi : \mathcal{S} \times \mathcal{S} \longrightarrow I\!R$ definida por

$$\phi(a, b) = \psi_{0^*}(\text{sym}[ab]) + \psi(a, b), \tag{4.4}$$

onde $\psi_{0^*} : \mathcal{S} \longrightarrow [0, \infty)$ é contínua tal que

$$\psi_{0^*}(t) = 0 \text{ se, e somente se, } t \preceq 0, \tag{4.5}$$

e $\psi : \mathcal{S} \times \mathcal{S} \longrightarrow [0, \infty)$ é contínua tal que

$$\psi(a, b) = 0,\ \langle a, b\rangle \leq 0 \quad \text{se, e somente se,} \quad (a, b) \in \mathcal{K} \times \mathcal{K},\ \langle a, b\rangle = 0. \tag{4.6}$$

Esta classe de funções é extensão do PCN [25].

Este capítulo é organizado como segue. Na segunda seção, mostramos que a nova classe de funções constitui um problema de minimização, dado por (4.1), equivalente ao $PCSD$. Na seção 3, provamos que g fornece cota de erro global sob a condição de F ser P-uniforme. Na seção 4, mostramos que os conjuntos de nível de g são limitados sob certas condições.

4.2 Problema de otimização equivalente

Nesta seção, mostramos que o problema de otimização irrestrito constituído por esta função é equivalente ao $PCSD$.

O seguinte Lema mostra que a função dada por (4.3)-(4.6) é uma função-$PCSD^*$.

Lema 4.2.1 *Para quaisquer ψ_{0^*} satisfazendo* (4.5) *e ψ satisfazendo* (4.6), *a função ϕ dada pela equação* (4.4) *é uma função-PCSD* não-negativa.*

Prova. Como ψ_{0^*} e ψ são funções não-negativas, por definição, então ϕ é não-negativa. Agora, vamos provar que ϕ é uma função-$PCSD^*$. Suponha que $\phi(a,b) = 0$. Como ψ_{0^*} e ψ são não-negativas temos que $\psi_{0^*}(sym[ab]) = 0$ e $\psi(a,b) = 0$. Mas, pela definição de ψ_{0^*} temos que $sym[ab] \preceq 0$, o que implica, pelo Lema 1.2.5 (b), $\langle a,b\rangle \leq 0$, e portanto, pela definição de ψ, temos que $a, b \in \mathcal{K}$, $\langle a,b\rangle = 0$. Reciprocamente, sejam $a, b \in \mathcal{K}$ e $\langle a,b\rangle = 0$. Pelo Lema 1.2.5 (c), temos que $ab = 0$ e $ba = 0$. Logo, $sym[ab] = 0$ e portanto pela definição de ψ_{0^*} temos que $\psi_{0^*}(sym[ab]) = 0$. Agora, pela definição de ψ, $\psi(a,b) = 0$. Assim, $\phi(a,b) = 0$. Portanto ϕ é uma função-$PCSD^*$. ■

O próximo teorema estabelece a equivalência entre o $PCSD$ e o problema de minimização dado pela equação (4.1) onde f é definida por (4.3) a (4.6).

Teorema 4.2.1 *Seja $\phi : \mathcal{S} \times \mathcal{S} \longrightarrow I\!R$ definida por* (4.4) *com ψ_{0^*} e ψ verificando* (4.5) *e* (4.6), *respectivamente, e $f : \mathcal{S} \longrightarrow I\!R$ dada por* (4.3). *Então, f é não-negativa. Além disso, $f(x) = 0$ se, e somente se, x é solução do PCSD.*

Prova. Segue diretamente do Lema 4.2.1. ■

A seguir daremos alguns exemplos de ψ_{0^*} e ψ.

Exemplo 4.2.0.4 . *Seja $\psi_{0^*} : \mathcal{S} \longrightarrow [0,\infty)$ definida por*

$$\psi_{0^*}(t) = \frac{1}{2}\|[t]_+\|^2, \tag{4.7}$$

e sejam ψ_1 e $\psi_2 : \mathcal{S} \times \mathcal{S} \longrightarrow [0,\infty)$ definidas por

$$\psi_1(a,b) \;=\; \frac{1}{2}\left(\|[-a]_+\|^2 + \|[-b]_+\|^2\right),$$

$$\psi_2(a,b) \;=\; \frac{1}{2}\left(\|[\sqrt{a^2+b^2}-a-b]_+\|^2\right).$$

É claro que ψ_{0^*} satisfaz (4.5). Observe que $\psi_i, i = 1,2$ são definidas como no Capítulo 3, Exemplo 3.2.0.1, satisfazem (4.6) e são diferenciáveis. O próximo resultado mostra que ψ_{0^*} é diferenciável.

Lema 4.2.2 *A função ψ_{0^*}, dada no Exemplo 4.2.0.4, é diferenciável e suas derivadas parciais são:*

$$\nabla_a \psi_{0^*}(sym[ab]) = sym\left[[sym[ab]]_+ b\right] \quad e \quad \nabla_b \psi_{0^*}(sym[ab]) = sym\left[[sym[ab]]_+ a\right].$$

Prova. Fixe $(a,b) \in \mathcal{S} \times \mathcal{S}$ arbitrário e tome $r = sym[ab]$ e $r' = sym[(a+u)b]$. Para qualquer $u \in \mathcal{S}$, temos que

$$\begin{aligned}
& |\psi_{0^*}(r') - \psi_{0^*}(r) - \langle [r]_+, r' - r\rangle| \\
= \; & \left|\frac{1}{2}\|[r']_+\|^2 - \frac{1}{2}\|[r]_+\|^2 - \langle [r]_+, r'-r\rangle\right| \\
= \; & |\langle [r'']_+ - [r]_+, r'-r\rangle| \\
\leq \; & \|[r'']_+ - [r]_+\|\|r'-r\| \leq \|r'-r\|^2, \qquad (4.8)
\end{aligned}$$

onde $r'' = (1-t)r + tr'$ para algum $t \in [0,1]$; a segunda igualdade segue da diferenciabilidade de $\|[\cdot]_+\|^2$ e do teorema do valor médio, a primeira desigualdade usa Cauchy e a última usa a não-expansividade de $[\cdot]_+$. Mas

$$r' - r = bu + ub = O\left(\|u\|\right), \qquad (4.9)$$

então, da primeira igualdade de (4.9) temos

$$\begin{aligned}
\langle [r]_+, r'-r\rangle \;&=\; \langle [sym[ab]]_+, bu + ub\rangle \\
&=\; \langle [sym[ab]]_+ b + b[sym[ab]]_+, u\rangle \\
&=\; \langle sym\left[[sym[ab]]_+ b\right], u\rangle. \qquad (4.10)
\end{aligned}$$

Portanto, por (4.8), (4.9) e (4.10) temos que

$$|\psi_{0^*}(sym[(a+u)b]) - \psi_{0^*}(sym[ab])| = \langle sym\,[[sym[ab]]_+b]\,, u\rangle + o(\|u\|),$$

implicando que

$$\nabla_a\psi_{0^*}(sym[ab]) = sym\,[[sym[ab]]_+b]\,.$$

De forma análoga obtém-se que

$$\nabla_b\psi_{0^*}(sym[ab]) = sym\,[[sym[ab]]_+a]\,.$$

E para quaisquer $u, v \in \mathcal{S}$,

$$\begin{aligned} &\psi_{0^*}(sym[(a+u)(b+v)]) - \psi_{0^*}(sym[ab]) \\ = \ & \langle\nabla_a\psi_{0^*}(sym[ab]), u\rangle + \langle\nabla_b\psi_{0^*}(sym[ab]), v\rangle + o(\|u\|) + o(\|v\|). \end{aligned}$$

Assim, ψ_{0^*} é diferenciável. ■

Usando as funções dadas no Exemplo 4.2.0.4, apresentamos alguns exemplos de funções-$PCSD^*$. Sejam $\phi_i,\ i = 1,2 : \mathcal{S}\times\mathcal{S} \longrightarrow [0,\infty)$ definidas por

$$\begin{aligned} \phi_1(a,b) &= \psi_{0^*}(sym[ab]) + \psi_1(a,b), \\ \phi_2(a,b) &= \psi_{0^*}(sym[ab]) + \psi_2(a,b), \end{aligned}$$

A funções $\phi_i,\ i = 1,2$ são diferenciáveis. Observe que temos dois casos especiais de g, ou seja

$$g_i(x) = \phi_i(x, F(x)), i = 1, 2. \tag{4.11}$$

4.3 Cota de erro

Nesta seção daremos condições sob as quais g fornece uma cota de erro para o $PCSD$. O termo $\psi_{0^*}(sym[xF(x)])$ é a chave para achar a cota de erro. No Capítulo

3, mostramos que f definida por (3.1)-(3.2) fornece cota de erro global se F é fortemente monótona. Aqui, mostraremos que a função de mérito g definida por (4.3) -(4.6) fornece uma cota de erro global para o $PCSD$ sob condição mais fraca, ou seja, quando F é uma função $P-$uniforme.

Seja

$$h(x) = \|[sym[xF(x)]]_+\| + \|[-x]_+\| + \|[-F(x)]_+\|.$$

A seguir mostraremos que h fornece uma cota de erro global para o $PCSD$.

Lema 4.3.1 *Suponha que F é uma função P-uniforme. Se $\overline{x}$ é a única solução do $PCSD$. Então existe uma constante $\alpha > 0$ tal que*

$$\alpha\|x - \overline{x}\|^2 \leq h(x),$$

para todo $x \in \mathcal{S}$.

Prova. Seja $p \in \mathcal{O}$. Como F é uma função P-uniforme, existe uma constante $\mu > 0$ e um índice i tal que

$$2\mu\,\|x - \overline{x}\|^2 \leq [p(x - \overline{x})(F(x) - F(\overline{x}))p^T + p(F(x) - F(\overline{x}))(x - \overline{x})p^T]_{ii}, \quad (4.12)$$

para todo $x \in \mathcal{S}$. Além disso, como $\overline{x}$ é solução do $PCSD$,

$$\lambda_i(\overline{x}) \geq 0, \lambda_i(F(\overline{x})) \geq 0, \quad \overline{x}F(\overline{x}) = 0, \quad \forall\ i, \quad (4.13)$$

onde as duas primeiras expressões seguem do fato de $\overline{x} \succeq 0$ e $F(\overline{x}) \succeq 0$, e a terceira expressão segue do Lema 1.2.5 (c). E, pelo Corolário 1.2.1, podemos tomar p tal que

$$p\overline{x}p^T = diag\{\lambda_1(\overline{x}), \ldots, \lambda_n(\overline{x})\} \quad (4.14)$$

e

$$pF(\overline{x})p^T = diag\{\lambda_1(F(\overline{x})), \ldots, \lambda_n(F(\overline{x}))\}. \tag{4.15}$$

Portanto, para todo $x \in \mathcal{S}$,

$$\begin{aligned}
& [p(x-\overline{x})(F(x)-F(\overline{x}))p^T + p(F(x)-F(\overline{x}))(x-\overline{x})p^T]_{ii} = \\
= \;& \left[sym[pxF(x)p^T]\right]_{ii} - \left[sym[p\overline{x}F(x)p^T]\right]_{ii} - \left[sym[pF(\overline{x})xp^T]\right]_{ii} \\
= \;& \left[sym[pxF(x)p^T]\right]_{ii} + \left[sym[p\overline{x}p^T(-pF(x)p^T)]\right]_{ii} + \left[sym[pF(\overline{x})p^T(-pxp^T)]\right]_{ii} \\
= \;& \left[sym[pxF(x)p^T]\right]_{ii} + 2\lambda_i(\overline{x})\left[-pF(x)p^T\right]_{ii} + 2\lambda_i(F(\overline{x})\left[-pxp^T\right]_{ii} \\
\leq \;& 2\left[p(sym[xF(x)])p^T\right]_{+}]_{ii} + 4\lambda_i(\overline{x})[[-pF(x)p^T]_+]_{ii} + 4\lambda_i(F(\overline{x})[[-pxp^T]_+]_{ii} \\
\leq \;& \tau\left(\left[p(sym[xF(x)])p^T]_+\right]_{ii} + [[-pF(x)p^T]_+]_{ii} + [[-pxp^T]_+]_{ii}\right) \\
= \;& \tau\left(\left[p[sym[xF(x)]]_+p^T\right]_{ii} + \left[p[-F(x)]_+p^T\right]_{ii} + \left[p[-x]_+p^T\right]_{ii}\right),
\end{aligned} \tag{4.16}$$

onde a primeira igualdade segue da terceira expressão de (4.13), a terceira igualdade segue de (4.14) e (4.15), a primeira desigualdade usa o fato que

$$2[a]_+ \succeq a, \; \forall\; a \in \mathcal{S}, \tag{4.17}$$

e a última desigualdade segue definindo τ da seguinte forma

$$\tau := \max\{2, 4\lambda_i(\overline{x}), 4\lambda_i(F(\overline{x}))\} > 0.$$

Então, combinando (4.12) e (4.16), obtemos

$$2\mu\left\|x-\overline{x}\right\|^2 \leq \tau\left(\|[sym[xF(x)]]_+\| + \|[-x]_+\| + \|[-F(x)]_+\|\right) = \tau h(x),$$

que é a desigualdade desejada. ■

A seguir mostraremos uma relação entre h e g, e com isso g fornece uma cota de erro global para o $PCSD$.

Teorema 4.3.1 *Suponha que F é P-uniforme e que*

$$\psi_{0^*}(t) \geq \alpha_1 \|[t]_+\|^2, \forall\ t \in \mathcal{S}, \tag{4.18}$$

e

$$\psi(a,b) \geq \alpha_2 \left(\|[-a]_+\| + \|[-b]_+\|\right)^2\ \ \forall\ (a,b) \in \mathcal{S} \times \mathcal{S}, \tag{4.19}$$

onde α_1 e α_2 são constantes positivas. Então existe um $\alpha > 0$ tal que

$$h(x) \leq 2\alpha g(x)^{1/2} \quad \forall \quad x \in \mathcal{S}.$$

Além disso, se o conjunto solução do $PCSD$ é não-vazio, então a função g dada por (4.3)-(4.6) *fornece uma cota de erro global para o $PCSD$.*

Prova. Pela hipótese sob ψ_{0^*} e ψ, existem constantes positivas k_1 e k_2 tais que

$$\|\left[sym[xF(x)]\right]_+\| \leq k_1 \psi_{0^*}^{1/2}\left(sym[xF(x)]\right)$$

e

$$\|[-x]_+\| + \|[-F(x)]_+\| \leq k_2 \psi^{1/2}(x, F(x)).$$

Seja $\alpha = \max\{k_1, k_2\}$, então

$$\begin{aligned} h(x) &\leq \alpha\left(\psi_{0^*}^{1/2}(sym[xF(x)]) + \psi^{1/2}(x, F(x))\right) \\ &\leq \alpha\left(g(x)^{1/2} + g(x)^{1/2}\right) \quad \text{para qualquer } \ x. \end{aligned}$$

Portanto, pelo Lema 4.3.1, g fornece uma cota de erro para o $PCSD$. ∎

Pela definição das funções dadas no Exemplo 4.2.0.4, ψ_{0^*} satisfaz (4.18) e ψ_1 satisfaz (4.19). Além disso, pelo Lema 2.5.9, ψ_2 satisfaz (4.19).

4.4 Conjunto de nível limitado

Nesta seção, consideramos em que condições os conjuntos de nível

$$\mathcal{L}(\gamma) = \{x \in \mathcal{S} / g(x) \leq \gamma\}$$

são limitados para qualquer $\gamma \geq 0$.

Para o $PCSD$, Shibata, Yamashita e Fukushima [30] mostraram que os conjuntos de nível da função definida por (2.15)-(2.17) são limitados sob a hipótese de F ser P-uniforme e $F(x) = \varpi(\lambda_{max}(x))$, onde $\varpi(t)$ é tal que $\lim_{t \longrightarrow \infty} \|\varpi(t)\| / t^2 = 0$ e $\lambda_{max}(x)$ é o maior autovalor de $x \in \mathcal{S}$. Devido a definição da g, o próximo resultado mostra que os conjuntos de nível de g são limitados desde que F seja P-uniforme apenas.

Teorema 4.4.1 *Suponha que F é uma função P-uniforme e considere ϕ definida por* (4.4) *tal que ψ_{0^*} satisfaz* (4.18), *e g definida por* (4.3). *Então os conjuntos de nível de g são limitados para todo $\gamma \geq 0$.*

Prova. Suponha que para algum $\gamma \geq 0$ o conjunto de nível de g, $\mathcal{L}(\gamma)$, é ilimitado. Então existe uma sequência ilimitada $\{x^k\} \subseteq \mathcal{L}(\gamma)$ tal que $\left\|x^k\right\| \to \infty$. Como F é uma função P-uniforme, existe um índice i tal que

$$[p(x^k F(x^k) - x^k F(0))p^T + p(F(x^k)x^k - F(0)x^k)p^T]_{ii} \geq 2\mu \|x^k\|^2, \tag{4.20}$$

onde μ é uma constante positiva e p é uma matriz ortogonal. E,

$$sym[x^k F(0)] = o(\|x^k\|^2). \tag{4.21}$$

Logo, segue que, de (4.20) e (4.21)

$$[psym[x^k F(x^k)]p^T]_{ii} \geq 2\mu \|x^k\|^2.$$

Dividindo ambos os lados da desigualdade acima por $\|x^k\|^2$ e tomando o limite, temos que

$$\liminf_{k\longrightarrow\infty} \frac{[p(sym[x^k[F(x^k)])p^T]_{ii}}{\|x^k\|^2} > 0.$$

Assim, existe um índice i tal que $[p(sym[x^kF(x^k)])p^T]_{ii} \to \infty$. Então por (4.18) temos que $\psi_{0^*}\left(sym[x^kF(x^k)]\right) \to \infty$, segue que $g(x^k) \to \infty$, o que contradiz o fato que $\{x^k\}\subseteq \mathcal{L}(\gamma)$. ■

A seguir, apresentaremos outra condição que garante a limitação dos conjuntos de nível de g. Para isso introduzimos outra classe de função-R_0.

Definição 4.4.1 *Dizemos que a função $F : \mathcal{S} \to \mathcal{S}$ é uma função-R_{02}, se para qualquer sequência $\{x^k\}$ satisfazendo*

$$\|x^k\| \longrightarrow \infty, \frac{[-x^k]_+}{\|x^k\|} \longrightarrow 0, \frac{[-F(x^k)]_+}{\|x^k\|} \longrightarrow 0 \tag{4.22}$$

e para qualquer $p \in \mathcal{O}$ existe um i tal que

$$\liminf_{k\longrightarrow\infty} \frac{[p[sym[x^kF(x^k)]]p^T]_{ii}}{\|x^k\|} > 0. \tag{4.23}$$

Esta classe de funções é uma extensão da classe utilizada no *PCN* [5, 8, 25]. A seguir mostraremos a relação entre as classes de funções-R_0 no sentido das Definições 3.4.1 e 4.4.1.

Lema 4.4.1 *Seja $F : \mathcal{S} \to \mathcal{S}$. Se F é uma função-R_{01}, então F é uma função-R_{02}.*

Prova. Note que

$$\liminf_{k\longrightarrow\infty} \frac{\langle x^k, F(x^k)\rangle}{\|x^k\|} > 0 \Longrightarrow \liminf_{k\longrightarrow\infty} \frac{\max_i[p[sym[x^kF(x^k)]]p^T]_{ii}}{\|x^k\|} > 0.$$

Então, pelas Definições 3.4.1 e 4.4.1, uma função-R_{01} é uma função-R_{02}. ■

O próximo Teorema é uma extensão do resultado obtido por Kanzow, Yamashita e Fukushima para o PCN [[25], Teorema 4.2]. A demonstração é uma modificação do Teorema 3.4.2.

Teorema 4.4.2 . *Suponha que F é uma função-R_{02} e considere ϕ definida por* (4.4) *tal que ψ_{0*} e ψ satisfazem* (4.18) *e* (4.19), *respectivamente, e g definida por* (4.3). *Então os conjuntos de nível de g são limitados para todo $\gamma \geq 0$.*

Prova. Suponha que para algum $\gamma \geq 0$ o conjunto de nível de g, $\mathcal{L}(\gamma)$, é ilimitado. Então existe uma sequência ilimitada $\{x^k\} \subseteq \mathcal{L}(\gamma)$ tal que $\left\|x^k\right\| \to \infty$. Pela condição (4.19) temos que as sequências de mínimo dos autovalores de $\{x^k\}$ e $\{F(x^k)\}$ são limitados inferiormente. Portanto,

$$\frac{[-x^k]_+}{\|x^k\|} \to 0, \frac{[-F(x^k)]_+}{\|x^k\|} \to 0$$

Logo, por hipótese, para p ortogonal, temos (4.23). Assim, existe um índice i tal que $[p(sym[x^k[F(x^k)])p^T]_{ii} \to \infty$. Então, por (4.17) e (4.18), temos que

$$\psi_{0*}(sym[x^k F(x^k)]) \to \infty,$$

e consequentemente $g(x^k) \to \infty$, o que contradiz o fato que $\{x^k\} \subset \mathcal{L}(\gamma)$. Portanto, os conjuntos de nível $\mathcal{L}(\gamma)$ são limitados. ■

Note que se F é uma função-R_{01}, pelo Lema 4.4.1, então F é uma função-R_{02}. Assim, as hipóteses para F no Teorema 4.4.2 são mais fracas do que a do Teorema 3.4.2. Além disso, pela Proposição 3.4.1, os conjuntos de nível de g são limitados quando F é monótona e o $PCSD$ é estritamente viável.

Conclusões

Neste trabalho estendemos alguns resultados do PCN [5, 25, 27] para o $PCSD$.

No Capítulo 3, apresentamos novas propriedades para a classe de funções Luo e Tseng, estendida por Tseng [32]. Utilizando-se da teoria do PCN, generalizamos para o $PCSD$ os conceitos de funções-R_{01} e R_{02}; e matrizes regulares e estritamente reuglares. Mostramos que f definida por (3.1)-(3.3) tem conjuntos de nível limitados desde que F seja uma função-R_{01} e que qualquer ponto estacionário é solução do $PCSD$ se, e somente se, x é regular (ou estritamente regular), condições estas mais fracas do que as obtidas anteriormente para o $PCSD$ [32, 38]. Também apresentamos dois algoritmos baseados nesta classe de funções para resolver o $PCSD$, para o caso onde a função de mérito f é definida com $\psi_0 \in \Psi_0$ e $\psi \in \Psi_{++}$; provamos a convergência global sob a hipótese de F ser fortemente monótona. Além disso, adicionando as hipóteses de F e $\nabla F(x)$ serem Lipschitz contínuas, provamos a convergência para o caso onde a função de mérito f é definida com $\psi_0 \in \Psi_0$ e $\psi \in \Psi_+$.

No Capítulo 4, propusemos uma nova classe de funções de mérito para o $PCSD$ e obtidas algumas propriedades. Em particular, mostramos que g definida por (4.3)-(4.6) fornece cota de erro global e tem conjuntos de nível limitados sob a condição de F ser P-uniforme. Além disso, garantimos a limitação dos conjuntos de nível de g se F é uma função-R_{02}, condição mais fraca do que F ser uma função-R_{01}.

Por outro lado, substituindo o segundo termo da função g_2, definida por (4.11), pela função Fischer-Burmeister quadrada obtemos a seguinte função

$$l(x) = \frac{1}{2}\|[sym[xF(x)]_+\|^2 + \frac{1}{2}\|\sqrt{x^2+(F(x))^2} - x - F(x)\|^2,$$

onde o segundo termo de $l(x)$, pelos Lemas 2.5.2 (b) e 2.5.9 satisfaz, respectivamente,

$$\sqrt{x^2+(F(x))^2} - x - F(x) = 0 \Longleftrightarrow x \in \mathcal{K} \quad F(x) \in \mathcal{K}, \quad \langle x, F(x)\rangle = 0$$

e

$$\sqrt{\frac{1}{2}\left\|\sqrt{x^2+(F(x))^2} - x - F(x)\right\|^2} \geq \frac{1}{2\sqrt{2}}\left(\|[-x]_+\| + \|[-F(x)]_+\|\right).$$

Logo, l goza das mesmas propriedades que g. Além disso, l é uma extensão da função de mérito do PCN obtida por Yamada, Yamashita e Fukushima [36].

A seguir apresentaremos algumas sugestões para pesquisas futuras:

1. Para o PCN, Yamada, Yamashita e Fukushima (2000) propuseram um algoritmo que converge desde que $F : I\!R^n \to I\!R^n$ seja apenas monótona. Uma questão em aberto é a obtenção deste resultado, baseado na classe de funções Luo e Tseng, com $\psi_0 \in \Psi_0$ e $\psi \in \Psi_{++}$.

2. Sabemos que se $F : I\!R^n \to I\!R^n$ é uma função P-uniforme, então F é um função-R_{02}. Está em aberto, para o espaço das matrizes simétricas, a verificação desta propriedade e a existência do $PCSD$ sob a condição de F ser P-uniforme apenas.

3. Para o PCN, Yamada, Yamashita e Fukushima obteram um algoritmo que converge sob certas condições, em particular sob a condição de F ser fortemente monótona. Usando a nova classe de funções ou a função $l(x)$ seria desejável obter resultados similares quando F é P-uniforme.

Referências Bibliográficas

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

[2] Auchmuty, G. "Variational principles for variational inequalities", *Numer. Func. Anal. Optim.*, v. 10, pp. 863-874, 1989.

[3] Auslender, A. *Optimisation: Méthodes Numériques*, 2 ed. Paris, Masson, 1976.

[4] Berman, A. "Cones, Matrices and Mathematical Programming", *Lectures Notes in Economics and Mathematical Systems*, v. 79, 1973.

[5] Chen, B. "Error bounds for R_0-type and Monotone Nonlinear Complementarity Problems", *Journal of Optimization Theory and Applications*, v. 108, n. 2, pp. 297-316, 2001.

[6] Chen, Xin and Qi, H. D. *An equivalent characterization of the monotone linear operator defined on the space of symmetric matrices.* Technical Report, Department of Mathematics, University of Washington, Seattle, Washington 98195-4350, 1998.

[7] Chen, X. and Tseng, P. *Non-interior continuation methods for solving semidefi-*

nite complementarity problems. Technical Report, Department of Mathematics, University of Washington, Seattle, Washington 98195, 1999.

[8] Cottle, R., Pang, J.S., e Stone, R., *The Linear Complementarity Problem.* New York, Academic Press, 1992.

[9] Fachinei, F. and Kanzow, C. "On Unconstrained and Constrained Stationary Points of the Implicit Lagrangian", *Journal of Optimization Theory and Applications*, v.92, pp. 99-115, 1997.

[10] Fachinei, F. and Soares, J. "A new function for nonlinear complementarity problems and a related algorithm", *SIAM Journal on Optimization*, v. 7, pp. 225-247, 1997.

[11] Fischer, A. "An NCP-function and its use for the solution of complementarity problems". In: D.-Z. Du, L. Qi and R. S. Womersley (eds), *Recent Advances in Nonsmooth Optimization*, Singapore, World Scientific Publishers, pp. 88-105, 1995.

[12] Fukushima, M. "Equivalent differentiable optimization problems and descent methods for assymetric variational inequality problems", *Mathematical Programming*, v. 53, pp. 99-110,1992.

[13] Fukushima, "M. Merit Functions for Variational Inequality and complementarity Problems". In: G. Di Pillo and F. Giannessi (eds), *Nonlinear Optimization and Applications*, New York, Plenum Publishing corporation, pp. 155-170, 1996.

[14] Geiger, C. and Kanzow, C. "On the resolution of monotone complementarity problems", *Comput. Optim. Appl.*, v. 5, pp. 155-173, 1996.

[15] Gowda, M. S. and Song, Y. "Semidefinite linear complementarity problems and a theorem of Lyapunov", Technical Report n$^{\bar{o}}$ TR99-04, Deparment of Mathematics and Statistics, University of Mayland, Baltimore, USA, 1999.

[16] Gowda, M. S. and Song, Y. "Semidefinite linear complementarity problems", *Mathematical Programming*, v. 88, pp. 575-587, 2000.

[17] Graybill, F. A. *Matrices with Applications in Statistics*, California, Wadsworth Internationa statistics/probability series, 1983.

[18] Harker, P. T. and Pang, J.-S. "Finite-dimensional variational inequality and nonlinear complementarity problem: A survey of theory, algoritms and aplications", *Mathematical Programming*, v. 48, pp. 161-220, 1990.

[19] Hearn, D. W. "The gap function of a convex program", *Operations Research Letters*, v.1, pp. 67-71, 1982.

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

[21] Horn, R. A. and Johnson, C. R. *Topics in Matrix Analysis.* Cambridge, Cambridge University Press, 1991.

[22] Isac, G. "Complementarity Problems", *Lectures notes in Mathematics*, New York, Springer-Verlag, 1992.

[23] Jiang, H. "Unconstrained minimization approaches to nonlinear complementarity problems", *Journal of Global Optimization*, v. 9, pp. 169-181, 1996.

[24] Kanzow, C. "Nonlinear complementarity as unconstrained optimization", *Journal of Optimization Theory and Applications*, v. 88, pp. 139-155, 1996.

[25] Kanzow, C., Yamashita, N. and Fukushima, M. "New NCP-Functions and Their Properties", *Journal of Optim. Theory and Appl.*, v. 94, pp. 115-135, 1997.

[26] Kojima, M., Shindoh, S. Hara, S. "Interior-point methods for the monotone semidefinite linear complementarity problems", *SIAM Journal on Optmization*, v. 7, pp. 86-125, 1997.

inequalities, Math.Programming, 64, 53-79, 1994.

[27] Luo, Z.-Q and Tseng, P. "A new class of merit functions for the nonlinear complementarity problem". In M. C. Ferris and J. -S. Pang (Eds.), *Complementarity and Variational Problems: State of the Art*, Philadelphia, SIAM, pp. 204-225, 1997.

[28] Ortega, J. M. and Rheinboldt, W. C. *Iterative Solution of Nonlinear Equations in Several Variables.* New York, Academic Press, 1970.

of Complementarity Problems. Journal of Optmi.

[29] Sun, D. and Qi, L. "On NCP-functions", *Computational of Optimization and Applications*, v. 13, n. 1/2/3, 1999.

[30] Shibata, M., Yamashita, N. Fukushima, M. "The Extended Semidefinite Linear Complementarity Problem: A Reformulation Approach". In: W. Takahashi and T. Tamaka (eds), *Non Analysis and Convex Analysis*, workscientific, pp. 326-332, 1999.

[31] Tiel, J. V. *Convex Analysis: An Introductory Text.* John Wiley and Sons, 1984.

[32] Tseng, P. "Merit functions for semi-definite complementarity problems", *Mathematical Programming*, v. 83, pp. 159-185, 1998.

[33] Tseng, P., Yamashita, N. and Fukushima, M. "Equivalence of complementarity problems to differentiable minimization: A unified approach," *SIAM Journal on Optimization*, v. 6, pp. 446-460, 1996.

[34] Vandenberghe, L. e Boyd, S. "Semidefinite Programming", *SIAM Review*, v. 38, pp. 49-95, 1996.

[35] Vandenberghe, L. e Boyd, S. "Applications of semidefinite programming", *Applied Numerical Mathematics*, v. 29, pp. 283-299, 1999.

[36] Yamada, K., Yamashita, N. and Fukushima, M. "A New Derivative-Free Descent Method for the Nonlinear Complementarity Problem". In G. D. Pillo and F. Giannessi (eds) *Nonlinear Optimization and Related Topics*, Boston, Massachusets, Kluwer Academic Publishers, pp. 463-487, 2000.

[37] Yamashita, N. and Fukushima, M. "On stationary points of the implicit Lagrangian for nonlinear complementarity problems", *Journal of Optimization Theory and Appl.*, v. 84, pp. 653-663, 1995.

[38] Yamashita, N. and Fukushima, M. "A New merit function and a descent method for semidefinite complementarity problems". In M. Fukushima and L. Qi (eds), *Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, Boston, Kluwer Academic Publishers, pp. 405-420, 1999.

[39] Yamashita, N. "Properties of Restricted NCP Functions for Nonlinear Complementarity Problems", *Journal Optim. Theory and Appl.*, v. 98, pp. 701-717, 1998.