

**Condições de qualificação para
Otimização Semidefinida**

Lucas Motta Freire

DISSERTAÇÃO APRESENTADA AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO TÍTULO DE
MESTRE EM MATEMÁTICA APLICADA

Programa: Matemática Aplicada
Orientador: Prof. Dr. Gabriel Haeser

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

São Paulo
Agosto de 2023

Condições de qualificação para Otimização Semidefinida

Lucas Motta Freire

Esta versão da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 28 de Agosto de 2023.

Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão julgadora:

Prof. Dr^a. Daiana Oliveira dos Santos – UNIFESP

Prof. Dr. Roberto Andreani – UNICAMP

Dr. Leonardo Makoto Mito – Mercado Livre

Prof. Dr. José Mario Martínez – UNICAMP

*O conteúdo deste trabalho é publicado sob a licença CC BY 4.0
(Creative Commons Attribution 4.0 International License)*

Ficha catalográfica elaborada com dados inseridos pelo(a) autor(a)
Biblioteca Carlos Benjamin de Lyra
Instituto de Matemática e Estatística
Universidade de São Paulo

FREIRE, LUCAS

Condições de qualificação para Otimização Semidefinida /
LUCAS FREIRE; orientador, Gabriel Haeser. - São Paulo, 2023.
55 p.: il.

Dissertação (Mestrado) - Programa de Pós-Graduação em
Matemática Aplicada / Instituto de Matemática e Estatística
/ Universidade de São Paulo.

Bibliografia

Versão corrigida

1. Programação Semidefinida. 2. Programação sobre
o Cone de Segunda Ordem. 3. Weak Nondegeneracy. I.
Haeser, Gabriel. II. Título.

Bibliotecárias do Serviço de Informação e Biblioteca
Carlos Benjamin de Lyra do IME-USP, responsáveis pela
estrutura de catalogação da publicação de acordo com a AACR2:
Maria Lúcia Ribeiro CRB-8/2766; Stela do Nascimento Madruga CRB 8/7534.

Este trabalho é dedicado a minha família, em especial, a Dilclidiane Fidelis Lira.

Agradecimentos

O jogo não acaba aqui.

— Clóvis de Barros Filho

Em uma dissertação, ou qualquer outro trabalho acadêmico, espera-se que o autor receba o maior reconhecimento pelo trabalho que fez, entretanto, para o meu caso, acredito que o maior reconhecimento deva ser atribuído essencialmente para as maravilhosas pessoas que estiveram ao meu lado durante todo este processo. Primeiramente, agradeço a Deus, e em especial, ao professor Gabriel Haeser, que sem sua grande colaboração e orientação, este trabalho não seria possível. Agradeço-lhe, por ser, além de um grande pesquisador e orientador, um grande ser humano para aqueles que estão ao seu lado.

Agradeço imensamente a toda minha família, na qual, sempre estive me estendendo todo o suporte e compreensão para eu realizar este feito. Um agradecimento mais que especial, ao meu grande amor Dilclidiane Fidelis Lira, que sempre demonstrou companheirismo, afeto e incentivo nesta empreitada, com certeza, sem ela, isto não seria possível. Agradeço também a minha mãe, dona Maria Ivanilza de Souza Motta, que jamais deixou de estar ao meu lado em conjunto com a Dona Maria Solange, que por vezes, a considero como uma segunda mãe.

Deixo por fim, os meus sinceros agradecimentos ao professor Roberto Andreani, que em conjunto com o Gabriel, sempre proporcionou ajuda a esclarecer e contribuir com este trabalho, de modo que, boa parte dos resultados aqui expostos, foram desenvolvidos com o seu auxílio e orientação, contribuindo assim, exponencialmente, para a minha formação. Com certeza, há inúmeras pessoas que deveriam ser mencionadas aqui, amigos, colegas de estudo, professores e alunos, entretanto, para não se tornar algo longo e cansativo, deixo aqui, meus sinceros agradecimentos a todos que, sob alguma parcela, contribuiu para minha formação e trabalho.

Resumo

Lucas Motta Freire. **Condições de qualificação para Otimização Semidefinida**. Dissertação (Mestrado). Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2023.

Nondegeneracy-CQ e Robinson's-CQ são duas das mais conhecidas condições de qualificação para problemas de otimização semidefinida, estas consistem em analisar a independência linear de um conjunto de $m(m+1)/2$ vetores que dependem do núcleo da restrição. O objetivo deste trabalho consiste em estudar uma nova maneira de avaliar a Nondegeneracy-CQ e Robinson's-CQ, de maneira a se analisar a independência de um conjunto de cardinalidade significativamente menor, originando assim o conceito de "Weak-Nondegeneracy-CQ" e "Weak-Robinson's-CQ". O intuito deste trabalho consiste em estudar novas condições de qualificação equivalentes a Nondegeneracy-CQ e Robinson's-CQ para Programação não linear Semidefinida (NSDP), estruturando-se inicialmente sobre algumas noções obtidas da Programação Não Linear sobre o Cone de Segunda Ordem (NSOCP). Estruturamos este trabalho, primeiramente, estudando condições de qualificação mais fracas para NSOCP, em especial, Nondegeneracy-CQ e Robinson's-CQ, com o intuito de estabelecer condições de qualificação equivalentes a estas, sobre o contexto de (NSOCP). Por fim, todas as definições, resultados e exemplos obtidos para NSOCP são estendidos naturalmente para Programação Semidefinida.

Palavras-chave: Programação Semidefinida. Programação Não Linear sobre o Cone de Segunda ordem. Weak Nondegeneracy.

Abstract

Lucas Motta Freire. **Constraint Qualifications for Semidefinite Programming.**
Thesis (Master's). Institute of Mathematics and Statistics, University of São Paulo, São Paulo, 2023.

Nondegeneracy and Robinson's-CQ are some of the best-known constraint qualifications for semidefinite optimization problems, which consist of analyzing the linear independence of a set of $m(m + 1)/2$ vectors that depend on the constraint kernel. The aim of this work is to study a new way of evaluating Nondegeneracy-CQ and Robinson's-CQ, to analyze the independence of a set of significantly smaller cardinality, thus originating the concept of "Weak-Nondegeneracy-CQ" and "Weak-Robinson's-CQ". The purpose of this paper is to study new qualification conditions equivalent to Nondegeneracy-CQ and Robinson's-CQ for Nonlinear Semi-Definite Positive Programming (NSDP), structuring initially on some notions obtained from Nonlinear Second Order Programming (NSOCP). We structure this work, first, by studying weaker qualification conditions for NSOCP, in particular, Nondegeneracy-CQ and Robinson's-CQ, to establish qualification conditions equivalent to these, on the context of (NSOCP). Finally, all the definitions, results, and examples obtained for NSOCP are naturally extended to Semidefinite Programming.

Keywords: Semidefinite programming. Second order cone programming. Weak Nondegeneracy.

Lista de abreviaturas

CQ	Condição de Qualificação (→Constraint Qualification)
MFCQ	Condição de Qualificação de <i>Mangasarian-Fromovitz</i> (→Mangasarian-Fromovitz's constraint qualification)
LICQ	Condição de Qualificação de Independência Linear (→Linear Independence constraint qualification)
NCP	Programação Não Linear Cônica (→Nonlinear Conic Programming)
NSOCP	Programação Não Linear sobre o Cone de Segunda Ordem (→Nonlinear Second-Order cone programming)
NSDP	Programação Não Linear Semidefinida (→Nonlinear Semidefinite Programming)
KKT	Karush - Kuhn - Tucker

Lista de símbolos

\mathbb{L}_m	Cone de <i>Lorentz</i> de dimensão m
\mathbb{S}^m	Espaço das Matrizes Simétricas de dimensão m
\mathbb{S}_+^m	Cone das Matrizes Simétricas Semidefinidas positivas de dimensão m
$\text{lin}(K)$	Espaço Linearizado de K (\rightarrow Linearity Space of K)
$\text{cone}(V)$	Fecho Cônico de V
C°	Cone Polar de C
$T_C(y)$	Cone Tangente a C em um ponto y
\mathbb{R}^m	Espaço Euclidiano de dimensão m
\mathbb{R}_+^m	Espaço Euclidiano de dimensão m com componentes maiores ou iguais a zero.
$\lambda_i(M)$	i -ésimo autovalor da matriz M
$\Pi_C(x)$	Projeção de x sobre C
$\text{dist}(x, C)$	Distância de x até o conjunto C
$[x]_+$	Máximo entre 0 e $x \in \mathbb{R}$
M^*	Aplicação/Matriz adjunta de M
$\nabla f(x)$	Gradiente de f aplicado em x

Lista de figuras

2.1	Decomposição em \mathbb{L}_m	18
-----	--	----

Sumário

Introdução	1
1 Conceitos Preliminares	3
1.1 Cones e Independência Linear	3
1.1.1 Cones	3
1.1.2 Independência linear	4
1.2 Bases paramétricas	5
1.3 Matrizes simétricas	6
1.4 Projeção	9
1.4.1 Funções espectrais	11
1.5 Problemas de Otimização Cônica	11
2 Nondegeneracy para NSOCP	15
2.1 O cone de segunda ordem	15
2.2 O problema sobre o cone de segunda ordem	19
2.3 Nondegeneracy e Bases Paramétricas	21
3 Nondegeneracy para NSDP	27
3.1 O Problema Não Linear Semidefinido Positivo	27
3.1.1 Elementos da Análise Convexa	31
3.1.2 Redução para NSDP	32
3.2 Weak-Nondegeneracy para NSDP	37
3.2.1 Sequência de Autovetores e <i>Nondegeneracy</i>	37
3.2.2 Exemplos	42
3.2.3 Resultados interessantes	46
4 Conclusão	51

Introdução

Uma das principais estratégias para o estudo de problemas de otimização em geral, consiste em estudar o conjunto viável de maneira tal, a encontrar pontos que satisfaçam as tão conhecidas condições de Karush-Kuhn-Tucker (KKT) (pontos estacionários). Estas condições foram inicialmente desenvolvidas por Karush em 1939 em sua dissertação de mestrado no Departamento de Matemática da Universidade de Chicago porém, só foram reconhecidas na década de 70 devido o grande desenvolvimento tecnológico instaurado e seu grande valor para a resolução de problemas militares e econômicos referentes a época. Contudo, é bem conhecido na literatura que se um ponto viável satisfizer as condições KKT este no entanto, não será necessariamente uma solução para o problema, ou seja, estas equações não constituem por si só uma condição de otimalidade para o problema.

Entretanto, algumas hipóteses de regularidade podem ser impostas sobre a descrição analítica do conjunto viável, de modo a garantir, que toda solução local satisfaz as equações KKT, ou como é comumente falado, é um ponto estacionário para o problema. Estas hipóteses sobre o conjunto viável são nomeadas pela literatura de Qualificações de Restrições, em inglês, *constraint qualifications*. Para o nosso trabalho utilizaremos o termo em português conhecido por, *condições de qualificação* (CQs). A grosso modo, estas condições garantem descrever o conjunto viável localmente em termos de uma aproximação de primeira ordem, isto é, permite caracterizar toda solução do problema em termos da primeira derivada das funções restrições que descrevem o conjunto viável. Neste trabalho, estaremos interessados em fazer um levantamento teórico sobre duas das mais relevantes condições de qualificação para problemas de otimização com restrições cônicas, *Nondegeneracy-CQ* e *Robinson's-CQ*. Contudo, nosso intuito principal será estudá-las no contexto de Programação Não Linear sobre o cone de Segunda Ordem e Programação Não Linear Semidefinida, isto é, problemas de otimização impostos sobre os cones:

$$\mathbb{L}_m \doteq \{x \doteq (x_0, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{m-1} : x_0 \geq \|\bar{x}\|\}$$

e

$$\mathbb{S}_+^m \doteq \{M \in \mathbb{S}^m : \forall x \in \mathbb{R}^m, x^T M x \geq 0\}$$

respectivamente.

Em especial, *Nondegeneracy-CQ* foi originalmente descrita sobre o contexto de problemas de otimização de autovalores por uma abordagem essencialmente geométrica, como é bem visto em [ROBINSON, 1976a](#), entretant0, nossa abordagem principal consiste em estudar *Nondegeneracy-CQ*, consequentemente *Robinson's CQ*, definindo-a propriamente

através da estrutura do problema, em contrapartida da definição usual dada pela transversalidade entre dois espaços. Assim, usaremos uma caracterização feita em SHAPIRO, 1997, Prop. 6 que consiste em escrever a *Nondegeneracy-CQ* em termos do gradiente das entradas de um "bloco ativo" isolado da restrição.

Durante este texto vamos seguir essencialmente a teoria construída nos trabalhos ANDREANI, HAESER, MITO, RAMÍREZ e SILVEIRA, 2022 e ANDREANI, HAESER, MITO e RAMÍREZ, 2023, que consistem em estudar condições de regularidade para NSOCP e NSDP através de, seqüências de autovetores obtidas por meio de uma decomposição pré estabelecida sobre o cone. Com isto, vamos propor o que os autores definiram como "*Weak-Nondegeneracy*" e "*Weak-Robinson's CQ*", cujas versões, fornecem uma interpretação mais fraca para a *Nondegeneracy-CQ* e *Robinson's-CQ* porém, apresentam grande valor teórico para a análise de pontos estacionários. É importante destacar, que a motivação para a construção destas versões mais fracas surgem primeiramente sobre o contexto de NSDP, oriundo do conceito de pontos aproximadamente KKT's para NSDP abordado em ANDREANI, HAESER e VIANA, 2020. Posteriormente em ANDREANI, HAESER, MITO, RAMÍREZ e SILVEIRA, 2022, foi estudado esta noção para o contexto de NSOCP com o intuito de estender os resultados obtidos até então para NSDP. No entanto, durante este trabalho será apresentada a construção de *Weak-Nondegeneracy* e "*Weak-Robinson's-CQ*" na ordem contrária dos fatos.

Para a leitura deste trabalho destacamos que, no capítulo 1 é levantado todo um arcabouço teórico de conceitos preliminares, cujo principal objetivo é contextualizar e justificar as proposições levantadas no capítulo 2 e capítulo 3. No capítulo 2, é levantado um breve estudo sobre a *Nondegeneracy-CQ* para NSOCP em conjunto com a definição de *Weak-Nondegeneracy* além de apresentar uma equivalência entre estes dois conceitos. Por fim, no capítulo 3 estudamos de fato as condições de qualificação para Programação Semi-definida cujo interesse maior será generalizar os resultados obtidos no capítulo 2.

Capítulo 1

Conceitos Preliminares

1.1 Cones e Independência Linear

Aqui discutiremos vários conceitos preliminares da análise convexa, análise matricial e álgebra linear, com ênfase sobre o conceito de independência linear, no qual demonstra ser de suma importância para um estudo estratégico das principais condições de qualificação para a otimização cônica.

Aqui, sempre que nos referirmos a espaços vetoriais estaremos restritos a espaços de dimensão finita equipados com algum produto interno, de modo que, para o produto interno usual, $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ será utilizada a notação, $\langle x, y \rangle = x^T y = \sum_{i=1}^n x_i y_i$, sejam quaisquer $x, y \in \mathbb{R}^n$. Para o produto interno em um espaço vetorial qualquer, manteremos a notação usual $\langle \cdot, \cdot \rangle$ ressaltando que, a menos de menção, o produto ficará implícito sob o contexto do espaço a ser trabalhado. Para vetores $x \in \mathbb{R}^n$ utilizaremos sua disposição em coluna onde, ocasionalmente poderá ser denotado por $x = (x_1, \dots, x_n)^T$.

1.1.1 Cones

Seja \mathbb{Y} um espaço linear de dimensão finita, um subconjunto não-vazio K de \mathbb{Y} é dito Cone quando, $\alpha x \in K$ para todo $x \in K$ e $\alpha \geq 0$, além disso, dados dois pontos $a, b \in K$, este é dito um cone convexo se, e somente se, conter o segmento de linha entre a, b , isto é, $\lambda a + (1 - \lambda)b \in K$ sempre que, $a, b \in K$ com $\lambda \in [0, 1]$. Vale lembrar que um cone $K \in \mathbb{Y}$ é dito "pontudo" (*pointed cone*) se, $x \in K$ e $-x \in K$ então, $x = 0$.

Definição 1.1 (*Lineality space*). Seja $K \subseteq \mathbb{Y}$ um cone, denotamos $\text{lin}(K)$ como o maior subespaço contido em K .

Note que, desta forma, $K \in \mathbb{Y}$ será um "pointed cone" se, e somente se, $\text{lin}(K) = \{0\}$.

A seguir definiremos importantes estruturas para a otimização cônica, que juntas, formaram o alicerce principal para nosso trabalho.

Definição 1.2 (Fecho Cônico). Seja $V \subseteq \mathbb{X}$, onde \mathbb{X} é um espaço vetorial, dada a família

$\{K_i; i \in I\}$ de todos os cones que contém V , onde I é um conjunto de índices define-se:

$$\text{cone}(V) = \bigcap_{i \in I} K_i \quad (1.1)$$

como o fecho cônico de V , isto é, o menor cone que contém V .

Uma importante estrutura para o estudo de condições de otimalidade para problemas de otimização nomeia-se *Cone Polar*, de modo a ser definido para um conjunto convexo, da seguinte forma:

Definição 1.3 (Cone Polar). *Dado um conjunto convexo $C \subseteq \mathbb{X}$ definimos o conjunto*

$$C^\circ \doteq \{y \in \mathbb{X} : \forall x \in C, \langle x, y \rangle \leq 0\}.$$

Como o Cone Polar de C .

Além do cone polar, é importante analisar, o conjunto de todas as direções tangentes em relação a um conjunto C , a partir de um ponto. Onde claramente este, satisfaz a estrutura de um cone, veja

Definição 1.4 (Cone Tangente). *Dado um conjunto C o cone tangente (Bouligand) em um ponto $y \in C$ é definido por:*

$$T_c(y) \doteq \left\{ d \in \mathbb{X} : \begin{array}{l} \exists \{d^k\}_{k \in \mathbb{N}} \rightarrow d, \exists \mathbb{R}_+ \ni \{t_k\}_{k \in \mathbb{N}} \rightarrow 0_+, \\ \forall k \in \mathbb{N}, y + t_k d^k \in C \end{array} \right\}.$$

Note que, tomar $y + t_k d^k \in C$ para as sequências acima definidas, nos garante, que dada uma direção tangente, ao diminuir o tamanho do passo t_k nesta direção, nos aproximamos do conjunto em uma ordem assintoticamente menor do que o passo, isto é, o cone tangente pode ser definido por

$$T_c(y) \doteq \left\{ d : \begin{array}{l} \exists \mathbb{R}_+ \ni \{t_k\}_{k \in \mathbb{N}} \rightarrow 0_+, \text{ tal que} \\ \text{dist}(y + t_k d^k, D) = o(t_k) \end{array} \right\}. \quad (1.2)$$

1.1.2 Independência linear

Seja $V = \{v_1, \dots, v_m\} \subset \mathbb{X}$, onde \mathbb{X} é um espaço vetorial, dizemos que os elementos de V são linearmente independentes se, e somente se, não existe $a \in \mathbb{R}^m$ não-nulo tal que,

$$\sum_{j=1}^m a_j v_j = 0 \quad (1.3)$$

Por outro lado, se existe $a \in \mathbb{R}^m$ não nulo, tal que (1.3) é válida, então os elementos de V , são ditos linearmente dependentes. Similarmente, se exigirmos $a_j \geq 0$ para todo $j \in \{1, \dots, m\}$, os vetores serão ditos positivo-linearmente independentes(dependentes). Vale lembrar, que o conceito de independência linear possui grande importância no estudo das aplicações lineares, visto que dada a matriz M que representa uma aplicação linear

$A : \mathbb{X} \rightarrow \mathbb{Y}$ sobre respectivas bases de \mathbb{X}, \mathbb{Y} . A é injetora se, e só se, M possui posto-coluna completo, isto é, as colunas de M formam um conjunto linearmente independente.

Em um primeiro momento, a definição de *cone* e o conceito de independência linear parecem ser dissociados entre si, entretanto, como será visto, relacionar estes conceitos pode nos fornecer uma nova abordagem para o estudo de condições de qualificação.

Definição 1.5 (Independência linear cônica). *Seja $K \subseteq \mathbb{R}^m$ um cone convexo e fechado. A matriz $M \in \mathbb{R}^{n \times m}$ é dita K -linearmente independente se não existe $v \in K$ não nulo, tal que $Mv = 0$.*

Observação 1.1. *Note que, dispondo os vetores $\{v_1, \dots, v_m\}$ como as colunas da matriz M com $v_i \in \mathbb{R}^m$, e assumindo $K = \mathbb{R}^m (\mathbb{R}_+^m)$, resgatamos o conceito usual de independência linear (positivo-independência linear) para vetores.*

Observação 1.2. *Note que a definição 1.5 nos garante uma caracterização de injetividade de M sobre o cone C .*

1.2 Bases paramétricas

Para um cone K convexo e fechado, verifica-se que é possível caracterizar a independência linear cônica de uma matriz M sobre K simplesmente analisando uma positivo-independência linear sobre conjuntos essencialmente menores que estruturam K . Veja:

Proposição 1.1. *Seja $K \subseteq \mathbb{R}^m$ um cone convexo e fechado, tal que existe um conjunto de índice J (possivelmente infinito) e, para cada $j \in J$, um subconjunto finito $\mathcal{E}_j \subseteq K$ cujos os elementos são linearmente independentes de tal forma que*

$$K = \bigcup_{j \in J} \text{cone}(\mathcal{E}_j). \quad (1.4)$$

Então, a matriz $M \in \mathbb{R}^{n \times m}$ é K -linearmente independente se, e somente se, a família $\{Mv\}_{v \in \mathcal{E}_j}$ é positivo-linearmente independente, para cada $j \in J$ fixado.

Prova. Suponha que M is K -linearmente independente, dado um índice $j \in J$ arbitrário e $\alpha_j \in \mathbb{R}_+$ escalares, $v \in \mathcal{E}_j$ tal que:

$$\sum_{v \in \mathcal{E}_j} \alpha_j Mv = M \left[\sum_{v \in \mathcal{E}_j} \alpha_j v \right] = 0 \quad (1.5)$$

Como K é um cone convexo, segue que a combinação linear positiva $\tilde{v} \doteq \sum_{v \in \mathcal{E}_j} \alpha_j v$ pertence a K , e por hipótese $\tilde{v} = 0$. Agora da independência linear de \mathcal{E}_j segue-se que $\alpha_j = 0$ para cada $v \in \mathcal{E}_j$. Portanto, a família $\{Mv\}_{v \in \mathcal{E}_j}$ é positivo-linearmente independente.

Por outro lado, assuma que $\{Mv\}_{v \in \mathcal{E}_j}$ é positivo-linearmente independente, daí, tome $\tilde{v} \in K$ tal que $M\tilde{v} = 0$. Então, existe algum índice $j \in J$ tal que $\tilde{v} \in \text{cone}(\mathcal{E}_j)$; isto é, existe

algum escalar $\alpha_j \geq 0$, $\tilde{v} \in \mathcal{E}_j$, de modo que $\tilde{v} = \sum_{v \in \mathcal{E}_j} \alpha_j v$ que satisfaz (1.5) daí, vale que $\alpha_j = 0$ para todo $v \in \mathcal{E}_j$; portanto $\tilde{v} = 0$. \square

Observação 1.3. Vale notar, que a proposição (1.1) pode ser facilmente estendida para o produto cartesiano de cones convexos fechados $K = \prod_{j \in J} K_j$ onde J é um conjunto de índice. Exigindo a K -independência linear da família de matrizes $\{M_j; j \in J\}$, partindo dos índices j , pode-se construir convenientemente a matriz bloco:

$$M \doteq \left[\begin{array}{c} \vdots \\ M_j \\ \vdots \end{array} \right]_{j \in J} \quad (1.6)$$

e exigir a K -independência linear de M .

Exemplo 1.1. Seja $K = \mathbb{R}^2$, note que utilizando a seguinte decomposição:

$$\mathbb{R}^2 \doteq \{(r \cos(\theta), r \sin(\theta)); r \geq 0 \text{ e } \theta \in [0, 2\pi]\} = \bigcup_{\theta \in [0, 2\pi]} \text{cone}(\cos(\theta), \sin(\theta))$$

Podemos então dizer, que uma matriz $M \in \mathbb{R}^{m \times 2}$ será \mathbb{R}^2 -linearmente independente (ou injetiva) se, e somente se, $M(\cos(\theta), \sin(\theta)) \neq 0$ para todo $\theta \in [0, 2\pi]$.

Observação 1.4. Perceba que o conceito de independência linear em \mathbb{R}^m pode ser facilmente posto em termos de bases paramétricas, basta tomarmos:

$$\mathbb{R}^m = \bigcup_{\substack{w \in \mathbb{R}^{m-1} \\ \|w\|=1}} \text{span}(\{(1, -w), (1, w)\}),$$

onde, $\text{span}(\{(1, -w), (1, w)\})$ denota o espaço linear gerado pelos vetores $(1, -w)$ e $(1, w)$. Então, uma matriz M é \mathbb{R}^m -linearmente independente (ou seja, injetiva) se, e somente se, os vetores

$$M(1, -w) \quad \text{e} \quad M(1, w)$$

são linearmente independentes para cada w na base paramétrica $\{w \in \mathbb{R}^{m-1} : \|w\| = 1\}$.

Note que no exemplo (1.1), trocamos a análise da independência linear de um conjunto com dois vetores, por analisar a independência linear de um conjunto com um único vetor para cada valor paramétrico θ , caracterizando a independência linear cônica por meio de bases paramétricas como é visto em (1.4). Este tipo de abordagem, será central para o estudo de condições de qualificação para problemas de otimização cônica que serão abordadas nos próximos capítulos.

1.3 Matrizes simétricas

Um dos nossos objetos de estudo será o espaço das matrizes simétricas a valores reais, isto é, matrizes M tais que, $M \in \mathbb{R}^{n \times n}$ e que $M^T = M$. Em especial, estamos interessados em estudar o subespaço das matrizes que satisfazem, $x^T M x \geq 0$ para todo $x \in \mathbb{R}^n$ não nulo,

isto é, o espaço das matrizes simétricas semi-definidas positiva (\mathbb{S}_+^m). Para a desigualdade estrita $x^T M x > 0$ diz-se que M é definida positiva.

A grosso modo, estas matrizes podem ser interpretadas como uma generalização de um número positivo qualquer de maneira a surgir naturalmente em diversos problemas na matemática, em especial em problemas de otimização. A seguir, será apresentado propriedades e caracterizações importantes que serão utilizadas por todo o restante do texto.

Proposição 1.2. *Seja uma matriz $M \in \mathbb{S}^m$ com $m \geq 2$. Se M é definida positiva, então:*

1. *Todos os elementos da sua diagonal são positivos;*
2. *Todos os autovalores de M são números reais positivos (não negativos, para $M \in \mathbb{S}_+^m$), e seus autovetores pertencem \mathbb{R}_*^+ (\mathbb{R}^m);*
3. *O determinante, o traço, e todos os determinantes principais menores de M são positivos;*
4. *Qualquer combinação linear não negativa de matrizes definida positivas é definida positiva;*
5. *Dado $U \in \mathbb{R}^{n \times m}$ então, $U^T M U$ é semi-definida positiva. Além que, $\text{rank}(U^T M U) = \text{rank}(U)$ e, $U^T M U$ é definida positiva se, e somente se, $\text{rank}(U) = m$;*

onde \mathbb{R}_*^+ denota o conjunto dos números reais positivos.

Demonstração. Os itens 1, 2 podem ser vistos em [SÜLI e MAYERS, 2007](#), Teo. 3.1. Para o item 3 e 4 recomendamos a leitura da seção 7.1 de [HORN e JOHNSON, 2012](#) em conjunto com o a observação 7.1.8 também de [HORN e JOHNSON, 2012](#). \square

Observação 1.5 ([SÜLI e MAYERS, 2007](#), Teo. 3.1). *As recíprocas de dois dos resultados da proposição (1.2) são verdadeiras:*

- *Se todos os autovalores de uma matriz simétrica M são positivos então, M é definida positiva;*
- *Se cada determinante principal menor de uma matriz simétrica M é positivo então, M é definida positiva;*

Proposição 1.3 ([HORN e JOHNSON, 2012](#), Teo. 7.2.6). *Seja $A \in \mathbb{S}_+^m$ e $k \in \{2, 3, \dots\}$, então existe uma única matriz $P \in \mathbb{S}_+^m$ tal que $P^k = A$. Mais ainda,*

- a) *$PA = AP$ e existe um polinômio real $p(t)$ tal que $B = p(A)$;*
- b) *$\text{rank}(A) = \text{rank}(P)$, isto é, B é positiva definida se A também for;*

O caso mais usual para a proposição acima é quando $k = 2$, de modo que utilizaremos para a caracterização da condição de complementariedade para um problema de otimização semidefinida.

Para uma matriz $M \in \mathbb{S}^m$ pode-se obter uma decomposição em termos de seus autovalores dispondo M por $M = U \Lambda U^T$ tal que $U = [u_1, \dots, u_m]$ é uma matriz ortogonal $m \times m$ cujas colunas são determinadas pelos autovetores de M e $\Lambda = \text{Diag}(\lambda_1(M), \dots, \lambda_m(M))$

uma matriz diagonal, cujos elementos são os respectivos autovalores de M . Comumente, pode-se escrever M da seguinte forma:

$$M = \sum_{i=1}^m \lambda_i(M) u_i(M) u_i(M)^T \quad (1.7)$$

onde, $\lambda_i(M) \in \mathbb{R}$ denota o i -ésimo autovalor de M dispostos em ordem não-crescente, (isto é $\lambda_1(M) \geq \lambda_2(M) \geq \dots \geq \lambda_m(M)$) e $u_i(M) \in \mathbb{R}^n$ corresponde a qualquer autovetor associado aos autovalores de modo que o conjunto $\{u_i(M) : i \in \{1, \dots, m\}\}$ é uma base ortormal de \mathbb{R}^m . Esta é a tão conhecida decomposição espectral para matrizes simétricas, para mais detalhes, veja [HORN e JOHNSON, 2012](#), Seção 4.4

Para o espaço das matrizes quadradas definimos a função autovalor $\lambda : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^n$ dada por $M \mapsto \lambda(M) = (\lambda_1(M), \dots, \lambda_n(M))^T \in \mathbb{R}^n$. Assim, a decomposição espectral para uma matriz M pode ser simplesmente escrita por $M = U \text{Diag}(\lambda(M)) U^T$.

Proposição 1.4 (Teo. 4.5.15 [HORN e JOHNSON, 2012](#)). *Sejam $A, B \in \mathbb{S}^m$ então $AB \in \mathbb{S}_+^m$ e $AB = BA$ se, e somente se, A e B compartilham a mesma decomposição espectral, isto é, existe uma matriz ortogonal $U \in \mathbb{R}^{m \times m}$ tal que $A = U \Lambda_A U^T$ e $B = U \Lambda_B U^T$.*

É bem conhecido que o espaço \mathbb{S}^m é um Espaço Euclidiano equipado com o produto interno de *Frobenius*, de modo que, dadas $M, N \in \mathbb{S}^m$ defini-se $\langle M, N \rangle \doteq \text{trace}(M^T N) = \text{trace}(MN) = \sum_{i,j=1}^m M_{i,j} N_{i,j}$, de modo a definir a norma $\|M\|_F \doteq \sqrt{\text{trace}(MN)}$. Neste espaço, pode-se definir diversas normas, relacionadas diretamente com os autovalores da matriz, veja a seção 5.6 de [HORN e JOHNSON, 2012](#), que caracteriza $\|M\|_F$ da seguinte forma

$$\|M\|_F = \sqrt{\lambda_1(M)^2 + \dots + \lambda_m(M)^2}$$

Outro exemplo para a norma que pode ser visto é dado por

Exemplo 1.2. *A l_∞ -norma é definida para uma matriz $M \in \mathbb{S}^m$ por*

$$\|M\|_\infty = \max_{1 \leq i, j \leq m} |m_{ij}|$$

Observação 1.6. *Vale lembrar que os autovalores de uma matriz simétrica são funções Lipschitz-contínuas em relação a entradas das matrizes, pois, se $\lambda(M)$ é vetor com os autovalores de M dispostos em ordem não crescente, então vale*

$$\begin{aligned} \|\lambda(M + N) - \lambda(M)\|_\infty &\leq \|\lambda(N)\|_\infty = \|N\|_2 \\ \|\lambda(M + N) - \lambda(M)\|_2 &\leq \|\lambda(N)\|_2 = \|N\|_F \end{aligned} \quad (1.8)$$

As duas desigualdades acima, são conhecidas por Teorema de Mirsky e Teorema de Hoffman-Wielandt respectivamente, veja [STEWART e SUN, 1990](#), Cor. 4.12

Uma propriedade importante para o produto interno de *Frobenius* é dada por um limite superior para o produto de duas matrizes que depende dos autovalores das mesmas, veja

Proposição 1.5 (THEOBALD, 1975). Dadas $A, B \in \mathbb{S}^m$ então, vale:

$$\langle M, N \rangle \leq \sum_{i=1}^m \lambda_i(A) \lambda_i(B)$$

Proposição 1.6. Se $A, B \in \mathbb{S}_+^m$ então as seguintes proposições são equivalentes

- a) $\langle A, B \rangle = 0$
- b) $AB = 0$

Demonstração. Pela proposição (1.3), sejam $A = P^2$ e $B = Q^2$ com $P, Q \in \mathbb{S}_+^m$. Então Para a) \Rightarrow b) basta notar que

$$\|PQ\|^2 = \langle PQ, PQ \rangle = \text{tr}(PQPQ) = \text{tr}(Q^T P^T PQ) = \text{tr}(QP^2Q) = \text{tr}(P^2Q^2) = \langle A, B \rangle = 0 \quad (1.9)$$

portanto, $AB = P(PQ)Q = 0$

Para b) \Rightarrow a) o resultado é imediato da definição do traço. \square

Para terminarmos esta seção, apresentaremos propriedades cônicas para o espaço das matrizes simétricas, semi-definidas positivas, como se segue

Proposição 1.7. Para o espaço \mathbb{S}_+^m vale:

1. \mathbb{S}_+^m é um cone convexo fechado.
2. \mathbb{S}_+^m é simétrico, isto é:
 - (a) \mathbb{S}_+^m é self-dual;
 - (b) \mathbb{S}_+^m é full-dimensional, isto é, o interior de \mathbb{S}_+^m é não vazio;
 - (c) \mathbb{S}_+^m é homogêneo, isto é, para todo par x, y no interior de \mathbb{S}_+^m , existe uma bijeção linear T , tal que $T(x) = y$ e $T(\mathbb{S}_+^m) = \mathbb{S}_+^m$;

1.4 Projeção

Dado um problema de otimização, veremos que uma importante estratégia para a construção de algoritmos que buscam soluções consiste em utilizar métricas que estimam a viabilidade de ponto qualquer dado, para o contexto desse trabalho, uma métrica bastante útil será determinada pelo *Operador Projeção*, cujo será bem definido sobre o contexto de cones convexos, fechados.

Definição 1.6 (Operador Projeção). Seja $x \in \mathbb{X}$ e um conjunto fechado $C \subseteq \mathbb{X}$ definimos o operador $\Pi_C(x): \mathbb{X} \rightarrow \mathbb{X}$ dado por

$$\Pi_C(x) \doteq \underset{y \in C}{\operatorname{argmin}} \|x - y\| \quad (1.10)$$

como o operador projeção de x sobre o conjunto C .

Este operador pode nos fornecer uma boa métrica, pois nos garante um artifício para medir a distância de um determinado ponto $x \in \mathbb{X}$ para um conjunto fechado C dado. Visto que

$$\text{dist}(x, C) \doteq \inf_{y \in C} \|x - y\| = \|x - \Pi_C(x)\|.$$

Sabe-se também, que para um conjunto convexo e fechado $C \subseteq \mathbb{X}$ temos que $\Pi_C(x) \in C$ é única. Além que, podemos caracterizá-la, através do ângulo que esta estabelece com todo elemento do conjunto:

Proposição 1.8 (IZMAILOV e SOLODOV, 2005, Teo. 3.2.4). *Seja $C \subseteq \mathbb{X}$, um conjunto convexo e fechado temos que $w = \Pi_C(z)$ se, e somente se,*

$$\langle z - w, c - w \rangle \leq 0$$

para todo $c \in C$.

Proposição 1.9. *Para todo $z \in \mathbb{X}$ vale que*

1. $\langle \Pi_C(z), \Pi_{C^\circ}(z) \rangle = 0$;
2. $\|\Pi_{C^\circ}(z)\| = \inf_{y \in C} \|z - y\|$;

Demonstração. Como $\alpha \Pi_C(z) \in C$ para todo $\alpha > 0$, vale, pela proposição (1.8) que

$$\langle z - \Pi_C(z), \alpha \Pi_C(z) - \Pi_C(z) \rangle \leq 0$$

isto é,

$$(\alpha - 1) \langle z - \Pi_C(z), \Pi_C(z) \rangle \leq 0$$

para valores $\alpha > 1$ e $\alpha < 1$, logo, a única possibilidade é tal que $\langle z - \Pi_C(z), \Pi_C(z) \rangle = 0$. Agora, definindo $w = z - \Pi_C(z)$, perceba que, para todo $c \in C^\circ$,

$$\langle z - w, c - w \rangle = \langle \Pi_C(z), c - w \rangle \leq 0 \quad (1.11)$$

onde, da última desigualdade é dada pela própria definição do polar, logo, de (1.8) conclui-se que $w = \Pi_{C^\circ}(z)$ e assim

$$\langle \Pi_C(z), \Pi_{C^\circ}(z) \rangle = 0 \quad \text{e} \quad \|\Pi_{C^\circ}(z)\| = \|z - \Pi_C(z)\| = \inf_{y \in C} \|z - y\|$$

□

No capítulo (3), será útil saber, dada uma matriz simétrica $M \in \mathbb{S}^m$, qual a matriz, melhor identifica M , no espaço das matrizes simétricas semi-definidas positivas, isto é, como determinar $\Pi_{\mathbb{S}_+^m}(M)$, assim, apresentaremos a seguir um importante resultado para o nosso estudo

Proposição 1.10 (Prop. 2.5 LOURENÇO *et al.*, 2018). *Seja $M \in \mathbb{S}^m$, então $\Pi_{\mathbb{S}_+^m}(M)$ é dado por*

$$\Pi_{\mathbb{S}_+^m}(M) = \sum_{i=1}^m [\lambda_i(M)]_+ u_i(M) u_i(M)^T \quad (1.12)$$

onde $[\lambda]_+ = \max\{0, \lambda\}$ para todo $\lambda \in \mathbb{R}$.

1.4.1 Funções espectrais

Definição 1.7 (Funções simétricas). *Uma função real $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ é dita simétrica, se $f(v) = f(Pv)$ para toda matriz permutação $P \in \mathbb{R}^{n \times n}$. Em particular, o conjunto Ω é dito simétrico se, $P\Omega = \Omega$*

Definição 1.8 (Funções espectrais). *Uma função $f : \Omega \subseteq \mathbb{S}^m \rightarrow \mathbb{R}$ é dita ortogonalmente invariante se, $f(U^T X U) = f(X)$ para toda matriz ortogonal U .*

Note que, uma função ortogonalmente invariante depende unicamente do conjunto dos autovalores de uma matriz, assim, nomearemos tais funções por "Funções Espectrais". Perceba que a função autovalor $X \mapsto \lambda(X) = (\lambda_1, \dots, \lambda_m(X))$ para uma matriz $X \in \mathbb{R}^{m \times m}$ é ortogonalmente invariante.

A classe de funções acima, será bastante útil durante a evolução dos próximos capítulos, desta forma, apresentaremos expressões úteis para a nossa análise. Veja.

Proposição 1.11 (LEWIS, 1996, Teo. 1.1). *Seja $\Omega \subseteq \mathbb{R}^n$ um conjunto aberto e simétrico, suponha que a função $f : \Omega \rightarrow \mathbb{R}$ seja simétrica. Então a função espectral $f(\lambda(M))$ é diferenciável em M se, e somente se, f é diferenciável no vetor $\lambda(M)$. Neste caso o gradiente de $f \circ \lambda$ em M é dado por*

$$(f \circ \lambda)'(M) = U^T (\text{Diag}(\nabla f(\lambda(M)))) U \quad (1.13)$$

para toda matriz ortogonal U tal que $M = U^T (\text{Diag}(\lambda(M))) U$

Note que, a proposição (1.11) nos fornece uma expressão para o gradiente de uma função espectral aplicada a uma matriz, assim, podemos então definir a aplicação derivada aplicada em matriz X como a aplicação $D(f \circ \lambda)(X) : \mathbb{S}^m \rightarrow \mathbb{R}$ tal que

$$Y \mapsto D(f \circ \lambda)(X)[Y] = \langle (f \circ \lambda)'(X), Y \rangle \quad (1.14)$$

1.5 Problemas de Otimização Cônica

Um problema de Otimização Cônica padrão é posto formalmente da seguinte forma:

$$\begin{aligned} & \underset{x \in \mathbb{X}}{\text{Minimizar}} && f(x), \\ & \text{sujeito a} && g(x) \in K \end{aligned} \quad (\text{NCP})$$

De modo que, $f : \mathbb{X} \rightarrow \mathbb{R}$, $g : \mathbb{X} \rightarrow \mathbb{Y}$ são funções continuamente diferenciáveis, \mathbb{X}, \mathbb{Y} são espaços vetoriais de dimensão finita equipados com um produto interno $\langle \cdot, \cdot \rangle$ e uma norma $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$, $K \subset \mathbb{Y}$ um cone, convexo e fechado. Para $x \in \mathbb{X}$, x será dito viável para (NCP) se, e só se, $g(x) \in K$ e, o conjunto de todos os pontos viáveis será denotado por \mathcal{F} .

Sob alguma hipótese de regularidade, dado um ponto $\bar{x} \in \mathcal{F}$, se este for uma solução local para o problema então, existirá um multiplicador de Lagrange $Y \in \mathbb{Y}$ tal que satisfaça a seguinte definição,

Definição 1.9 (KKT). *Seja $\bar{x} \in \mathcal{F}$, este será dito um ponto KKT, se, e somente se, existir $Y \in \mathbb{Y}$ tal que:*

$$\begin{aligned} \nabla f(\bar{x}) + Dg(\bar{x})^*[Y] &= 0 \\ \langle g(\bar{x}), Y \rangle &= 0 \\ Y &\in K^\circ \end{aligned} \quad (1.15)$$

Onde, $Dg(x)^*$ é a matriz adjunta da função derivada $Dg(x) : \mathbb{X} \rightarrow \mathbb{Y}$, de modo que, Y será dito um Multiplicador de Lagrange associado a \bar{x} para (NCP).

Aqui, o gradiente $\nabla f(\bar{x}) \in \mathbb{X}$ é único elemento, tal que, $Df(\bar{x})[d] = \langle \nabla f(\bar{x}), d \rangle$. A primeira equação de (1.9) é dita equação de Lagrange, a segunda equação será dita condição de complementariedade e a terceira é conhecida como viabilidade dual.

Observação 1.7. *Vale lembrar que, dada uma aplicação linear $A : \mathbb{X} \rightarrow \mathbb{Y}$ define-se a Ad-
junta de A (A^*) como a única aplicação linear $A^* : \mathbb{Y} \rightarrow \mathbb{X}$ tal que:*

$$\langle Ax, y \rangle = \langle x, A^*y \rangle \quad (1.16)$$

para todo $(x, y) \in \mathbb{X} \times \mathbb{Y}$.

Para qualquer ponto $x \in \mathcal{F}$ que satisfaça (1.15), este será dito ponto estacionário, ou simplesmente *ponto KKT*. Além disso, para o problema (NCP), denota-se $\Lambda(x, f)$ como o conjunto de todos os multiplicadores de Lagrange associados a x para uma função objetivo f . Como já foi mencionado, qualquer solução local será necessariamente um ponto estacionário, se esta satisfizer uma determinada condição de regularidade, motivando assim nossa próxima definição:

Definição 1.10 (Condição de qualificação). *Uma condição de qualificação (CQ) é uma proposição acerca da descrição analítica do conjunto \mathcal{F} , tal que sob essa proposição, todo minimizador local é ponto estacionário.*

Importantes classes de problemas podem ser postos da forma (NCP), entre eles, um problema não linear de otimização com m restrições de igualdades e p restrições de desigualdades geralmente descrito da seguinte forma

$$\begin{aligned} &\text{Minimizar}_{x \in \mathbb{R}^n} f(x), \\ &\text{sujeito a} \quad h(x) = 0 \\ &\quad \quad \quad g(x) \leq 0 \end{aligned} \quad (\text{PNL})$$

onde $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ são funções continuamente diferenciáveis. Neste caso, definimos o cone $K = \{0\}^m \times -\mathbb{R}_+^p \subset \mathbb{R}^{m+p}$, onde \mathbb{R}_+^p é o conjunto de números reais não negativos, assim neste contexto, seu polar é dado por $K^\circ = \mathbb{R}^m \times \mathbb{R}_+^p$, e as condições KKT para um ponto viável \bar{x} serão

$$\begin{aligned} \nabla f(\bar{x}) + Dh(\bar{x})^T[\lambda] + Dg(\bar{x})^T[\mu] &= 0 \\ \langle g(\bar{x}), \mu \rangle &= 0 \\ (\lambda, \mu) &\in \mathbb{R}^m \times \mathbb{R}_+^p \end{aligned} \quad (1.17)$$

Na literatura, existem diversas condições de qualificações para problemas (PNL), visto que este é um campo de pesquisa bastante vasto e furtivo, assim vale destacar duas das principais condições de qualificação para PNL, a saber, "Condição de Independência Linear - (LICQ)- ou regularidade - e a "Condição de Mangasarian-Fromovitz - (MFCQ)" que podem ser postas da seguinte forma,

Definição 1.11 (LICQ). Dizemos que vale LICQ em um ponto viável \bar{x} se

$$(Dh(\bar{x})^T[\lambda], D\tilde{g}(\bar{x})^T[\mu]) = (0, 0) \Rightarrow (\lambda, \mu) = 0 \quad (1.18)$$

ou, equivalentemente, os vetores

$$\{\nabla h_i(\bar{x}), \nabla g_j(\bar{x}) : i \in \{1, \dots, m\}, j \in A(\bar{x})\} \quad (1.19)$$

são linearmente independentes. Onde $A(\bar{x}) = \{j \in \{1, \dots, p\} : g_j(\bar{x}) = 0\}$ e $D\tilde{g}(\bar{x})^T$ é matriz cuja j -ésima coluna é $\nabla g_j(\bar{x})$ para $j \in A(\bar{x})$.

E de maneira equivalente definimos

Definição 1.12 (MFCQ). Dizemos que vale MFCQ em um ponto viável \bar{x} se os vetores

$$\{\nabla h_i(\bar{x}), \nabla g_j(\bar{x}) : i \in \{1, \dots, m\}, j \in A(\bar{x})\} \quad (1.20)$$

são positivo-linearmente independentes. Onde $A(\bar{x}) = \{j \in \{1, \dots, p\} : g_j(\bar{x}) = 0\}$.

Isto é, a regularidade do problema é determinada exigindo a (positivo)independência linear das colunas das matrizes $Dh(\bar{x})^T$ e $D\tilde{g}(\bar{x})^T$ definidas a partir das restrições analíticas do problema realmente relevantes para a solução local. Desta forma, para o contexto cônico (NCP) de maneira geral, a regularidade, ou independência linear para um ponto $\bar{x} \in \mathcal{F}$, pode ser posta por:

$$\left. \begin{array}{l} DG(\bar{x})^*[\alpha] = 0 \\ \langle G(\bar{x}), \alpha \rangle = 0 \\ \alpha \in K^\circ \end{array} \right\} \Rightarrow \alpha = 0 \quad (1.21)$$

Conhecida, na literatura como "Condição de Qualificação de Robinson- (*Robinson's-CQ*). Veja ROBINSON, 1976b. Outra importante classe de problema cônico, que de certa maneira generaliza o problema não linear (PNL), é estabelecido ao exigir que as restrições sejam postas sobre o cone de segunda ordem (NSOCP) no qual é definido da seguinte forma:

$$\mathbb{L}_m \doteq \begin{cases} \{x \doteq (x_0, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{m-1} : x_0 \geq \|\bar{x}\|\}, & \text{se } m > 1, \\ \mathbb{R}_+, & \text{se } m = 1 \end{cases} \quad (1.22)$$

Outra classe de problemas cônicos bem conhecida é obtida impondo as restrições sobre o cone das matrizes simétricas semi-definida positivas (NSDP) já mencionado anteriormente. Dadas estas ponderações, nosso principal objetivo será o estudo de condições de regularidade para NSOCP e NSDP através do conceito de independência linear tão difundido para PNL, mas com enfoque no contexto cônico.

Durante este texto, será enfatizado o estudo das tão conhecidas condições de qualificação, *Nondegeneracy-CQ* e *Robinson's-CQ*. Entretanto vale ressaltar, que diferente de PNL, estas condições demonstram apresentar um grau de generalidade alto para problemas de grande escala quando tratadas sobre NSDP, ou seja, para estes problemas, as soluções tendem a não satisfazer tais condições regularidade, entretanto, em [ANDREANI, HAESER, MITO, RAMÍREZ e SILVEIRA, 2022](#) e [ANDREANI, HAESER, MITO e RAMÍREZ, 2023](#) os autores, através do conceito de decomposição em autovetores, desenvolveram um interessante abordagem para a independência linear tratada sobre (NSOCP) e (NSDP) de maneira tal, a aproveitar com mais eficácia a estrutura do problema cônico em si, evitando de certa forma, tal generalidade. Assim, os capítulos se sucedem de modo a discutir essencialmente esta nova abordagem com intuito de fornecer aporte teórico e possivelmente novos resultados a partir deste.

Capítulo 2

Nondegeneracy para NSOCP

Neste capítulo, nosso principal interesse será discutir a ideia de independência linear para problemas de otimização não lineares postos sobre o cone de segunda ordem (NSOCP), neste contexto, a condição de qualificação de independência linear é de certa forma, transcrita por meio da conhecida condição de qualificação *Nondegeneracy-CQ*. Com este intuito, inicialmente serão levantadas algumas propriedades algébricas e geométricas para o cone de Segunda Ordem cuja a intenção será estabelecer o conceito de decomposição sobre este cone e então motivar a definição de uma versão fraca para *Nondegeneracy-CQ*, a saber, *Weak-Nondegeneracy*. Ao fim, estabeleceremos algumas condições de equivalências sobre estas definições. A discussão aqui, será feita, inteiramente sobre os resultados obtidos pelo o trabalho [ANDREANI, HAESER, MITO, RAMÍREZ e SILVEIRA, 2022](#) com a intenção de conectar as principais ideias obtidas aqui ao contexto de (NSDP).

2.1 O cone de segunda ordem

Nosso principal objeto de estudo desta seção será o cone de segunda ordem $\mathbb{L}_m \subset \mathbb{R}^m$, (ou *cone de Lorentz*), que usualmente é definido por:

$$\mathbb{L}_m \doteq \begin{cases} \{x \doteq (x_0, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{m-1} : x_0 \geq \|\bar{x}\|\}, & \text{se } m > 1, \\ \mathbb{R}_+, & \text{se } m = 1 \end{cases} \quad (\text{C.L.})$$

Costuma-se, para o cone \mathbb{L}_m , denotar por:

$$bd_+ \mathbb{L}_m \doteq \{x \in \mathbb{L}_m : x_0 = \|\bar{x}\| \text{ e } \bar{x} \neq 0\} \quad (2.1)$$

$$int \mathbb{L}_m \doteq \{x \in \mathbb{L}_m : x_0 > \|\bar{x}\|\} \quad (2.2)$$

respectivamente, a borda de \mathbb{L}_m excluindo-se a origem 0, e o interior de \mathbb{L}_m .

Assim como no espaço das matrizes simétricas semidefinidas positivas, os elementos do cone de *Lorentz* podem ser estruturados por meio de uma decomposição, deste modo, seguindo a abordagem de [ALIZADEH e GOLDFARB, 2003](#) Seção 4, com o intuito de construir

uma álgebra particular para \mathbb{L}_m com propriedades geométricas e algébricas semelhantes ao cone das matrizes simétricas, considere $x = (x_0, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{m-1}$ e $y = (y_0, \bar{y}) \in \mathbb{R} \times \mathbb{R}^{m-1}$, assim pode-se definir a seguinte operação binária $*$: $\mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ dada por:

$$x * y \doteq \begin{pmatrix} \bar{x}^T \bar{y} \\ x_0 y_1 + y_0 x_1 \\ \dots \\ x_0 y_n + y_0 x_n \end{pmatrix} \quad (2.3)$$

Esta multiplicação irá fornecer importantes propriedades algébricas bastante úteis para o nosso estudo:

Proposição 2.1. *Para todo x, y e $z \in \mathbb{R}^m$ vale:*

1. $x * (\alpha y + \beta z) = \alpha x * y + \beta x * z$ and $(\alpha y + \beta z) * x = \alpha y * x + \beta z * x$ para todo $\alpha, \beta \in \mathbb{R}$;
2. $x * y = y * x$;
3. O vetor $e \doteq (1, \mathbf{0})$ é o único elemento identidade: $x * e = x$;
4. $*$ não é associativo para $m > 2$;
5. No produto de k cópias de x : $x * x * \dots * x$, a ordem em que se realiza a multiplicação não importa, isto é, $*$ é associativa para a potência de modo que, $x^p * x^q = x^{p+q}$.

Onde $\mathbf{0}$ denota o vetor nulo de \mathbb{R}^{m-1} .

A verificação destas propriedades se faz imediatamente a partir da definição (2.3), por isto será omitida por agora. As propriedades acima nos garantem uma boa sustentação para trabalhar com a álgebra $(\mathbb{R}^m, *)$ e permitirá fazer interessantes analogias entre o cone \mathbb{L}_m e o cone \mathbb{S}_+^m . Vejamos:

Para todo $x = (x_0, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{m-1}$ podemos escrever:

$$x = \frac{1}{2}(x_0 + \|\bar{x}\|) \begin{pmatrix} 1 \\ \frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix} + \frac{1}{2}(x_0 - \|\bar{x}\|) \begin{pmatrix} 1 \\ -\frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix} \quad (2.4)$$

de modo que, pondo:

$$\lambda_i(x) \doteq x_0 + (-1)^i \|\bar{x}\| \quad \text{e} \quad u_i(x) \doteq \begin{cases} \frac{1}{2} \begin{pmatrix} 1 \\ (-1)^i \frac{\bar{x}}{\|\bar{x}\|} \end{pmatrix}, & \text{se } \bar{x} \neq 0 \\ \frac{1}{2}(1, (-1)^i w) & , \text{c.c.} \end{cases} \quad (2.5)$$

com $w \in \mathbb{R}^{m-1}$ um vetor unitário qualquer e $i \in \{1, 2\}$, podemos então reescrever a equação (2.4) no seguinte formato:

$$x = \lambda_1(x)u_1(x) + \lambda_2(x)u_2(x) \quad (2.6)$$

Comparando as expressões (2.6) e (1.7) nota-se que, podemos interpretar a grosso

modo, os elementos de (2.5) como uma espécie de autovalor e autovetor associados a um elemento de \mathbb{R}^m . Entretanto note que, dada uma matriz simétrica $M \in \mathbb{S}^m$ qualquer, onde:

$$M = U\Lambda U^T = \lambda_1 u_1 u_1^T + \dots + \lambda_n u_n u_n^T \quad (2.7)$$

os elementos que decompõem a matriz M são as matrizes de posto 1, dadas por $u_i u_i^T$, construídas a partir dos autovetores de u_i de M , para $i \in \{1, \dots, m\}$. Além disso, definindo o seguinte polinômio,

$$p(\lambda, x) \doteq \lambda^2 - 2x_0 \lambda + (x_0^2 - \|\bar{x}\|^2) \quad (2.8)$$

verifica-se facilmente que os escalares $\lambda_1(x) = x_0 - \|\bar{x}\|$ e $\lambda_2(x) = x_0 + \|\bar{x}\|$ são raízes para $p(\lambda, x)$, assim o polinômio (2.8) é dito polinômio característico de x .

Observação 2.1. *É válido lembrar que da proposição (1.2) uma matriz é dita semi-definida positiva se, e somente se, seus autovalores são não-negativos, isto é, $\lambda_i \geq 0$ para todo $i \in \{1, \dots, n\}$, sendo assim, perceba que da definição (C.L.) dado um elemento $x \in \mathbb{R}^m$, à luz da decomposição (2.4) garante-se que, $x \in \mathbb{L}_m$ se, e somente se, $\lambda_1(x) \geq 0$ e $\lambda_2(x) \geq 0$.*

Perceba também que, da decomposição estabelecida acima para o cone de segunda ordem, para os vetores que decompõem $x \in \mathbb{R}^m$ $u_1(x)$ e $u_2(x)$ em (2.6) vale as seguintes relações:

$$u_1(x) * u_2(x) = 0, \quad (2.9)$$

$$u_1(x)^2 = u_1(x) \quad \text{e} \quad u_2(x)^2 = u_2(x), \quad (2.10)$$

$$u_1(x) + u_2(x) = e, \quad (2.11)$$

$$u_1(x) = R u_2(x) \quad \text{e} \quad u_2(x) = R u_1(x) \quad (2.12)$$

$$u_1(x), u_2(x) \in bd \mathbb{L}_m \quad (2.13)$$

onde

$$R \doteq \begin{bmatrix} 1 & 0 \\ 0 & -\mathbb{I}_{m-1} \end{bmatrix} \quad (2.14)$$

com \mathbb{I}_{m-1} a matriz identidade de ordem $m - 1$.

Para qualquer par de vetores $\{u_1, u_2\}$ que satisfazem as propriedades (2.9), (2.10) e (2.11) dizemos que estes constituem um *Jordan Frame*.

Observação 2.2. *Vale notar, que as propriedades postas sobre o elementos $u_1(x)$ e $u_2(x)$ da decomposição (2.4), também são estabelecidas para a decomposição espectral de uma matriz simétrica (2.7). Pois, definindo a operação simétrica*

$$X * Y \doteq \frac{1}{2}(XY + YX) \quad (2.15)$$

podemos verificar, que as matrizes de posto 1 $u_i u_i^T$ em (2.7) constituem também um *Jordan Frame*. Esta é uma característica intrínseca à cones simétricos, como pode ser visto na próxima proposição

Proposição 2.2 (FARAUT, 1994, Teo. 3.1.2). *Seja $(\mathcal{E}, *)$ uma álgebra de Jordan Euclidiana e seja $y \in \mathcal{E}$. Então, existem idempotentes primitivos c_1, \dots, c_r satisfazendo:*

$$c_i * c_j = 0, \quad \text{para } i \neq j \quad (2.16)$$

$$c_i * c_i = c_i \quad i = 1, \dots, r \quad (2.17)$$

$$c_1 + \dots + c_r(x) = e, \quad i = 1, \dots, r \quad (2.18)$$

$$(2.19)$$

e, únicos números reais $\sigma_1, \dots, \sigma_r$ tais que

$$y = \sigma_1 c_1 + \dots + \sigma_r c_r$$

Terminaremos esta seção, apresentando uma representação geométrica para os elementos de \mathbb{L}_m , em vista das propriedades (2.13) e (2.12), em conjunto com a decomposição espectral (2.4).

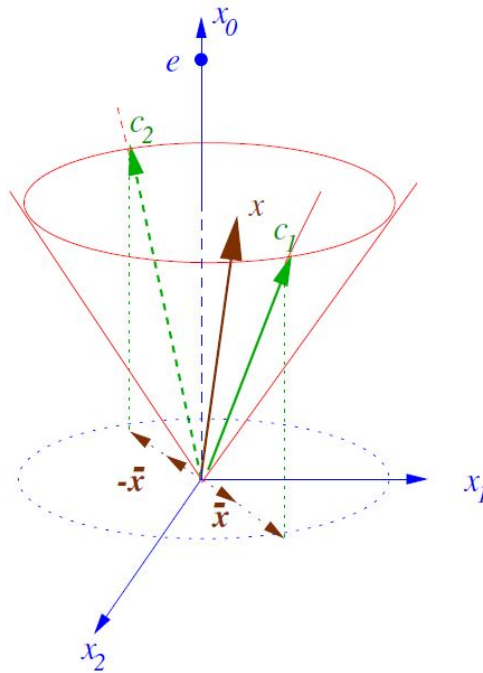


Figura 2.1: Decomposição em \mathbb{L}_m

2.2 O problema sobre o cone de segunda ordem

O problema de otimização, com múltiplas restrições cônicas de segunda ordem, é frequentemente posto da seguinte forma:

$$\begin{aligned} & \underset{x \in \mathbb{R}^n}{\text{Minimizar}} && f(x), \\ & \text{sujeito a} && g_j(x) \in \mathbb{L}_{m_j}, \quad \forall j \in \{1, \dots, q\}, \end{aligned} \tag{NSOCP}$$

Onde, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $g_j : \mathbb{R}^n \rightarrow \mathbb{R}^{m_j}$ são funções contínuas e diferenciáveis para todo $j \in \{1, \dots, q\}$ com $\mathbb{L}_{m_j} \subset \mathbb{R}^{m_j}$. Usualmente, para um ponto $x \in \mathbb{R}^n$, modificando a notação, as restrições cônicas serão denotadas por $g_j(x) = (g_{j,0}(x), \hat{g}_j(x)) \in \mathbb{R} \times \mathbb{R}^{m_j-1}$. Denotaremos o conjunto de todos os pontos viáveis de (NSOCP) por \mathcal{F} . Como de usual, para $x \in \mathcal{F}$ oriundo das equações (2.1), (2.2) costuma-se particionar o conjunto $\{1, \dots, q\}$ da seguinte forma:

$$\begin{aligned} I_0(x) &\doteq \{j \in \{1, \dots, q\}; g_j(x) = 0\}, \\ I_B(x) &\doteq \{j \in \{1, \dots, q\}; g_j(x) \in bd_+ \mathbb{L}_{m_j}\}, \\ I_{int}(x) &\doteq \{j \in \{1, \dots, q\}; g_j(x) \in int \mathbb{L}_{m_j}\} \end{aligned}$$

Uma importante abordagem para o estudo de problemas de otimização, consiste na reformulação do problema para um contexto tecnicamente mais simples, para a otimização cônica, reformulamos o problema, utilizamos uma técnica de redução de cones. Esta abordagem nos rende um estudo mais qualificado sobre o conjunto viável por meio de parametrizações e de modo a facilitar o estudo das condições de regularidade para o conjunto viável. Veja [BONNANS e SHAPIRO, 2000, sec. 3.4.4] para um estudo mais aprofundado.

Definição 2.1 (Redução de cones). *Sejam \mathbb{Y} e \mathbb{Z} espaços de dimensão finita, sejam $K \subseteq \mathbb{Y}$ e $C \subseteq \mathbb{Z}$ conjuntos convexos e fechados. Diremos que K é C^l -redutível para o conjunto C em um ponto $y^* \in K$ se, existe uma vizinhança V de y^* e uma aplicação $\psi : V \rightarrow \mathbb{Z}$, l -vezes continuamente diferenciável (C^l), tal que:*

1. Para todo $y \in V$ então, $y \in K$ se, e somente se, $\psi(y) \in C$;
2. $D\psi(y^*) : \mathbb{Y} \rightarrow \mathbb{Z}$ é sobrejetiva;

Rudemente falando, um conjunto convexo e fechado é redutível a outro em um dado ponto \bar{y} , se este pode ser parametrizado localmente em torno de y^* através do outro. A definição (2.1) nos trará uma nova abordagem sobre (NSOCP) pois, a a partir desta conseguiremos reduzir o cone \mathbb{L}_m em qualquer ponto y^* de \mathbb{L}_m de maneira a simplificar a construção de condições de qualificação para o problema.

Proposição 2.3. *O Cone de segunda ordem \mathbb{L}_m é redutível em todo ponto $y^* \in \mathbb{L}_m$.*

Demonstração.

- Se $y^* = 0$, reduzimos \mathbb{L}_m para o próprio $L = \mathbb{L}_m$ tomando $\psi(y) = y$;

- Se $y \in \text{int } \mathbb{L}_m(y_0^* > \|\bar{y}^*\|)$, reduzimos \mathbb{L}_m para o cone $L = \{0\}$ tomando $\phi(y) = 0$;
- Se $y^* \in \text{bd}_+ \mathbb{L}_m(y_0^* = \|\bar{y}^*\|)$, reduzimos \mathbb{L}_m para o cone $L = \mathbb{R}_-$ tomando $\phi(y) = \|\bar{y}\| - y_0$. \square

Por meio da abordagem de redução acima, podemos então agora, estudar duas importantes condições de qualificação na literatura de (NSOCP), conhecidas por *Nondegeneracy-CQ* e *Robinson's CQ*. Entretanto é válido lembrar, que na literatura, sob a ótica da otimização cônica, esta pode ser definida de diversas formas, contudo para o contexto de (NSOCP) as definições apresentam-se semelhantes.

Definição 2.2 (*Nondegeneracy*). *Considere o problema:*

$$\begin{aligned} & \underset{x \in \mathbb{X}}{\text{Minimizar}} && f(x), \\ & \text{sujeito a} && g(x) \in K \end{aligned} \tag{CP}$$

onde f, g são funções duas vezes continuamente diferenciáveis, \mathbb{X}, \mathbb{Y} e \mathbb{Z} são espaços de dimensão finita e $K \subseteq \mathbb{Y}$ um cone convexo fechado, redutível ao cone convexo fechado $C \subseteq \mathbb{Z}$ em um ponto $y^* = g(x^*) \in K$ por meio de ψ . O ponto viável x^* é dito não-degenerado (com respeito a ψ) quando a derivada da função $\mathcal{A} \doteq \psi \circ g$ é sobrejetiva.

Assim, utilizando o Lema (2.3) e a definição (2.2), poderemos caracterizar "Nondegeneracy-CQ" para (NSOCP) da seguinte maneira

Proposição 2.4 (*Nondegeneracy-CQ*). *Seja $\bar{x} \in \mathcal{F}$ para (NSOCP) \bar{x} satisfaz Nondegeneracy-CQ se:*

- A seguinte família

$$\{Dg_j(\bar{x})^T R_j g_j(\bar{x})\}_{j \in I_B(\bar{x})} \cup \{Dg(\bar{x})^T\}_{I_0(\bar{x})} \tag{2.20}$$

é K -linearmente independente, onde $K = \mathbb{R}^{|I_B(\bar{x})|} \times \prod_{j \in I_0(\bar{x})} \mathbb{R}^{m_j}$.

Demonstração. Seja $\bar{x} \in \mathcal{F}$ e $j \in I_0(\bar{x})$ isto é, $g_j(\bar{x}) = 0$, à luz do Lema 2.3 e da definição (2.2), pondo $A_j : \mathbb{R}^n \rightarrow \mathbb{R}^m$ dada por $A_j = \psi \circ g_j$ com $\psi(y) = y$, teremos que \bar{x} satisfaz *Nondegeneracy-CQ* se a aplicação $DA_j(\bar{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ for sobrejetiva. Então, com o auxílio da regra da cadeia têm-se:

$$DA_j(\bar{x}) = D\psi(g_j(\bar{x}))Dg_j(\bar{x}) = \mathbb{I}_n Dg_j(\bar{x}) = Dg_j(\bar{x})$$

assim, sabe-se que $DA_j(\bar{x}) = Dg_j(\bar{x})$ é sobrejetiva, se e só se, possui posto linha completo, isto é, pela definição (1.5) $Dg_j(\bar{x})^T$ precisa ser \mathbb{R}_{m_j} -Linearmente Independente. De maneira análoga, considerando $j \in I_B(\bar{x})$ (ou seja, $g_{j,0}(\bar{x}) = \hat{g}_j(\bar{x})$) reduziremos o cone por meio de $\psi(y) = \|\bar{y}\| - y_0$, assim

$$\nabla\psi(y) = \left(-1, \frac{\bar{y}}{\|\bar{y}\|}\right) = \frac{1}{y_0} (y_0, -\bar{y})$$

ou utilizando (2.14), $\nabla\psi(y) = \frac{1}{y_0} Ry$, daí, teremos:

$$DA_j(\bar{x}) = D\psi(g_j(\bar{x}))Dg_j(\bar{x}) = Dg_j(\bar{x})^T \frac{1}{g_{j,0}(\bar{x})} Rg_j(\bar{x})$$

Ou seja, a sobrejetividade de $DA_j(\bar{x})$ é dada pelo vetor $Dg_j(\bar{x})^T Rg_j(\bar{x})$. Por fim, para $j \in I_{int}(\bar{x})$ reduzimos \mathbb{L}_m por $\psi(y) = 0$ em $L = \{0\}$, que é trivialmente irrelevante para a análise da *Nondegeneracy*. Assim, posto a redução acima em conjunto, e considerando a observação (1.3) têm-se que \bar{x} satisfaz *Nondegeneracy* se, e só se, \bar{x} satisfazer (2.20). \square

2.3 Nondegeneracy e Bases Paramétricas

Da proposição (2.4) pode-se observar que, avaliar *Nondegeneracy-CQ* em um dado ponto viável de (NSOCP) consiste intrinsecamente em estabelecer uma independência linear cônica para um determinado conjunto que, essencialmente depende da derivada das funções restrições. Entretanto, como é feito em [ANDREANI, HAESER, MITO, RAMÍREZ e SILVEIRA, 2022](#), a avaliação desta independência linear pode ser posta sobre um conjunto com dimensão menor porém, com um sacrifício de um número maior de avaliações. Assim, motivados pela observação (1.4) e explorando o conceito de bases paramétricas dado na seção (1.2) podemos obter a seguinte caracterização para *Nondegeneracy-CQ*:

Proposição 2.5. *Seja \bar{x} um ponto viável de (NSOCP) então, \bar{x} satisfaz "Nondegeneracy-CQ" se, e somente se, a família de vetores*

$$\{Dg_j(\bar{x})^T u_1(g_j(\bar{x}))\}_{j \in I_B(\bar{x})} \cup \{Dg_j(\bar{x})^T(1, -w_j), Dg_j(\bar{x})^T(1, w_j)\}_{j \in I_0(\bar{x})} \quad (2.21)$$

é linearmente independente para todo $w_j \in \mathbb{R}^{m_j-1}$ tal que $\|w_j\| = 1$, $j \in I_0(\bar{x})$.

Demonstração. Considere o produto $C = \prod_{j \in J} C_j$, $J \doteq I_B(\bar{x}) \cup I_0(\bar{x})$, de modo que

$$C_j \doteq \begin{cases} \mathbb{R} & , \text{ se } j \in I_B(\bar{x}) \\ \mathbb{R}^{m_j-1} & , \text{ se } j \in I_0(\bar{x}) \end{cases}$$

desta forma, com o auxílio das observações (1.3) e (1.4) tomando a matriz bloco $M = [M_j]_{j \in J}$ disposta como em (1.6) cujos blocos são dados por:

$$M_j \doteq \begin{cases} Dg_j(\bar{x})^T u_1(g_j(\bar{x})) & , \text{ se } j \in I_B(\bar{x}) \\ Dg_j(\bar{x})^T & , \text{ se } j \in I_0(\bar{x}) \end{cases}$$

assim, na proposição (2.4), a família (2.20) é K -linearmente se, e somente se, a matriz M é C -linearmente independente, ou seja, (2.21) é linearmente independente para todo $w_j \in \mathbb{R}^{m_j-1}$ tal que $\|w_j\| = 1$, $j \in I_0(\bar{x})$. \square

Apesar do conjunto (2.21) ser essencialmente menor para se analisar a independência linear, aumentamos o número de análises ao trabalhar com a base paramétrica $S = \{w_j \in \mathbb{R}^{m_j-1} : \|w_j\| = 1\}$ para $j \in I_0(\bar{x})$. Do ponto de vista teórico, torna-se coerente, visto que, o

cone de segunda ordem, pode ser posto na forma

$$\mathbb{L}_m = \bigcup_{\substack{w \in \mathbb{R}^{m-1} \\ \|w\|=1}} \text{cone}(\{(1, -w), (1, w)\}). \quad (2.22)$$

Entretanto, como será visto, toda a avaliação sobre a base S pode ser desnecessariamente generalista para alguns exemplos que possuem estruturas específicas. Com isto, observamos que é possível refinar ainda mais a análise da *Nondegeneracy-CQ* de modo a aproveitar melhor a estrutura do problema. Seguiremos com um exemplo motivador:

Exemplo 2.1 (ANDREANI, HAESER, MITO, RAMÍREZ e SILVEIRA, 2022, Ex.: 3.1). *Sejam $g_0, g_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ funções continuamente diferenciáveis. Definindo $g(x) = (g_0(x), g_1(x))$ a função restrição para o problema NSOCP ($g(x) \in \mathbb{L}_2$) considere \bar{x} tal que:*

- $g(\bar{x}) = 0$
- $\{\nabla g_0(\bar{x}), \nabla g_1(\bar{x})\}$ são linearmente independentes.

Assim, note que, \bar{x} satisfaz *Nondegeneracy-CQ* visto que, $Dg(\bar{x})^T$ é \mathbb{R}^2 -linearmente independente. Por outro olhar, considere agora o problema equivalente (NSOCP) cuja função restrição é dada por:

$$\tilde{g}(\bar{x}) = (g_0(x), g_1(x), 0, \dots, 0) \in \mathbb{L}_m$$

Agora, para o problema reformulado, utilizando a proposição (2.5) estabelecemos *Nondegeneracy* por meio da independência linear dos seguintes vetores:

$$D\tilde{g}(\bar{x})^T(1, -\bar{w}) = \nabla g_0(\bar{x}) - \bar{w}_1 \nabla g_1(\bar{x}), \quad e \quad D\tilde{g}(\bar{x})^T(1, \bar{w}) = \nabla g_0(\bar{x}) + \bar{w}_1 \nabla g_1(\bar{x})$$

para todo $\bar{w} = (\bar{w}_1, \dots, \bar{w}_{m-1})$ com $\|\bar{w}\| = 1$. Observe que, *Nondegeneracy* falha para $\bar{w}_1 = 0$.

Note que, $\tilde{g}(x)$ pode admitir inúmeras decomposições de autovetores do tipo $\frac{1}{2}(1, w)$ e $\frac{1}{2}(1, -w)$ entretanto, para cada x tal que $g_1(x) \neq 0$, utilizando (2.5) podemos determinar, unicamente os autovetores de $\tilde{g}(x)$ por:

$$u_1(\tilde{g}(x)) = \frac{1}{2} \left(1, -\frac{g_1(x)}{\|g_1(x)\|}, 0, \dots, 0 \right)^T \quad e \quad u_2(\tilde{g}(x)) = \frac{1}{2} \left(1, \frac{g_1(x)}{\|g_1(x)\|}, 0, \dots, 0 \right)^T.$$

Isto nos faz pensar, que apesar das muitas decomposições possíveis para $\tilde{g}(x)$, as únicas relevantes para a estrutura do problemas são tais que $\bar{w} = (\pm 1, 0, \dots, 0)$. De fato, seguimos por meio de uma análise local em torno de \bar{x} , isto é, considere qualquer sequência $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$, deste modo, temos os casos:

- Se $g_1(x^k) \neq 0$ para cada $k \in \mathbb{N}$, podemos definir unicamente as sequências $\{u_1(\tilde{g}(x^k))\}$ e $\{u_2(\tilde{g}(x^k))\}$ dadas por meio da decomposição, de modo que os pontos $\frac{1}{2}(1, \bar{w})$ e $\frac{1}{2}(1, -\bar{w})$ tais que $\bar{w}_1 = 0$ não fazem parte dos seus pontos limites.
- Se $g_1(x^k) = 0$ para algum índice $k \in \mathbb{N}$, podemos então escolher convenientemente a decomposição de $\tilde{g}(x^k)$ de maneira tal, a evitar os pontos $\frac{1}{2}(1, \bar{w})$ e $\frac{1}{2}(1, -\bar{w})$ e $\bar{w}_1 = 0$ como pontos

limites.

Assim, para toda sequência $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$, avaliar a "Nondegeneracy-CQ" localmente em torno de \bar{x} , se dá pela análise sobre os pontos limites $\bar{w} = (\pm 1, 0, \dots, 0)$, assim teremos os vetores

$$D\tilde{g}(\bar{x})^T(1, -\bar{w}) = \nabla g_0(\bar{x}) \mp \nabla g_1(\bar{x}), \quad e \quad D\tilde{g}(\bar{x})^T(1, \bar{w}) = \nabla g_0(\bar{x}) \pm \nabla g_1(\bar{x})$$

. que são linearmente independentes.

De acordo com os autores, o exemplo anterior nos induz a pensar que, exigir a independência linear de (2.21) para todo \bar{w}_j pode ser desnecessariamente forte para uma condição de qualificação. Com isto, somos motivados a analisar a construir uma condição de qualificação estruturada em Nondegeneracy-CQ, de maneira tal, a ser avaliada em um conjunto relativamente menor de vetores, preservando as informações da estrutura do problema. Assim, podemos definir:

Definição 2.3 (Weak-Nondegeneracy). *Seja \bar{x} um ponto viável para (NSOCP). Dizemos que \bar{x} satisfaz Weak-Nondegeneracy se, para cada sequência $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$ existe $I \subseteq_{\infty} \mathbb{N}$ e sequências de autovetores convergentes $\{u_1(g_j(x^k))\}_{k \in I} \rightarrow \frac{1}{2}(1, -\bar{w}_j)$ e $\{u_2(g_j(x^k))\}_{k \in I} \rightarrow \frac{1}{2}(1, \bar{w}_j)$, para cada $j \in I_0(\bar{x})$ de maneira tal que (2.21) é linearmente independente sobre estes pontos limites.*

Pela proposição (2.5) e pela definição (2.3) acima, perceba que, para $x \in \mathcal{F}$, se x satisfizer Nondegeneracy então x satisfaz weak-Nondegeneracy, dado que os pontos limites das sequências de autovetores estão contidos na base paramétrica $S = \{w_j \in \mathbb{R}^{m-1} : \|w_j\| = 1\}$ entretanto, vamos nos perguntar, se o contrário é necessariamente verdadeiro, veja a seguir

Exemplo 2.2 (ANDREANI, HAESER, MITO, RAMÍREZ e SILVEIRA, 2022, Ex.: 3.2). *Considere a seguinte restrição para o problema (NSOCP)*

$$g(x) \doteq (x_1, x_2, x_2) \in \mathbb{L}_3$$

Dado o ponto viável $\bar{x} = (0, 0)$ veja que este, não satisfaz Nondegeneracy-CQ. Veja por exemplo, que para $\bar{w} \doteq \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$ ou $\bar{w} \doteq \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ teremos

$$Dg(\bar{x})^T(1, -\bar{w}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad e \quad Dg(\bar{x})^T(1, \bar{w}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Por outro lado, podemos evitar tais pontos, considerando os pontos limites das sequências de autovetores. Assim, para avaliar Weak-Nondegeneracy considere uma sequência qualquer $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$ com isto temos três possíveis casos a considerar:

- Existe $I \subseteq_{\infty} \mathbb{N}$ tal que $x_2^k > 0$ para todo $k \in I$;
- O caso 1 falha, porém existe $I \subseteq_{\infty} \mathbb{N}$ tal que $x_2^k < 0$ para todo $k \in I$;
- Os casos anteriores falham e teremos $x_2^k = 0$ para k suficientemente grande;

Estudando o caso 1, note que, por meio de (2.4), $g(x^k)$ pode ser determinado por:

$$g(x^k) = 1/2(x_1^k - \sqrt{2}x_2^k) \begin{pmatrix} 1 \\ -1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} + 1/2(x_1^k + \sqrt{2}x_2^k) \begin{pmatrix} 1 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

logo, seus autovetores são unicamente determinados por

$$u_1(g(x^k)) = \frac{1}{2} \left(1, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)^T \quad e \quad u_2(g(x^k)) = \frac{1}{2} \left(1, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)^T$$

para todo $k \in I$. Com isto, definindo $\bar{w} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$, veja que

$$\lim_{k \in I} u_1(g(x^k)) = \frac{1}{2}(1, -\bar{w}) \quad e \quad \lim_{k \in I} u_2(g(x^k)) = \frac{1}{2}(1, \bar{w})$$

Assim, avaliando weak-Nondegeneracy nestes pontos limites teremos

$$Dg(\bar{x})^T(1, -\bar{w}) = \frac{1}{2} \begin{pmatrix} 1 \\ -\sqrt{2} \end{pmatrix} \quad e \quad Dg(\bar{x})^T(1, \bar{w}) = \frac{1}{2} \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}$$

que são linearmente independentes. O caso 2 é inteiramente análogo, com um ajuste apenas no sinal. Para o caso 3, de (2.5) temos que a decomposição de $g(x^k)$ não é única, e a partir da definição de Weak-Nondegeneracy podemos tomar uma decomposição conveniente para cada $g(x^k)$, escolhendo assim a mesma decomposição dos casos anteriores, de modo que \bar{x} satisfaz Weak-Nondegeneracy.

Como é visto em [ANDREANI, HAESER, MITO, RAMÍREZ e SILVEIRA, 2022](#), neste contexto de (NSOCP), Nondegeneracy demonstra-se não ser equivalente a Weak-Nondegeneracy, contudo, adicionando uma hipótese sobre a derivada das restrições conseguimos obter uma nova forma de estabelecer uma relação direta de equivalência entre Weak-Nondegeneracy e Nondegeneracy, vejamos:

Proposição 2.6. *Seja $\bar{x} \in \mathcal{F}$ para NSOCP SOCP tal que $g(\bar{x}) = 0$ com \bar{x} satisfazendo weak-Nondegeneracy, além disso suponha que:*

- Para todo $w \in \mathbb{R}^{m-1}$ com $\|w\| = 1$, existe $d \in \mathbb{R}^n$ tal que $Dg(\bar{x})d = \lambda_1(1, -w) + \lambda_2(1, w)$ com $\lambda_1 \neq \lambda_2$.

Então, \bar{x} satisfaz Nondegeneracy.

Demonstração. A prova será essencialmente por contradição. Supomos que \bar{x} não satisfaz Nondegeneracy. Isto é, existe um vetor unitário $\bar{w} \in \mathbb{R}^{m-1}$ tal que

$$\{Dg(\bar{x})^T(1, -\bar{w}), Dg(\bar{x})^T(1, \bar{w})\} \quad (1)$$

são linearmente independentes. Seja d tal que $Dg(\bar{x})d = \lambda_1(1, -\bar{w}) + \lambda_2(1, \bar{w})$ com $\lambda_1 \neq \lambda_2$. Defina $x^k \doteq \bar{x} + t_k d$ para qualquer $t_k \rightarrow 0^+$. Deste modo,

$$g(x^k) = g(\bar{x}) + t_k Dg(\bar{x})d + o(t_k).$$

Desde que $g(\bar{x}) = 0$, segue-se que

$$\begin{aligned} (g_0(x^k), \hat{g}(x^k)) &= t_k(\lambda_1(1, -\bar{w}) + \lambda_2(1, \bar{w})) + o(t_k) \\ \hat{g}(x^k) &= t_k(\lambda_2 - \lambda_1)\bar{w} + o(t_k) \end{aligned}$$

assim

$$\frac{\hat{g}(x^k)}{\|\hat{g}(x^k)\|} = \frac{(\lambda_2 - \lambda_1)\bar{w} + o(t_k)/tk}{\|(\lambda_2 - \lambda_1)\bar{w} + o(t_k)/tk\|} \rightarrow \bar{w}$$

e como $\bar{w} \in \mathbb{R}^{m-1}$ com $\|\bar{w}\| = 1$, é tal que (1) é linearmente dependente, concluímos que, \bar{x} não satisfaz *weak-nondegeneracy*, o que contradiz nossa hipótese, logo, \bar{x} satisfaz *Nondegeneracy*. \square

Perceba que, pela proposição acima, para obtermos uma equivalência direta entre *Nondegeneracy-CQ* e *Weak-Nondegeneracy-CQ* é necessário exigir uma decomposição única para $Dg(\bar{x})d$. Esta equivalência já foi estabelecida em [ANDREANI, HAESER, MITO, RAMÍREZ e SILVEIRA, 2022](#) entretanto, os autores exigiram uma hipótese de sobrejetividade sobre a derivada das restrições em adição com *Weak-Nondegeneracy*. O próximo resultado mostrará uma conexão entre a sobrejetividade da derivada das restrições com a diagonalização exigida em (2.6).

Proposição 2.7. *Sejam $g(x) = (g_0(x), \hat{g}(x)) \in \mathbb{R} \times \mathbb{R}^{m-1}$ e matriz*

$$D\hat{g}(x) \doteq \begin{bmatrix} \cdots \\ \nabla g_j(x) \\ \cdots \end{bmatrix}_{j \in \{1, \dots, m-1\}}$$

de modo que as linhas da matriz derivada $D\hat{g}(x)$ sejam correspondentes aos vetores gradientes de $\hat{g}(x)$. Então, $D\hat{g}(x)$ é sobrejetiva se, e somente se, para todo $w \in \mathbb{R}^{m-1}$ com $\|w\| = 1$, existe $d \in \mathbb{R}^n$ tal que $Dg(\bar{x})d = \lambda_1(1, -w) + \lambda_2(1, w)$ com $\lambda_1 \neq \lambda_2$.

Demonstração. Primeiramente, suponha que $D\hat{g}(x)$ seja sobrejetiva, desta forma, seja um vetor unitário $w \in \mathbb{R}_{m-1}$ qualquer, por hipótese existe uma direção $d \in \mathbb{R}^n$ tal que $D\hat{g}(x)d = w$. Com isto, basta definir

$$y \doteq Dg(x)d = (\nabla g_0(x)^T d, D\hat{g}(x)d) = (y_0, w)$$

Deste modo, utilizando a decomposição (2.4) têm-se:

$$\begin{aligned} Dg(x)d &= (y_0, w) \\ &= \frac{1}{2}(y_0 - 1) \begin{pmatrix} 1 \\ -w \end{pmatrix} + \frac{1}{2}(y_0 + 1) \begin{pmatrix} 1 \\ w \end{pmatrix} \\ &= \lambda_1(1, -w) + \lambda_2(1, w) \end{aligned}$$

como queríamos mostrar. Por outro lado, suponha que para todo $w \in \mathbb{R}_{m-1}$ unitário, exista uma direção $d \in \mathbb{R}^n$ tal que $Dg(x)d = \lambda_1(1, -w) + \lambda_2(1, w)$ com $\lambda_1 \neq \lambda_2$. Isto é,

$$(\nabla g_0(x)^T d, D\hat{g}(x)d) = (\lambda_1 + \lambda_2, (\lambda_2 - \lambda_1)w)$$

ou seja, para qualquer $w \in \mathbb{R}_{m-1}$ com $\|w\| = 1$ têm-se que existe $d \in \mathbb{R}^n$ tal que

$$D\hat{g}(x)d = (\lambda_2 - \lambda_1)w$$

assim, para provar a sobrejetividade de $D\hat{g}(x)d$, considere $y \in \mathbb{R}^{m-1}$ qualquer. Se $y = 0$ basta tomar $d = 0$, por outro lado, se $y \neq 0$ escrevemos $w = y/\|y\|$, com isto, existem d , λ_1 e λ_2 tal que

$$D\hat{g}(x)d = (\lambda_2 - \lambda_1) \frac{y}{\|y\|}$$

logo, escrevendo

$$\bar{d} = \frac{d\|y\|}{\lambda_2 - \lambda_1}$$

teremos $D\hat{g}(x)\bar{d} = y$. Portanto $D\hat{g}(x)$ é sobrejetiva. \square

Veja que, das proposições (2.6) e (2.7) originalmente neste trabalho, obtemos uma nova caracterização da Proposição 3.1 de [ANDREANI, HAESER, MITO, RAMÍREZ e SILVEIRA, 2022](#) cuja é posta por:

Proposição 2.8 ([ANDREANI, HAESER, MITO, RAMÍREZ e SILVEIRA, 2022](#), Prop.: 3.1). *Seja $\bar{x} \in \mathcal{F}$ para (NSOCP), \bar{x} satisfaz Nondegeneracy se, e somente se, \bar{x} satisfaz weak-Nondegeneracy e a matriz*

$$D\hat{g}(x) \doteq \begin{bmatrix} \cdots \\ \nabla g_j(x) \\ \cdots \end{bmatrix}_{j \in \{1, \dots, m-1\}}$$

é sobrejetiva.

A proposição (2.6) nos será bastante útil ao estudarmos estes mesmos conceitos sobre *Nondegeneracy-CQ* e *Weak-Nondegeneracy-CQ* voltados para o contexto de programação semidefinida, cujo apresenta curiosas semelhanças com a programação posta sobre o cone de segunda ordem.

Capítulo 3

Nondegeneracy para NSDP

3.1 O Problema Não Linear Semidefinido Positivo

O nosso principal interesse neste capítulo é estudar como estender os resultados obtidos no capítulo (2) para *Nondegeneracy-CQ* e *Weak-Nondegeneracy* para o contexto de um problema de otimização não linear semidefinido positivo (NSDP). Isto é, para problemas do tipo

$$\begin{aligned} & \underset{x \in \mathbb{R}^n}{\text{Minimize}} && f(x), \\ & \text{sujeito a} && G(x) \succeq 0 \end{aligned} \tag{NSDP}$$

De modo que, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $G : \mathbb{R}^n \rightarrow \mathbb{S}^m$ são funções continuamente diferenciáveis, \mathbb{S}^m é o espaço das matrizes simétricas, onde \succeq é uma relação parcial induzida sobre o cone convexo, fechado das matrizes $m \times m$ simétricas semi-definida positivas \mathbb{S}_+^m . Uma caracterização para esta relação é dada por:

$$M \succeq N \Leftrightarrow M - N \in \mathbb{S}_+^m.$$

O conjunto viável para (NSDP) será denotado por $\mathcal{F} \doteq G^{-1}(\mathbb{S}_+^m)$ assim como no capítulo (2) ficando subentendido qual problema está sendo tratado pelo o contexto.

Definição 3.1. *Seja $x \in \mathbb{R}^n$ a derivada de $G(x)$ é a aplicação linear $DG(x) : \mathbb{R}^n \rightarrow \mathbb{S}^m$ definida pela seguinte ação (em relação a base canônica):*

$$d \mapsto DG(x)[d] \doteq \sum_{i=1}^n D_{x_i} G(x) d_i \tag{3.1}$$

Para todo $d \in \mathbb{R}^n$, onde $D_{x_i} G(x) \in \mathbb{S}^m$ é a derivada parcial de G em relação a variável x_i em $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

Veamos que, dada a aplicação $DG(x) : \mathbb{R}^n \rightarrow \mathbb{S}^m$ para o cálculo da sua adjunta

$DG(x)^* : \mathbb{S}^m \rightarrow \mathbb{R}^n$ temos, para $d \in \mathbb{R}^n$ e $Y \in \mathbb{S}^m$ vale

$$\begin{aligned} d^T DG(x)^*[Y] &= \langle DG(x)[d], Y \rangle \\ &= \left\langle \sum_{i=1}^n D_{x_i} G(x) d_i, Y \right\rangle \end{aligned} \quad (3.2)$$

onde para simplificar a notação escreveremos o vetor $DG(x)^*[Y] = D^*$ e a matriz $D_{x_i} G(x) = DG_{x_i}$ assim, de (3.2) têm-se

$$\begin{aligned} d^T D^* &= \left\langle \sum_{i=1}^n d_i DG_{x_i}, Y \right\rangle \\ &= \sum_{i=1}^n d_i \langle DG_{x_i}, Y \rangle \end{aligned}$$

logo concluímos que as componentes de D^* serão dadas por $D_i^* = \langle D_{x_i} G(x), Y \rangle$ isto é,

$$\begin{aligned} DG(x)^*[Y] &= \left[\langle DG_{x_1}, Y \rangle, \dots, \langle DG_{x_n}, Y \rangle \right]^T \\ &= \left[\sum_{i,j=1}^m \partial_{x_1} g_{ij}(x) y_{ij}, \dots, \sum_{i,j=1}^m \partial_{x_n} g_{ij}(x) y_{ij} \right]^T \\ &= \sum_{i,j=1}^m \nabla g_{ij}(x) y_{ij} \end{aligned} \quad (3.3)$$

Assim, definimos

Definição 3.2. Seja $DG(x) : \mathbb{R}^n \rightarrow \mathbb{S}^m$ a aplicação derivada de G aplicada em x define-se a adjunta $DG(x)^* : \mathbb{S}^m \rightarrow \mathbb{R}^n$ como a aplicação tal que

$$Y \mapsto DG(x)^*[Y] = \sum_{i,j=1}^m \nabla g_{ij}(x) y_{ij} \quad (3.4)$$

Na proposição a seguir, analisaremos a proposição (1.11) aplicada sobre a restrição cônica de (NSDP) para obtermos uma importante expressão para o cálculo da derivada de funções espectrais que dependem especificamente dos autovalores da função restrição matricial $G : \mathbb{R}^n \rightarrow \mathbb{S}^m$. Isto é, para funções do tipo $f(x) = g(\lambda(G(x)))$.

Proposição 3.1. Seja $G : \mathbb{R}^n \rightarrow \mathbb{S}^m$ tal que $x \mapsto G(x) \in \mathbb{S}^m$. Dada $g : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função simétrica qualquer. Defina a função espectral $f \doteq (g \circ \lambda) \circ G$ tal que $f(x) = g(\lambda(G(x)))$. Então,

$$\nabla f(x) = DG^*(x)[U^T \text{Diag}(g'(\lambda(G(x))))U] \quad (3.5)$$

para toda matriz ortogonal U tal que $G(x) = U^T (\text{Diag}(\lambda(G(x))))U$

Demonstração. Note que, como g é uma função simétrica pela proposição (1.11) f é diferenciável se, e somente se, $g(\lambda(G(x)))$ é diferenciável em $\lambda(G(x))$. Além de que, a derivada

de $g \circ \lambda$ aplicada em $G(x)$ é a aplicação $D(g \circ \lambda)(G(x)) : \mathbb{S}^m \rightarrow \mathbb{R}$ tal que

$$\begin{aligned} Z \mapsto D(g \circ \lambda)(G(x))[Z] &= \langle (g \circ \lambda)'(G(x)), Z \rangle \\ &= \langle U^T \text{Diag}(g'(\lambda(G(x))))U, Z \rangle \end{aligned}$$

Para concluirmos o resultado, basta observar, que para calcular a derivada da função f podemos aplicamos a regra da cadeia

$$D(H(G(x)))[\cdot] = DH(G(x))[DG(x)[\cdot]] \quad (3.6)$$

para a composição das aplicações $D(g \circ \lambda)(G(x)) : \mathbb{S}^m \rightarrow \mathbb{R}$ e $DG(x) : \mathbb{S}^m \rightarrow \mathbb{R}^n$, portanto $Df : \mathbb{R}^n \rightarrow \mathbb{R}$ é tal que

$$\begin{aligned} d \mapsto Df(x)[d] &= D(g \circ \lambda)(G(x))[DG(x)[d]] : \\ &= \langle U^T \text{Diag}(g'(\lambda(G(x))))U, DG(x)[d] \rangle \\ &= d^T DG^*(x)[U^T \text{Diag}(g'(\lambda(G(x))))U] \end{aligned}$$

assim, considerando que $\nabla f(x)$ é o vetor tal que $Df(x)[d] = d^T \nabla f(x)$ então

$$\nabla f(x) = DG^*(x)[U^T \text{Diag}(g'(\lambda(G(x))))U] \quad (3.7)$$

como queríamos mostrar. \square

A equação (3.5) será de extrema importância para o estudo dos multiplicadores de Lagrange para (NSDP) que será nossa próxima discussão.

Seja $\bar{x} \in \mathcal{F}$ com $r = \text{rank}(G(\bar{x}))$ e $\bar{U} \in \mathbb{R}^{m \times m}$ qualquer matriz ortogonal cujas colunas são compostas pelos autovetores de $G(\bar{x})$, tomando a partição $\bar{U} = [\bar{P} \bar{E}]$ tal que $\bar{P} \in \mathbb{R}^{m \times r}$ possui colunas que correspondem aos autovetores associados com os autovalores positivos de $G(\bar{x})$ e as colunas de $\bar{E} \in \mathbb{R}^{m \times m-r}$ correspondentes aos autovetores associados com os autovalores nulos de $G(\bar{x})$. Note que,

$$G(\bar{x})\bar{E} = 0 \quad \text{e} \quad \bar{E}^T \bar{E} = \mathbb{I}_{m-r}. \quad (3.8)$$

onde \bar{e}_i denota a i -ésima coluna de \bar{E} para $i \in \{1, \dots, m-r\}$. Assim, para uma matriz \bar{E} que satisfaz (3.8), faremos um abuso de notação e diremos que $\text{Ker}(G(x))$ é gerado por \bar{E} .

Um dos nossos objetivos é estudar condições de otimalidade baseadas nas clássicas equações de *Karush-Kuhn-Tucker* (KKT), que por meio das equações da definição (1.9) para (NCP), para o contexto de (NSDP) podem ser postas da seguinte maneira:

Definição 3.3. (KKT) Dizemos que um ponto $\bar{x} \in \mathcal{F}$ satisfaz as equações KKT quando existe algum $\bar{Y} \succeq 0$ tal que:

$$\begin{aligned} \nabla_x L(\bar{x}, \bar{Y}) &\doteq \nabla f(\bar{x}) - DG(\bar{x})^*[\bar{Y}] = 0, \\ \langle G(\bar{x}), \bar{Y} \rangle &= 0, \end{aligned} \quad (\text{KKT})$$

Onde $L : \mathbb{R}^n \times \mathbb{S}^m \rightarrow \mathbb{R}$ é a função Lagrangiana para (NSDP) dada por:

$$L(x, Y) \doteq f(x) - \langle G(x), Y \rangle \quad (3.9)$$

Definiremos $\Lambda(\bar{x}, f)$ como o conjunto de todos os multiplicadores de Lagrange associados a \bar{x} para o problema (NSDP) com f como função objetivo. Se $\Lambda(\bar{x}, f) \neq 0$ então seus elementos serão ditos pontos KKT para (NSDP). A segunda equação de (KKT) é dita condição de complementariedade, e esta pode ser caracterizada da seguinte forma:

Proposição 3.2. *Seja $\bar{x} \in \mathcal{F}$, seja $r = \text{rank}(G(\bar{x}))$ e $\bar{E} \in \mathbb{R}^{m \times m-r}$ tal que $\text{Ker}(G(\bar{x}))$ é gerado por \bar{E} . Para todo $\bar{Y} \in \Lambda(\bar{x}, f)$ temos que $\langle G(\bar{x}), \bar{Y} \rangle = 0$ se, e somente se, \bar{Y} é da forma*

$$\bar{Y} = \bar{E}\tilde{Y}\bar{E}^T \quad (3.10)$$

onde $\tilde{Y} \in \mathbb{S}_+^{m-r}$.

Demonstração. Como \bar{Y} e $G(\bar{x})$ são semi-definidas positivas, pela proposição (1.6) a complementariedade $\langle G(\bar{x}), \bar{Y} \rangle = 0$ é equivalente a $G(\bar{x})\bar{Y} = 0$. Isto é,

$$\text{Im}\bar{Y} \subseteq (\text{Im}G(\bar{x}))^\perp = \text{Ker}G(\bar{x}) \quad (3.11)$$

ou seja, $\bar{Y}G(\bar{x})v = 0$ para todo $v \in \text{Ker}G(\bar{x}) \supseteq \text{Im}\bar{Y}$, logo $\bar{Y}G(\bar{x}) = 0 = G(\bar{x})\bar{Y}$. Assim, $G(\bar{x})\bar{Y}$ comutam e pela proposição (1.4) existe $U = [\bar{P} \ \bar{E}]$ tal que $G(\bar{x}) = U\Lambda_\rho U^T$ e $\bar{Y} = U\Lambda_\gamma U^T$. Agora, seja \bar{E} associado ao $\text{Ker}G(\bar{x})$ têm-se

$$G(\bar{x})\bar{Y} = G(\bar{x})(\bar{P}\Lambda_\rho\bar{P}^T + \bar{E}\Lambda_\gamma\bar{E}^T) = G(\bar{x})(\bar{P}\Lambda_\rho\bar{P}^T) = 0 \quad (3.12)$$

isto é, $\Lambda_\rho\bar{P}^T = 0$, onde Λ_ρ e Λ_γ são sub-blocos da matriz diagonal Λ_γ , assim $\bar{Y} = \bar{E}\Lambda_\gamma\bar{E}^T$. Por outro lado, seja \tilde{E} tal que $\text{Ker}G(\bar{x})$ é gerado por \tilde{E} então, repetindo o argumento acima teremos, novamente

$$\bar{Y} = \tilde{E}\Lambda_{\tilde{E}}\tilde{E}^T = \sum_{i=1}^{m-r} \tilde{\theta}_i \tilde{e}_i \tilde{e}_i^T$$

desta forma, fixando os vetores \bar{e}_i e escrevendo $\tilde{e}_i = \sum_{j=1}^{m-r} \lambda_j \bar{e}_j$ para cada $i \in \{1, \dots, m-r\}$ temos,

$$\begin{aligned} \bar{Y} &= \sum_{i=1}^{m-r} \tilde{\theta}_i \left(\sum_{j=1}^{m-r} \lambda_j \bar{e}_j \right) \left(\sum_{j=1}^{m-r} \lambda_j \bar{e}_j \right)^T \\ &= \sum_{i=1}^{m-r} \tilde{\theta}_i \left(\sum_{i,j=1}^{m-r} \lambda_{ij} \bar{e}_i \bar{e}_j^T \right) \\ &= \sum_{i,j=1}^{m-r} \tilde{\lambda}_{ij} \bar{e}_i \bar{e}_j^T \\ &= \bar{E}\tilde{Y}\bar{E}^T \end{aligned} \quad (3.13)$$

com $[\tilde{Y}]_{ij} = \tilde{\lambda}_{ij} = \tilde{\theta}_i \lambda_{ij}$. Portanto, como $\bar{Y} \in \mathbb{S}_+^m$ então $\tilde{Y} \in \mathbb{S}_+^m$. Perceba que \tilde{Y} não precisa

ser necessariamente diagonal. \square

Observação 3.1. *É válido lembrar que as condições (KKT) não constituem por si só, uma condição necessária para a solução local do problema, a menos, que estas estejam postas em conjunto com alguma condição de qualificação, por exemplo, uma das mais conhecidas e estudadas condições de qualificação para (NSDP) é conhecida como Robinson's CQ. Veja na definição a seguir*

Definição 3.4 (SHAPIRO, 1997, Def. 3). *Seja $\bar{x} \in \mathcal{F}$, \bar{x} satisfaz Robinson's CQ se, existe uma direção $d \in \mathbb{R}^n$ tal que*

$$G(\bar{x}) + DG(\bar{x})[d] \in \text{int}(\mathbb{S}_+^m) \quad (3.14)$$

onde $\text{int}(\mathbb{S}_+^m)$ denota o interior topológico de \mathbb{S}_+^m .

Entretanto, podemos usar a proposição 2.97 de BONNANS e SHAPIRO, 2000 para caracterizar Robinson's CQ em sua forma dual, dizendo que, $\bar{x} \in \mathcal{F}$ satisfaz Robinson's CQ se, e somente se,

$$\left. \begin{array}{l} DG(\bar{x})^*[Y] = 0 \\ \langle G(\bar{x}), Y \rangle = 0 \\ Y \succeq 0 \end{array} \right\} \Rightarrow Y = 0 \quad (3.15)$$

utilizaremos, daqui em diante a caracterização acima para propormos novos resultados.

3.1.1 Elementos da Análise Convexa

Aqui escreveremos alguns elementos da análise convexa para o contexto de (NSDP).

Partindo da definição (1.4) para o cone $C = \mathbb{S}_+^m$ têm-se para uma matriz $M \in \mathbb{S}_+^m$ temos

$$T_{\mathbb{S}_+^m}(M) = \{N \in \mathbb{S}^m : d^T N d \geq 0, \forall d \in \text{Ker } M\} \quad (3.16)$$

Portanto, para cada ponto viável \bar{x} de (NSDP) e para \bar{E} tal que $\text{Ker } G(\bar{x})$ é gerado por \bar{E} têm-se

$$\begin{aligned} T_{\mathbb{S}_+^m}(M) &= \{N \in \mathbb{S}^m : (\bar{E}y)^T N (\bar{E}y) \geq 0, \forall y \in \mathbb{R}_*^n\} \\ &= \{N \in \mathbb{S}^m : y^T \bar{E}^T N \bar{E} y \geq 0, \forall y \in \mathbb{R}^n\} \\ &= \{N \in \mathbb{S}^m : \bar{E}^T N \bar{E} \succeq 0\} \end{aligned} \quad (3.17)$$

Assim, teremos que o espaço linearidade, ou *lineality space*, definido como o maior subespaço contido em $T_{\mathbb{S}_+^m}(M)$ poderá ser caracterizado por

$$\text{lin}(T_{\mathbb{S}_+^m}(M)) = \{N \in \mathbb{S}^m : \bar{E}^T N \bar{E} = 0\} \quad (3.18)$$

com os elementos acima, podemos discutir uma condição de regularidade, posta originalmente sobre o contexto de otimização para autovalores em SHAPIRO e FAN, 1995 sobre o nome de *Transversalidade*. Veja

Definição 3.5 (SHAPIRO, 1997, Def. 4). *Seja \bar{x} um ponto viável para o problema (NSDP)*

diz-se que \bar{x} satisfaz *Nondegeneracy-CQ* se existe $d \in \mathbb{R}^n$ tal que

$$\text{Im } DG(\bar{x}) + \text{lin}(T_{\mathbb{S}_+^m}(G(\bar{x}))) = \mathbb{S}_+^m$$

3.1.2 Redução para NSDP

Nossa estratégia para definir a *Nondegeneracy-CQ* para o contexto de (NSDP) será análoga ao que foi feito no capítulo (2), isto é, utilizaremos a ideia de redução de cone, definida em (2.1), de modo a parametrizar localmente o cone \mathbb{S}_+^m em torno de $\bar{y} = G(\bar{x})$ para o cone \mathbb{S}_+^{m-r} onde $r = \text{rank}(G(\bar{x}))$.

Exemplo 3.1. Seja $\bar{M} \in \mathbb{S}_+^m$ com $r = \text{rank}(\bar{M})$ de modo que $\bar{M} = [\bar{P} \ \bar{E}] \text{Diag}(\lambda(\bar{M})) [\bar{P} \ \bar{E}]^T$ onde $\lambda_1(\bar{M}) \geq \dots, \lambda_m(\bar{M})$ e \bar{E} a matriz associada aos $m - r$ autovalores nulos de \bar{M} , ou seja, suas colunas $\bar{e}_1(\bar{M}), \dots, \bar{e}_{m-r}(\bar{M})$ constituem uma base ortogonal e $\bar{M}\bar{E} = 0$. Como a função autovalor é uma função contínua, para qualquer matriz $N \in \mathbb{S}^m$ suficientemente próxima de \bar{M} vale que $r = \text{rank}(N)$ e

$$N = [P_N \ E_N] \text{Diag}(\lambda(N)) [P_N \ E_N]^T \quad (3.19)$$

de modo que E_N é uma matriz ortogonal, associada aos $m - r$ menores autovalores de N . Assim, percebe-se que, $N \in \mathbb{S}_+^m$, se e somente se, a matriz diagonal $E_N^T N E_N \in \mathbb{S}_+^{m-r}$, já que os outros r autovalores são não negativos. Logo, para uma vizinhança $V(\bar{M})$ suficientemente pequena de \bar{M} , o cone \mathbb{S}_+^m pode ser posto como

$$\{N \in \mathbb{S} : E_N^T N E_N \in \mathbb{S}_+^{m-r}\} \quad (3.20)$$

desta forma, considerando a aplicação $\psi : V(\bar{M}) \rightarrow \mathbb{S}_+^{m-r}$ tal que

$$N \mapsto \psi(N) = E_N^T N E_N \quad (3.21)$$

teremos que $\psi(\cdot)$ é uma aplicação continuamente diferenciável, cuja derivada em \bar{M} é aplicação $D\psi(\bar{M}) : \mathbb{S}^m \rightarrow \mathbb{S}_+^{m-r}$ dada por:

$$D\psi(\bar{M})[N] = (D\bar{E}^T [N])\bar{M}\bar{E} + \bar{E}^T N \bar{E} + \bar{E}^T M (D\bar{E}[N]) \quad (3.22)$$

portanto, como $\bar{M}\bar{E} = 0$, temos que $D\psi(\bar{M})[N] = \bar{E}^T N \bar{E}$. Segue então que $D\psi(\bar{M})[N]$ é sobrejetiva e \mathbb{S}_+^m é C^∞ -redutível em \bar{M} para $C = \mathbb{S}_+^{m-r}$.

Portanto, por meio da definição (2.2), temos que, para $\bar{x} \in \mathcal{F}$ este satisfaz *Nondegeneracy-CQ* se, e somente se, $\text{Ker}G(\bar{x}) = 0$ ou a aplicação linear $\psi_{\bar{x}} : \mathbb{R}^n \rightarrow \mathbb{S}_+^{m-r}$ tal que

$$d \mapsto \psi_{\bar{x}}[d] \doteq \bar{E}^T DG(\bar{x})[d] \bar{E} \quad (3.23)$$

é sobrejetiva, para uma matriz fixada \bar{E} tal que $\text{Ker}G(\bar{x})$ é gerado por \bar{E} . Assim, dizer que

a aplicação

$$\begin{aligned}\psi_{\bar{x}}[d] &\doteq \bar{E}^T DG(\bar{x})[d]\bar{E} = \bar{E}^T \left(\sum_{i=1}^n D_{x_i} d_i \right) \bar{E} \\ &= \sum_{i=1}^n (\bar{E}^T D_{x_i} G(\bar{x}) \bar{E}) d_i\end{aligned}\quad (3.24)$$

é sobrejetiva, é o mesmo que afirmar que:

Proposição 3.3 (SHAPIRO, 1997, Prop. 3.2). *Seja \bar{x} um ponto viável para (NSDP), denotando $r = \text{rank}(G(\bar{x}))$. Então, \bar{x} satisfaz "Nondegeneracy-CQ" se, e somente se, $\text{Ker } G(\bar{x}) = 0$ ou os vetores*

$$v_{ij}(\bar{x}, \bar{E}) \doteq \left[\bar{e}_i^T D_{x_1} G(\bar{x}) \bar{e}_j, \dots, \bar{e}_i^T D_{x_n} G(\bar{x}) \bar{e}_j \right]^T \quad 1 \leq i \leq j \leq m - r \quad (3.25)$$

são linearmente independentes, onde $\bar{E} \in \mathbb{R}^{m \times m-r}$ é uma matriz fixa arbitrária, tal que $\text{Ker}(G(\bar{x}))$ é gerado por \bar{E} .

Veja que,

$$\begin{aligned}v_{ij}(\bar{x}, \bar{E}) &\doteq \left[\bar{e}_i^T D_{x_1} G(\bar{x}) \bar{e}_j, \dots, \bar{e}_i^T D_{x_n} G(\bar{x}) \bar{e}_j \right]^T \\ &= \left[\langle D_{x_1} G(x), \bar{e}_i \bar{e}_j^T \rangle, \dots, \langle D_{x_n} G(x), \bar{e}_i \bar{e}_j^T \rangle \right]^T \\ &= \left[\left\langle D_{x_1} G(x), \frac{\bar{e}_i \bar{e}_j^T + \bar{e}_j \bar{e}_i^T}{2} \right\rangle, \dots, \left\langle D_{x_n} G(x), \frac{\bar{e}_i \bar{e}_j^T + \bar{e}_j \bar{e}_i^T}{2} \right\rangle \right]^T \\ &= DG(\bar{x})^* \left[\frac{\bar{e}_i \bar{e}_j^T + \bar{e}_j \bar{e}_i^T}{2} \right]\end{aligned}\quad (3.26)$$

Assim, *Nondegeneracy-CQ* é equivalente à independência linear dos vetores

$$DG(\bar{x})^* \left[\frac{\bar{e}_i \bar{e}_j^T + \bar{e}_j \bar{e}_i^T}{2} \right], \quad 1 \leq i \leq j \leq m - r \quad (3.27)$$

Da proposição acima, apresentaremos agora uma proposição equivalente, de modo que poderemos avaliar *Nondegeneracy-CQ* sobre a independência linear de um conjunto essencialmente menor, entretanto com o custo de aumentar as observações para todas as possíveis escolhas de \bar{E} ao invés de olharmos para apenas uma \bar{E} fixada.

Proposição 3.4 (ANDREANI, HAESER, MITO e RAMÍREZ, 2023, Prop. 3.2). *Seja \bar{x} um ponto viável para (NSDP) e $r = \text{rank}(G(\bar{x}))$. Considere*

$$\{v_{ii}(\bar{x}, \bar{E}) : i \in \{1, \dots, m - r\}\} \quad (3.28)$$

então,

1. \bar{x} satisfaz *Nondegeneracy-CQ* se, e somente se $r = m$, ou o conjunto (3.28) é linearmente independente para cada matriz \bar{E} tal que $\text{Ker}(G(\bar{x}))$ é gerado por \bar{E} .

2. \bar{x} satisfaz Robinson's CQ se, e somente se $r = m$ ou o conjunto (3.28) é positivo linearmente independente para cada matriz \bar{E} tal que $\text{Ker}(G(\bar{x}))$ é gerado por \bar{E} .

Demonstração. Para $r = m$ o resultado verifica-se trivialmente, o caso de maior interesse consiste quando $r < m$. Primeiramente, perceba que dada $\bar{E} \in \mathbb{R}^{m \times (m-r)}$ tal que $\text{Ker}(G(\bar{x}))$ é gerado por \bar{E} , perceba que (3.28) é (positivo) linearmente independente se, e somente se, dados os escalares $\alpha_1, \dots, \alpha_{m-r} \in \mathbb{R}$ (com $\alpha_1 \geq 0, \dots, \alpha_{m-r} \geq 0$) respectivamente tal que

$$\sum_{i=1}^{m-r} \alpha_i \nu_{ii}(\bar{x}, \bar{E}) = \sum_{i=1}^{m-r} \alpha_i DG(\bar{x})^* [\bar{e}_i \bar{e}_i^T] = DG(\bar{x})^* \left[\sum_{i=1}^{m-r} \alpha_i \bar{e}_i \bar{e}_i^T \right] = 0, \quad (3.29)$$

então, devemos ter $\alpha_1 = \dots = \alpha_{m-r} = 0$. Desta maneira, (3.28) é (positivo) linearmente independente se, e somente se, para cada matriz $Y \in \mathbb{S}^m$ do tipo

$$Y = \sum_{i=1}^{m-r} \alpha_i \bar{e}_i \bar{e}_i^T = \bar{E} \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_{m-r} \end{bmatrix} \bar{E}^T = \bar{E} \text{Diag}(\alpha) \bar{E}^T \quad (3.30)$$

com $\alpha = (\alpha_1, \dots, \alpha_{m-r}) \in \mathbb{R}^{m-r}$ ($\alpha = (\alpha_1, \dots, \alpha_{m-r}) \in \mathbb{R}_+^{m-r}$) devemos ter

$$DG^*[Y] = 0 \Rightarrow Y = 0 \quad (3.31)$$

Posto isto, primeiramente suponha \bar{x} satisfaz *Nondegeneracy-CQ* isto é, (3.31) vale para toda matriz $Y \in \mathbb{S}^m$ do tipo

$$Y = \sum_{i,j=1}^{m-r} \alpha_{ij} \frac{\bar{e}_i \bar{e}_j^T + \bar{e}_j \bar{e}_i^T}{2} = \bar{E} \begin{bmatrix} \alpha_{11} & \dots & \alpha_{1m-r} \\ \vdots & \ddots & \vdots \\ \alpha_{m-r1} & \dots & \alpha_{m-rm-r} \end{bmatrix} \bar{E}^T \quad (3.32)$$

com $\alpha_{ij} \in \mathbb{R}$ e $\alpha_{ij} = \alpha_{ji}$. Isto é, para $Y = \bar{E} Z \bar{E}^T$ com $Z \in \mathbb{S}^{m-r}$ e \bar{E} uma matriz fixada. Em particular, (3.31) vale para matrizes Y tais que $Y = \bar{E} \text{Diag}(\alpha) \bar{E}^T$ logo 3.28 é linearmente independente. Agora por outro lado, suponha que (3.28) é linearmente independente para toda matriz \bar{E} tal que $\text{ker}G(\bar{x})$ é gerado por \bar{E} , desta forma, considere $Y = \bar{E} Z \bar{E}^T$ com $Z \in \mathbb{S}^m$. Como Z é uma matriz simétrica, considere uma matriz ortogonal $C \in \mathbb{R}^{m-r \times m-r}$ tal que $C^T Z C = \text{Diag}(z_1, \dots, z_{m-r})$, logo

$$Y = \bar{E} C \text{Diag}(z_1, \dots, z_{m-r}) C^T \bar{E}^T = (\bar{E} C) \text{Diag}(z_1, \dots, z_{m-r}) (C \bar{E})^T \quad (3.33)$$

Agora, perceba que a matriz $\bar{E} C$ também é uma matriz que gera o espaço $\text{ker}G(\bar{x})$ logo, por hipótese Y satisfaz (3.31) e (3.30), ou seja, (3.28) é linearmente independente. Por fim, para provarmos a equivalência referente a *Robinson's CQ* note que, de (3.15) \bar{x} satisfaz *Robinson's CQ* se, e somente se, (3.31) vale para toda matriz $Y \in \mathbb{S}_+^m$ tal que $\langle G(\bar{x}), Y \rangle = 0$ ou melhor, de (3.2), para toda matriz $Y \in \mathbb{S}_+^m$ da forma $Y = \bar{E} \tilde{Y} \bar{E}^T$ com $\tilde{Y} \in \mathbb{S}_+^{m-r \times m-r}$, ou seja, (3.31) vale para matrizes $Y = \bar{E} \text{Diag}(\alpha) \bar{E}^T$ com $\alpha \geq 0$, logo (3.28) são positivo-linearmente independentes.

Para a recíproca, suponha que (3.28) são positivo-linearmente independentes e toma-

mos Y tal que $DG^*(\bar{x})[Y] = 0$, $Y \in \mathbb{S}_+^m$ e $\langle G(\bar{x}), Y \rangle = 0$, repetindo o argumento feito para *Nondegeneracy-CQ* pode-se obter uma matriz \bar{E} que gera o espaço $\text{Ker}G(\bar{x})$ e que seja da forma $Y = \bar{E}\text{Diag}(\alpha)\bar{E}^T$ com $\alpha \geq 0$. Logo, por hipótese $Y = 0$ e como Y é arbitrária, \bar{x} satisfaz *Robinson's CQ*. \square

Motivados pela proposição (3.1) encontraremos uma importante relação entre a soma dos menores autovalores de $G(\bar{x})$ com uma matriz $\bar{E} \in \mathbb{R}^{m \times m-r}$ tal que $\text{Ker}(G(\bar{x}))$ é gerado por \bar{E} . Veja

Proposição 3.5. *Seja $G : \mathbb{R}^n \rightarrow \mathbb{S}^m$ tal que $x \mapsto G(x) \in \mathbb{S}^m$. Considere $r = \text{rank}(G(x))$ e $E \in \mathbb{R}^{m \times m-r}$ uma matriz tal que $\text{Ker}(G(x))$ é gerado por E . Seja $\sigma_{m-r} : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que*

$$\sigma_{m-r}(x) \doteq \sum_{k=1}^{m-r} \lambda_k(G(x)) \quad (3.34)$$

a soma dos "m - r" menores autovalores de $G(x)$. Então

$$\nabla \sigma_{m-r}(x) = DG^*(x)[EE^T] \quad (3.35)$$

Demonstração. Note que $\sigma_{m-r} \doteq (g \circ \lambda) \circ G$ é uma função espectral onde $g : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função simétrica dada pela soma das "m - r" menores componentes. Assim, podemos utilizar a equação (3.5) para o cálculo de $\nabla \sigma_{m-r}(x)$. Assim para toda U ortogonal tal que $G(x) = U\text{Diag}(\lambda(G(x)))U^T$ com $\lambda_1(G(x)) \geq \dots, \lambda_m(G(x))$ teremos:

$$\nabla \sigma_{m-r}(x) = DG^*(x)[U^T \text{Diag}(g'(\lambda(G(x))))U] \quad (3.36)$$

Assim, perceba que g é uma função diferenciável em \mathbb{R}^n e $g'(\lambda(G(x)))$ é o vetor, tal que, as primeiras r componentes são nulas e as últimas $m-r$ são iguais a 1. Deste modo, tomando a partição $U = [P \ E]$ tal que $\text{ker}(G(x))$ é gerado por E , então a matriz $U^T \text{Diag}(g'(\lambda(G(x))))U$ será dada pela matriz bloco

$$\begin{aligned} [P \ E]\text{Diag}(g'(\lambda(G(x))))[P \ E]^T &= [P \ E] \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{I}_{m-r} \end{bmatrix} [P \ E]^T \\ &= EE^T \end{aligned} \quad (3.37)$$

portanto

$$\nabla \sigma_{m-r}(x) = DG^*(x)[EE^T] \quad \square$$

Para um problema de otimização não linear (PNL), em [WACHSMUTH, 2013](#), Teo. 2 prova-se que, se o multiplicador de Lagrange associado a um dado ponto viável é único, então têm-se que vale (LICQ), ou seja, os gradientes das restrições do problema são linearmente independentes. Por meio das duas proposições acima, abordaremos aqui uma nova caracterização para a prova deste mesmo resultado para o contexto de (NSDP) por meio da *Nondegeneracy-CQ*, veja:

Proposição 3.6. *Seja \bar{x} um ponto viável para (NSDP), se existe um único multiplicador de Lagrange $\bar{Y} \in \Lambda(\bar{x}, f)$ para toda função f que assume mínimo em \bar{x} então, \bar{x} satisfaz Nondegeneracy-CQ.*

Demonstração. Primeiramente, é importante observar que o único multiplicador de Lagrange varia de acordo com a função objetivo f . Com isto, considere $r = \text{rank}(G(\bar{x}))$ e uma matriz arbitrária $\bar{E} \in \mathbb{R}^{m \times m-r}$ tal que $\text{Ker}G(\bar{x})$ é gerado por \bar{E} . Defina $f : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que $f(x) = -\sigma_{m-r}(x)$ como em (3.34). Assim, têm-se que $f(x) \geq 0$ para todo $x \in \mathbb{R}^n$ além que, pelo item 2 da proposição (1.2) $f(\bar{x}) = 0$. Portanto, $f(\bar{x}) \leq f(x) \forall x \in \mathbb{R}^n$, isto é, \bar{x} é um minimizador local para o problema de função objetivo f . Assim, $\bar{Y} \in \Lambda(\bar{x}, f)$ se, e somente se, $\bar{Y} \in \mathbb{S}_+^m$ e

$$\begin{aligned} \nabla f(\bar{x}) + DG^*(\bar{x})[\bar{Y}] &= 0 \\ \langle G(\bar{x}), \bar{Y} \rangle &= 0 \end{aligned} \quad (3.38)$$

Da proposição (3.5) e utilizando o fato de que $\langle G(\bar{x}), \bar{Y} \rangle = 0$ se, e somente se, $\bar{Y} \in \mathbb{S}_+^m$ com \bar{Y} da forma $\bar{Y} = \bar{E}\tilde{Y}\bar{E}^T$ (3.2) para alguma $\tilde{Y} \in \mathbb{S}_+^{m-r \times m-r}$, as equações (3.38) são dadas por

$$\begin{aligned} -DG^*(\bar{x})[\bar{E}\bar{E}^T] + DG^*(\bar{x})[\bar{E}\tilde{Y}\bar{E}^T] &= 0 \\ DG^*(\bar{x})[\bar{E}\tilde{Y}\bar{E}^T - \bar{E}\bar{E}^T] &= 0 \end{aligned} \quad (3.39)$$

Assim, perceba que $\bar{Y} = \bar{E}\bar{E}^T \in \Lambda(\bar{x}, f)$. Por outro lado, seja $Y \in \mathbb{S}^m$ da forma $Y = \bar{E}\text{Diag}(\alpha)\bar{E}^T$ tal que $DG^*(\bar{x})[Y] = 0$ para qualquer $\alpha = (\alpha_1, \dots, \alpha_{m-r}) \in \mathbb{R}^n$. Denote $C = 1 + \max\{|\alpha_i|\}$ com $i \in \{1, \dots, m-r\}$ desta forma, note que

$$\tilde{Y} = \bar{Y} + \bar{E}\text{Diag}(\alpha/C)\bar{E}^T$$

é um multiplicador de Lagrange logo, supondo a unicidade para $\Lambda(\bar{x}, f)$ têm-se

$$\bar{Y} = \bar{Y} + \bar{E}\text{Diag}(\alpha/C)\bar{E}^T$$

ou seja

$$\bar{E}\mathbb{I}\bar{E}^T = \bar{E}\bar{E}^T + \bar{E}\text{Diag}(\alpha/C)\bar{E}^T = \bar{E}(\mathbb{I} + \text{Diag}(\alpha/C))\bar{E}^T$$

assim, $1 = 1 + \frac{\alpha_i}{C}$ para todo $i \in \{1, \dots, m-r\}$ ou seja $\alpha_i = 0$, portanto $v_{ii}(\bar{x}, \bar{E})$ são linearmente independentes, isto é, \bar{x} satisfaz Nondegeneracy-CQ. \square

Até então, veja que a caracterização da Nondegeneracy-CQ feita em (3.4) possui relação direta com a análise dos multiplicadores de Lagrange, tornando a proposição (3.4) uma importante ferramenta teórica para o estudo da Nondegeneracy-CQ. Entretanto, assim como foi estabelecido no capítulo (2), procuraremos estudar, se a análise sobre todas as possíveis escolhas de \bar{E} poderá ser, ainda mais refinada, a ponto de termos uma condição de qualificação essencialmente mais fraca, que aproveite melhor a estrutura do problema e, que nos forneça um bom valor teórico para o estudo dos pontos KKT. Com isto, a próxima seção nos trará uma motivação para definirmos o conceito de *Weak-Nondegeneracy* para (NSDP).

3.2 Weak-Nondegeneracy para NSDP

3.2.1 Sequência de Autovetores e Nondegeneracy

Na literatura, há diversas técnicas e conceitos para determinar a existência de multiplicadores de Lagrange para (NSDP). Em [ANDREANI, HAESER e VIANA, 2020](#) utiliza-se uma interessante técnica, que consiste essencialmente, em aplicar um método de penalidade externa para (NSDP) regularizando-o localmente em torno de um minimizador local através de uma função penalidade $P(x)$. Esta função penalidade, essencialmente, determina uma medida para a viabilidade de um ponto qualquer. O próximo resultado, trata-se de uma fórmula bem conhecida para o estudo desta função penalidade, na qual terá grande relevância para a determinação da existência de pontos KKT.

Proposição 3.7. *Seja $G : \mathbb{R}^n \rightarrow \mathbb{S}^m$ uma função diferenciável e $P(x) = \|\Pi_{\mathbb{S}_+^m} G(x)\|^2$. Então o gradiente de P em $x \in \mathbb{R}^n$ é dado por*

$$\nabla P(x) = 2DG(x)^* \left[\Pi_{\mathbb{S}_+^m} G(x) \right] \quad (3.40)$$

Demonstração. Primeiramente perceba que $P(x) = \|\Pi_{\mathbb{S}_+^m} G(x)\|^2 = \text{tr} \left((\Pi_{\mathbb{S}_+^m} G(x))^2 \right)$ assim, utilizando o fato de que o traço de uma matriz é caracterizado pela soma dos seus autovalores têm-se,

$$\begin{aligned} P(x) &= \sum_{i=1}^m \lambda_i \left((\Pi_{\mathbb{S}_+^m} G(x))^2 \right) \\ &= \sum_{i=1}^m \lambda_i \left(\Pi_{\mathbb{S}_+^m} G(x) \right)^2 \end{aligned} \quad (3.41)$$

ou seja, de (1.10)

$$P(x) = \sum_{i=1}^m [\lambda_i(G(x))]_{+}^2 \quad (3.42)$$

Assim perceba que P é uma função espectral da forma $P \doteq (g \circ \lambda) \circ G$ com a função simétrica $g : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que

$$x \mapsto g(x) = \sum_{i=1}^n [x_i]_{+}^2, \quad (3.43)$$

onde $[x_i]_{+} = \max\{0, x_i\}$. Logo, $\nabla g(x) = 2([x_1]_{+}, \dots, [x_n]_{+}) = 2[x]_{+}$ e utilizando a fórmula (3.5) temos

$$\nabla P(x) = DG^*(x) [U^T \text{Diag}(g'(\lambda(G(x)))) U] \quad (3.44)$$

onde $g'(\lambda(G(x))) = 2[\lambda(G(x))]_{+}$ daí

$$\begin{aligned} U \text{Diag}(g'(\lambda(G(x)))) U &= U \text{Diag}(2[\lambda(G(x))]_{+}) U^T \\ &= U \text{Diag} \left(2\lambda \left(\Pi_{\mathbb{S}_+^m} G(x) \right) \right) U^T \\ &= 2 \Pi_{\mathbb{S}_+^m} G(x) \end{aligned} \quad (3.45)$$

Portanto,

$$\nabla P(X) = 2DG^*(x) \left[\Pi_{\mathbb{S}_+^m} G(x) \right] \quad (3.46)$$

□

Através da função penalidade apresentada na proposição (3.7) os autores em [ANDREANI, HAESER e VIANA, 2020](#) constroem o conceito de pontos aproximadamente KKT (AKKT) para (NSDP), cujos são pontos, que determinam uma sequência de modo a convergir para uma solução do problema. Esta técnica se baseia nas tão conhecidas "Condições de otimalidade sequenciais" exploradas para problemas não lineares de otimização (PNL) como é visto em [ANDREANI, HAESER e MARTÍNEZ, 2011](#). Assim, exploraremos um importante resultado para nos motivar a nossa próxima definição.

Proposição 3.8 ([ANDREANI, GÓMEZ et al., 2022](#), Teo. 3.2). *Seja \bar{x} um minimizador local de (NSDP). Então, para qualquer sequência $\{\rho_k\}_{k \in \mathbb{N}} \rightarrow +\infty$, existe alguma sequência $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$ tal que para cada $k \in \mathbb{N}$, x^k é um minimizador local para a função penalizada regularizada*

$$F_k(x) = f(x) + \frac{1}{2} \|x - \bar{x}\|^2 + \frac{\rho_k}{2} \left\| \Pi_{\mathbb{S}_+^m}(-G(x)) \right\|^2 \quad (3.47)$$

Em particular, tomando a sequência $Y^k \doteq \rho_k \Pi_{\mathbb{S}_+^m}(-G(x^k))$ têm-se que $\nabla_x L(x^k, Y^k) \rightarrow 0$.

Demonstração. Primeiramente, considere uma vizinhança local para \bar{x} garantindo sua otimalidade, assim definiremos o seguinte subproblema regularizado

$$\begin{aligned} & \underset{x \in \mathbb{R}^n}{\text{Minimize}} && F_k(x), \\ & \text{sujeito a} && \|x - \bar{x}\| \leq \delta \end{aligned} \quad (3.48)$$

para $\delta > 0$. Agora, como o problema é limitado, garantimos $\{x^k\}_{k \in \mathbb{N}}$ uma sequência gerada pelos minimizadores locais de (3.48) e um ponto limite qualquer \bar{s} , isto é, $\lim_{k \in I} x^k = \bar{s}$ para

$I \subseteq_{\infty} \mathbb{N}$. Como x^k é um minimizador local, temos $F_k(x^k) \leq F_k(\bar{x})$ para todo $k \in \mathbb{N}$, assim

$$f(\bar{x}) + \frac{\rho_k}{2} \left\| \Pi_{\mathbb{S}_+^m}(-G(\bar{x})) \right\|^2 \geq f(x^k) + \frac{1}{2} \|x^k - \bar{x}\|^2 + \frac{\rho_k}{2} \left\| \Pi_{\mathbb{S}_+^m}(-G(x^k)) \right\|^2$$

o que nos leva a

$$\left\| \Pi_{\mathbb{S}_+^m}(-G(\bar{x})) \right\|^2 \geq \frac{2(f(x^k) - f(\bar{x}))}{\rho_k} + \frac{\|x^k - \bar{x}\|^2}{\rho_k} + \left\| \Pi_{\mathbb{S}_+^m}(-G(x^k)) \right\|^2$$

assim, como \bar{x} é um ponto viável para (NSDP), teremos que $\left\| \Pi_{\mathbb{S}_+^m}(-G(\bar{x})) \right\|^2 = 0$, além de que, $\{x^k\}_{k \in \mathbb{N}}$ é limitada e $\rho_k \rightarrow \infty$, portanto, teremos

$$\lim_{k \in I} \left\| \Pi_{\mathbb{S}_+^m}(-G(x^k)) \right\| = 0$$

ou seja, $G(\bar{s}) \in \mathbb{S}_+^m$. Por fim, perceba que, para qualquer $z \in \mathcal{F}$ tal que $\|z - \bar{x}\| \leq \delta$ temos

$$f(x^k) + \frac{1}{2}\|x^k - \bar{x}\|^2 \leq F_k(x^k) \leq F_k(z) = f(z) + \frac{1}{2}\|z - \bar{x}\|^2$$

com isto, tomando o limite para $k \in I$

$$f(\bar{s}) + \frac{1}{2}\|\bar{s} - \bar{x}\|^2 \leq f(z) + \frac{1}{2}\|z - \bar{x}\|^2$$

ou seja, \bar{s} é um minimizador para o problema local

$$\begin{aligned} & \underset{x \in \mathbb{R}^n}{\text{Minimize}} && f(x) + \frac{1}{2}\|x - \bar{x}\|^2, \\ & \text{sujeito a} && G(x) \in \mathbb{S}_+^m \\ & && \|x - \bar{x}\| \leq \delta \end{aligned} \quad (3.49)$$

portanto, $\bar{s} = \bar{x}$ e para $k \in I$ suficientemente grande, teremos $\|x^k - \bar{x}\| \leq \delta$ e para cada $k \in \mathbb{N}$ vale que, x^k é um minimizador local para o problema irrestrito para a função objetivo $F_k(x)$. \square

Com a proposição acima, prova-se que, sobre *Nondegeneracy-CQ* a sequência de multiplicadores de Lagrange para x^k é limitada, veja

Proposição 3.9 (ANDREANI, HAESER, MITO e RAMÍREZ, 2023, Sec. 3.1). *Seja \bar{x} um minimizador local de (NSDP). Sobre Nondegeneracy-CQ vale que a sequência $Y^k \doteq \rho_k \Pi_{\mathbb{S}_+^m}(-G(x^k))$ é limitada.*

Demonstração. Suponha que $\{Y^k\}_{k \in \mathbb{N}}$ é ilimitada, desta forma definimos a sequência limitada $\{Y^k / \|Y^k\|\}$ de modo que existe $I \subseteq_{\infty} \mathbb{N}$ tal que $\{Y^k\}_{k \in I} \rightarrow \bar{Y}$. Assim, de (3.40)

$$\nabla_x L(x^k, Y^k) = \nabla f(x^k) + (x^k - \bar{x}) + DG(x^k)^* [Y^k] \quad (3.50)$$

utilizando a proposição (3.8), isto é, de $\nabla_x L(x^k, Y^k) \rightarrow 0$ temos

$$DG(x^k)^* [Y^k / \|Y^k\|] \rightarrow 0$$

ou seja, $DG(\bar{x})^* [\bar{Y}] = 0$, logo $\bar{Y} \in \text{Ker } DG(\bar{x})^* = \text{Im } DG(\bar{x})^\perp$. Da definição de Y^k , note que, $0 \neq \bar{Y} \succeq 0$ e $\langle G(\bar{x}), \bar{Y} \rangle = 0$, então $\bar{Y} \in \text{lin}(T_{\mathbb{S}_+^m}(G(\bar{x})))^\perp$. Portanto,

$$\bar{Y} \in \text{Im } DG(\bar{x})^\perp \cap \text{lin}(T_{\mathbb{S}_+^m}(G(\bar{x})))^\perp \quad (3.51)$$

que contradiz a *Nondegeneracy-CQ*. \square

Observação 3.2. *Vamos observar que a proposição acima pode ser posta em termos da proposição (3.3), pois, tomando a decomposição espectral de Y^k , perceba que seus $m - r$ menores autovalores são nulos, isto é, $\lambda_i(Y^k) = \left[\rho_k \lambda_i(-G(x^k)) \right]_+ = 0$ para todo $i \in \{m - r + 1, \dots, m\}$ e todo $k \in \mathbb{N}$ suficientemente grande, visto que*

$$\lambda_i(-G(x^k)) = -\lambda_{m-i+1}(G(x^k)). \quad (3.52)$$

assim, de (3.8) e da equação (3.50) teremos

$$\nabla_x L(x^k, Y^k) = \nabla f(x^k) - \sum_{i=1}^{m-r} [\rho_k \lambda_i(-G(x^k))]_{+} v_{ii}(x^k, E^k) \rightarrow 0 \quad (3.53)$$

onde $E^k \in \mathbb{R}^{m \times m-r}$ é uma matriz cuja a i -ésima coluna é dada por $u_{m-i+1}(G(x^k))$. Com isto, percebe-se que, se pudermos escolher E^k de tal maneira, que seus pontos limites \bar{E} garantem que $\{v_{ii}(\bar{x}, \bar{E}) : i \in \{1, \dots, m-r\}\}$ sejam linearmente independentes, então, à luz do que foi visto na proposição (3.9), teremos que Y^k será limitada. Assim, analogamente ao capítulo (2), analisaremos o valor teórico ao se exigir a independência linear de $\{v_{ii}(\bar{x}, \bar{E}) : i \in \{1, \dots, m-r\}\}$ para os pontos limites de alguma sequência $\{E^k\}_{k \in \mathbb{N}}$ de autovetores de $G(x^k)$.

Definição 3.6 (Weak-Nondegeneracy). *Seja \bar{x} um ponto viável para (NSDP) e seja $r = \text{rank}(G(\bar{x}))$. Dizemos que \bar{x} satisfaz Weak-Nondegeneracy(respectivamente, Weak-Robinson's CQ) se, $\text{Ker } G(\bar{x}) = 0$ ou: Para toda sequência $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$ existe alguma sequência de matrizes com colunas ortogonais $\{E^k\}_{k \in \mathbb{N}} \subseteq \mathbb{R}^{m \times m-r}$ tal que*

1. As colunas de E^k são os autovetores associados aos $m-r$ menores autovalores de $G(x^k)$ para cada $k \in \mathbb{N}$.
2. Existe um ponto limite de $\{E^k\}_{k \in \mathbb{N}}$ tal que o conjunto $\{v_{ii}(\bar{x}, \bar{E}) : i \in \{1, \dots, m-r\}\}$ definido em (3.28) são (positivo) linearmente independentes.

Observação 3.3 (ANDREANI, HAESER, MITO e RAMÍREZ, 2023, Obs. 3.1). *Intuitivamente a definição acima, pode ser postulada por uma noção de continuidade de autovetores de $G(x)$. Note que, para \bar{x} viável com $r = \text{rank}(G(\bar{x}))$ como $r < m$ para cada x suficientemente próximo de \bar{x} têm-se $\lambda_r(G(x)) > \lambda_{r+1}(G(x))$, com isto podemos definir o conjunto*

$$\mathcal{B}(x) \doteq \left\{ E \in \mathbb{R}^{m \times m-r} : \begin{array}{l} G(x)e_i = \lambda_{m-i+1}(G(x))e_i, \quad \forall i \in \{1, \dots, m-r\} \\ E^T E = \mathbb{I}_{m-r} \end{array} \right\} \quad (3.54)$$

que estabelece todas as matrizes cujas colunas são autovetores ortogonais associados aos menores autovalores de $G(x)$, isto é, $E \doteq [e_1, \dots, e_{m-r}]$. A definição de Weak-Nondegeneracy busca pontos limites de sequências de matrizes ortogonais que estão em \mathcal{B} construindo assim $\mathcal{B}(x^k)$ para qualquer sequência $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$. De [BONNANS e SHAPIRO, 2000, Def. 2.52] podemos estabelecer o limite superior da sequência de imagens $\{\mathcal{B}(x^k)\}_{k \in \mathbb{N}}$ definido como:

$$\limsup_{k \rightarrow \infty} \mathcal{B}(x^k) \doteq \{z : \exists I \subseteq_{\infty} \mathbb{N}, \exists \{z^k\}_{k \in I} \rightarrow z, \forall k \in I, z \in \mathcal{B}(x^k)\}$$

Como os autores ressaltam em ANDREANI, HAESER, MITO e RAMÍREZ, 2023, determinar o conjunto $\mathcal{B}(x)$ na prática pode ser uma tarefa complicada. Logo, Weak-Nondegeneracy pode não ser uma boa ferramenta computacional para a maioria dos casos, entretanto pode ser um bom ponto de partida teórico para estabelecer uma nova condição de qualificação estruturada em Nondegeneracy-CQ, para isto vejamos qual é a associação entre a definição de (3.6) para a análise das condições KKT.

Proposição 3.10 (ANDREANI, HAESER, MITO e RAMÍREZ, 2023, Teo. 3.2). *Seja $\bar{x} \in \mathcal{F}$ um minimizador local para (NSDP) que satisfaz weak-Robinson's CQ então, \bar{x} satisfaz as condições KKT.*

Demonstração. Seja \bar{x} um minimizador local que satisfaz Robinson's-CQ. Considere a sequência $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$ tal que, x^k é minimizador local da função

$$f(x) + \frac{1}{2} \|x - \bar{x}\|^2 + \frac{\rho_k}{2} \left\| \prod_{\mathbb{S}_+^m} (-G(x)) \right\|^2. \quad (3.55)$$

Seja $\{Y^k\}_{k \in \mathbb{N}}$ tal que $Y^k \doteq \rho_k \prod_{\mathbb{S}_+^m} (-G(x^k))$ para uma sequência arbitrária $\{\rho_k\}_{k \in \mathbb{N}}$. Se $r = \text{rank}G(\bar{x})$ e $r < m$, teremos, como x^k é solução local vale que

$$\nabla f(x^k) + (x^k - \bar{x}) + DG(x^k)^* [Y^k] = 0. \quad (3.56)$$

Sem perda de generalidade podemos assumir $\lambda_1(-G(x^k)) \geq \dots \geq \lambda_m(-G(x^k))$ para cada k e, como $G(\cdot)$ e $\lambda(\cdot)$ são funções contínuas temos que $G(x^k) \rightarrow G(\bar{x})$. Assim, tomando k suficientemente grande, a saber $K_0 \in \mathbb{N}$, temos que $\lambda_i(-G(x^k)) < 0$ para todo $i \in \{m-r+1, \dots, m\}$. Logo para cada $k > K_0$ temos

$$Y^k = \sum_{i=1}^{m-r} \alpha_i^k e_i^k (e_i^k)^T, \quad (3.57)$$

com $\alpha_i^k = [\rho_k \lambda_i(-G(x^k))]_+ \geq 0$ e $e_i^k = u_{m-i+1}(G(x^k))$ uma sequência arbitrária de autovetores associados com $\lambda_{m-i+1}(G(x^k))$, para cada $i \in \{1, \dots, m-r\}$. Por fim, tome $E^k \doteq [e_1^k, \dots, e_{m-r}^k]$ que é limitada, assim, tomando uma subsequência se necessário, seja qualquer ponto limite $\bar{E} \doteq [\bar{e}_1, \dots, \bar{e}_{m-r}]$ tal que, $\text{Ker} G(\bar{x})$ é gerado por \bar{E} , assim de (3.56)

$$\begin{aligned} \nabla f(x^k) - \sum_{i=1}^{m-r} DG(x^k)^* [\alpha_i^k e_i^k (e_i^k)^T] &= 0 \\ \nabla f(x^k) - \sum_{i=1}^{m-r} \alpha_i^k DG(x^k)^* [e_i^k (e_i^k)^T] &\rightarrow 0 \end{aligned} \quad (3.58)$$

onde $Dg(x^k)^* [e_i^k (e_i^k)^T] = v_{ii}(x^k, E^k)$ assim teremos

$$\nabla f(x^k) - \alpha_i^k v_{ii}(x^k, E^k) \rightarrow 0 \quad (3.59)$$

Para a sequência $\{\alpha_i^k, \dots, \alpha_{m-r}^k\}$, se esta tiver qualquer subsequência convergente, tome qualquer ponto limite $\bar{\alpha} \doteq (\bar{\alpha}_i, \dots, \bar{\alpha}_{m-r})$ e perceba que $\bar{\alpha}$ gera um multiplicador de Lagrange para \bar{x} , cujo

$$\bar{Y} = \sum_{i=1}^{m-r} \bar{\alpha}_i \bar{e}(\bar{e})^T = \bar{E} \bar{Y} \bar{E}^T \quad (3.60)$$

Assim, nos resta mostrar, que a sequência $\{\alpha_i^k\}_{k \in \mathbb{N}}$ com $i \in \{1, \dots, m-r\}$, deve ser limitada se \bar{x} satisfaz Weak-Robinson's CQ. Suponhamos que $\{x^k\}_{k \in \mathbb{N}}$ seja ilimitada, desta forma

considere

$$m^k \doteq \max\{\alpha_i^k : i \in \{1, \dots, m-r\}\} \rightarrow \infty \quad (3.61)$$

assim, construa a sequência limitada $\{(\alpha_1^k, \dots, \alpha_{m-r}^k)/m^k\}_{k \in \mathbb{N}}$ e tomando uma subsequência se necessário, considere um ponto limite não nulo $(\alpha_1, \dots, \alpha_{m-r}) \geq 0$. Portanto, dividindo (3.59) por m^k para cada $k \in \mathbb{N}$, tomando $k \rightarrow +\infty$ obtemos

$$\sum_{i=1}^{m-r} \tilde{\alpha}_i v_{ii}(\bar{x}, \bar{E}) = 0 \quad (3.62)$$

logo, $\{v_{ii}(\bar{x}, \bar{E}) : i \in \{1, \dots, m-r\}\}$ é positivo linearmente dependente. Portanto, como assumimos qualquer escolha de $\{E^k\}$ e qualquer \bar{E} , \bar{x} não satisfaz *Weak-Robinson's CQ*. \square

3.2.2 Exemplos

Nesta seção, será discutido com um pouco mais de detalhes vários exemplos abordados em [ANDREANI, HAESER, MITO e RAMÍREZ, 2023](#), cujos se demonstram essenciais para um estudo detalhado sobre como obter uma equivalência entre *Weak-Nondegeneracy* e *Nondegeneracy-CQ*.

Exemplo 3.2 ([ANDREANI, GÓMEZ et al., 2022](#), Sec. 3.1). *Considere o problema (NSDP) com a seguinte restrição*

$$G(x) = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_1 \end{bmatrix}$$

sobre o ponto $\bar{x} = (0, 0)$. Assim, temos

- Claramente \bar{x} não satisfaz *Nondegeneracy* já que $G(\bar{x}) = 0$ (basta considerar a base canônica de \mathbb{R}^2 em (3.28));
- Por outro lado, note que os autovalores de $G(x)$ são dados por $\lambda_1 = x_1 + x_2$ e $\lambda_2 = x_1 - x_2$ para todo $x \in \mathbb{R}^2$ e seus respectivos autovetores $v \in \mathbb{R}^2$ são dados tais que $v_1 = v_2$ e $v_1 = -v_2$ se $x_2 \neq 0$, ou qualquer $v \in \mathbb{R}^2$ se $x_2 = 0$. Logo de (3.54)

$$\mathcal{B}(x) = \begin{cases} \left\{ \frac{\pm 1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}, \frac{\pm 1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \right\}, & \text{se } x_2 \neq 0 \\ \{E \in \mathbb{R}^{2 \times 2} : E^T E = \mathbb{I}_2\}, & \text{se } x_2 = 0 \end{cases}$$

deste modo, para toda sequência $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$ têm-se

$$E^k \doteq \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \quad e \quad \bar{E} \doteq \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \in \limsup_{k \rightarrow \infty} \mathcal{B}(x^k)$$

logo, localmente $G(\bar{x})$ gera os vetores $v_{11}(\bar{x}, \bar{E}) = [1, -1]$ e $v_{22}(\bar{x}, \bar{E}) = [1, 1]$ que são linearmente independentes. Logo, \bar{x} satisfaz *Weak-Nondegeneracy*.

- Considere o exemplo acima para a função objetivo $f(x) = 2x_1$, veja que $\bar{x} = (0, 0)$ é

um minimizador local para o problema, já que, $G(x) \in \mathbb{S}_+^m \Leftrightarrow x_1 \geq 0$, com isto, da definição (3.3), os multiplicadores de Lagrange para \bar{x} são dados por:

$$\begin{bmatrix} 2 \\ 0 \end{bmatrix} - \begin{bmatrix} \bar{y}_{11} + \bar{y}_{22} \\ 2\bar{y}_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

isto é, cada multiplicador de Lagrange $\bar{Y} \in \mathbb{S}_+^m$ associado a \bar{x} é da forma:

$$\bar{Y} = \begin{bmatrix} 1 - \theta & 0 \\ 0 & 1 + \theta \end{bmatrix} \quad \text{para} \quad \theta \in [-1, 1] \setminus \{0\}$$

O exemplo (3.2) exhibe duas importantes observações:

- *Nondegeneracy* e *Weak-Nondegeneracy* não são equivalentes;
- *Weak-Nondegeneracy* não garante a unicidade dos multiplicadores de Lagrange.

Para melhorarmos a análise, vejamos agora um caso com dimensão maior, mas que possui uma estrutura bastante específica.

Exemplo 3.3 (ANDREANI, GÓMEZ *et al.*, 2022, Sec. 3.1). Consideremos agora o problema (NSDP) com a seguinte restrição

$$G(x) \doteq \begin{bmatrix} x_{11} & 0 & x_{13} \\ 0 & x_{22} & 0 \\ x_{13} & 0 & x_{33} \end{bmatrix} \succeq 0$$

Note que o ponto viável $\bar{x} = (0, 0, 0, 0)$ de maneira equivalente ao exemplo (3.2) não satisfaz *Nondegeneracy-CQ*, para ver isso basta tomar o conjunto $\{v_{ii}(\bar{x}, \bar{E}) : i \in \{1, 2, 3\}\}$ em (3.28) com

$$\bar{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}.$$

Entretanto, \bar{x} satisfaz *Weak-Nondegeneracy*. Considere uma sequência $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$. Vejamos, se $x_{13}^k = 0$ recaímos ao caso trivial, de modo que os autovalores de $G(x^k)$ são dados por $\lambda_1 = x_{11}^k$, $\lambda_2 = x_{22}^k$ e $\lambda_3 = x_{33}^k$ e $E^k = \mathbb{I}_3$ com $\bar{E} = \mathbb{I}_3$ gerando assim o conjunto $\{v_{ii}(\bar{x}, \bar{E})\} = \{[1, 0, 0, 0]^T, [0, 1, 0, 0]^T, [0, 0, 1, 0]^T\}$ que é linearmente independente. Por outro lado, seja $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$ tal que $x_{13}^k \neq 0$ para todo $k \in \mathbb{N}$ temos que os autovalores de $G(x^k)$ são dados por:

$$\begin{aligned} \lambda_1^k &= \frac{1}{2} (x_{11}^k + x_{33}^k - \delta^k) \\ \lambda_2^k &= x_{22}^k \\ \lambda_3^k &= \frac{1}{2} (x_{11}^k + x_{33}^k + \delta^k) \end{aligned}$$

onde $\delta^k = \sqrt{4(x_{13}^k)^2 + (x_{11}^k)^2 + (x_{33}^k)^2 - 2x_{11}^k x_{33}^k}$, cujo seus respectivos autovetores serão da-

dos por:

$$v_1^k = \begin{bmatrix} -\frac{-x_{11}^k + x_{33}^k + \delta^k}{2x_{13}^k} \\ 0 \\ 1 \end{bmatrix} \quad v_2^k = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad v_3^k = \begin{bmatrix} -\frac{-x_{11}^k + x_{33}^k - \delta^k}{2x_{13}^k} \\ 0 \\ 1 \end{bmatrix}$$

assim, para $x_{13}^k > 0$ (o caso $x_{13}^k < 0$ é análogo) teremos que:

$$\lim_{k \rightarrow \infty} \frac{-x_{11}^k + x_{33}^k + \delta^k}{2x_{13}^k} = \lim_{k \rightarrow \infty} \frac{\|x_{13}^k\|}{x_{13}^k} = 1 \quad e \quad \lim_{k \rightarrow \infty} \frac{-x_{11}^k - x_{33}^k - \delta^k}{2x_{13}^k} = \lim_{k \rightarrow \infty} -\frac{\|x_{13}^k\|}{x_{13}^k} = -1$$

portanto $v_1^k \rightarrow [-1, 0, 1]$, $v_2^k \rightarrow [0, 1, 0]$ e $v_3^k \rightarrow [0, 0, 1]$. Deste modo obtemos $E^k \in \mathbb{S}^{3 \times 3}$ de maneira que:

$$E^k \doteq \begin{bmatrix} \frac{v_1^k}{\|v_1^k\|} & \frac{v_2^k}{\|v_2^k\|} & \frac{v_3^k}{\|v_3^k\|} \end{bmatrix}$$

e,

$$E^k \rightarrow \bar{E} \doteq \begin{bmatrix} \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

e obtendo os vetores de interesse, teremos:

$$v_{11}(\bar{x}, \bar{E}) = \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ -2 \end{bmatrix} \quad v_{22}(\bar{x}, \bar{E}) = \frac{1}{2} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad v_{33}(\bar{x}, \bar{E}) = \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

que facilmente, verificam-se linearmente independentes.

Note que o exemplo anterior, nos diz algo sobre a *Weak-Nondegeneracy*, visto que a estrutura local das sequências $\{E^k\}_{k \in \mathbb{N}}$ coincide com a estrutura da restrição, ou seja, possuem a mesma esparsidade.

Aproveitando esta noção de como se comporta a esparsidade da restrição em relação a *Weak-Nondegeneracy*, vejamos agora como verificar *Weak-Nondegeneracy* para o caso onde a estrutura da restrição é uma matriz diagonal.

Proposição 3.11 (ANDREANI, HAESER, MITO e RAMÍREZ, 2023, Obs. 3.2). *Considere a seguinte restrição para o problema (NSDP).*

$$G(x) \doteq \begin{bmatrix} g_1(x) & & \\ & \ddots & \\ & & g_m(x) \end{bmatrix} \succeq 0$$

Então um ponto \bar{x} viável satisfaz *Weak-Nondegeneracy* se, e somente se, o conjunto $\{\nabla g_i(\bar{x}) : g_i(\bar{x}) = 0\}$ é linearmente independente.

Demonstração. De fato, note que para $r = m$ verifica-se trivialmente. No entanto, o caso mais interessante é dado quando $r < m$. Desta forma podemos assumir, sem perda de generalidade que $\{i \in \{1, \dots, m\} : g_i(\bar{x}) = 0\} = \{r + 1, \dots, m\}$ onde $r = \text{rank}(G(\bar{x}))$ assim:

$$G(\bar{x}) \doteq \begin{bmatrix} g_1(\bar{x}) & & & \\ & \ddots & & \\ & & g_r(\bar{x}) & \\ & & & 0 \end{bmatrix} \succeq 0$$

assim para toda sequência $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$ podemos tomar

$$E^k \doteq \begin{bmatrix} 0 \\ \mathbb{I}_{m-r} \end{bmatrix} \rightarrow \bar{E} = \begin{bmatrix} 0 \\ \mathbb{I}_{m-r} \end{bmatrix} \in \mathbb{R}^{m \times m-r}$$

Agora, note que

$$\{v_{ii}(\bar{x}, \bar{E}) : i \in \{1, \dots, m-r\}\} = \{\nabla g_{r+1}(\bar{x}), \dots, \nabla g_m(\bar{x})\}$$

assim, se $\{\nabla g_i(\bar{x}) : g_i(\bar{x}) = 0\}$ é linearmente independente então, \bar{x} satisfaz *Weak-Nondegeneracy*. Por outro lado, suponha que \bar{x} satisfaça *Weak-Nondegeneracy*, isto é, para qualquer sequência $\{x^k\}_{k \in \mathbb{N}} \rightarrow \bar{x}$ e qualquer $\{E^k\}_{k \in \mathbb{N}} \rightarrow \bar{E} \doteq [\bar{e}_1, \dots, \bar{e}_{m-r}]$ têm-se que conjunto $\{v_{ii}(\bar{x}) : i \in \{1, \dots, m-r\}\}$ é linearmente independente. A demonstração consiste em observar que \bar{E} terá a forma

$$\bar{E} = \begin{bmatrix} 0 \\ Q \end{bmatrix} \quad \text{onde } Q \in \mathbb{R}^{m-r \times m-r} \text{ é uma matriz ortogonal}$$

oriunda diretamente da estrutura diagonal de $G(\bar{x})$ onde $g_i(\bar{x}) \neq 0$ para todo $i \in \{1, \dots, m-r\}$. Assim, perceba que

$$v_{ii}(\bar{x}, \bar{E}) = \sum_{j=r+1}^m \nabla g_j(\bar{x}) Q_{i,j-r}^2 = Gg(\bar{x})^T (Q_i \odot Q_i) \quad (3.63)$$

onde Q_i é a i -ésima coluna de Q e $Dg(\bar{x})$ é a matriz Jacobiana de $g(x) \doteq (g_{r+1}(x), \dots, g_m(x))$ em \bar{x} com $i \in \{1, \dots, m-r\}$. (" $Q_i \odot Q_i$ " denota o produto de *Hadamard*, que se têm, multiplicando entrada por entrada do vetor Q_i). Logo,

$$\text{span}\{v_{ii}(\bar{x}) : i \in \{1, \dots, m-r\}\} \subseteq \text{Im} Dg(\bar{x})^T$$

assim

$$\begin{aligned} m-r &= \dim(\text{span}\{v_{ii}(\bar{x}) : i \in \{1, \dots, m-r\}\}) \\ &\leq \dim(\text{Im} Dg(\bar{x})^T) \\ &= \text{rank}(Dg(\bar{x})^T) \leq m-r \end{aligned}$$

portanto, $\text{rank}(Dg(\bar{x})^T) = m-r$ ou seja $\{\nabla g_{r+1}(\bar{x}), \dots, \nabla g_m(\bar{x})\}$ é linearmente independente. \square

Por fim, para o término deste capítulo, serão propostos novos exemplos e resultados, cujo principal intuito será estabelecer uma correspondência entre *Nondegeneracy-CQ* e *Weak-Nondegeneracy* assim como disposto para *NSOCP*, algo que, até então, ainda não foi estabelecido na literatura.

3.2.3 Resultados interessantes

Seja $G(x) \in \mathbb{S}_+^m$ a restrição para o problema (*NSDP*), dado $\bar{x} \in \mathcal{F}$ considere $\{x^k\}_{k \in \mathbb{N}}$ tal que $x^k = \bar{x} + t^k d$ onde $\{t^k\}_{k \in \mathbb{N}} \rightarrow 0_+$ logo, localmente em torno de \bar{x} teremos

$$G(x^k) = G(\bar{x}) + t_k DG(\bar{x})d + o(t_k)$$

Deste modo, seja \bar{x} um ponto viável tal que $G(\bar{x}) = 0$, têm-se

$$G(x^k) = t^k \left(DG(\bar{x})d + \frac{o(t_k)}{t^k} \right) \quad (3.64)$$

assim, definindo a sequência $\{R_d^k\}_{k \in \mathbb{N}}$ tal que, $R_d^k \doteq DG(\bar{x})d + \frac{o(t_k)}{t^k}$ teremos $G(x^k) = t^k R_d^k$ e $R_d^k \rightarrow R_d \doteq DG(\bar{x})d$. Com isto, podemos definir

$$\begin{aligned} \mathcal{D}_d^k &\doteq \{U \in \mathbb{R}^{m \times m} : U^T R_d^k U = \text{Diag}(z), U^T U = \mathbb{I}\} \\ \mathcal{D}_d &\doteq \{U \in \mathbb{R}^{m \times m} : U^T R_d U = \text{Diag}(z), U^T U = \mathbb{I}\} \end{aligned} \quad (3.65)$$

os conjuntos das matrizes ortogonais que diagonalizam R_d^k e R_d respectivamente, logo, percebe-se que $\mathcal{D}_d^k \subseteq \mathcal{D}_d$.

Observação 3.4. Vale ressaltar que uma correspondência entre *Robinson-CQ* e *Weak-Robinson CQ* ainda não foi bem estabelecida nem sobre o contexto de *NSOCP*, tal qual sobre o contexto de *NSDP* até a presente data deste trabalho.

Os conjuntos dados em (3.65) serão úteis para a discussão de alguns exemplos interessantes, cujos terão como principal motivação, caracterizar, embora em casos particulares, o conceito de *Weak-Robinson CQ* que possivelmente poderá ser útil para correlacioná-lo com *Robinson-CQ*.

Exemplo 3.4. Considere uma restrição $G(x) \in \mathbb{S}_+^2$ e uma sequência $\{x^k\}_{k \in \mathbb{N}}$ de modo que $\{x^k\} \rightarrow \bar{x}$ com \bar{x} um ponto viável como definido na observação acima, suponha ainda que

$$R_d^k \doteq \begin{bmatrix} 1 & \frac{1}{k} \\ \frac{1}{k} & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbb{I}_2$$

Note que os autovalores de R_d^k são dados por $\lambda_i(R_d^k) = 1 \pm \frac{1}{k}$ distintos, dessa forma

$$\mathcal{D}_d^k \doteq \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$$

ou seja, $\mathcal{D}_d^k \subset \mathcal{D}_{\mathbb{I}_2} = \mathbb{R}^{2 \times 2}$.

Exemplo 3.5. Considere o problema (NSDP) com a seguinte restrição

$$G(x) \doteq \begin{bmatrix} g_{11}(x) & g_{12}(x) \\ g_{12}(x) & g_{22}(x) \end{bmatrix} \succeq 0$$

e um ponto viável \bar{x} , tal que os gradientes da diagonal $\{\nabla g_{11}(\bar{x}), \nabla g_{22}(\bar{x})\}$, são positivo-linearmente dependentes. Ou seja, $\exists \theta > 0$ tal que $\nabla g_{11}(\bar{x}) = -\theta \nabla g_{22}(\bar{x})$. Deste modo, perceba que se, $\{\nabla g_{11}(\bar{x}), \nabla g_{12}(\bar{x})\}$ são linearmente independentes, têm-se:

$$R_d = \begin{bmatrix} \nabla g_{11}(\bar{x})^T d & \nabla g_{12}(\bar{x})^T d \\ \nabla g_{12}(\bar{x})^T d & \nabla g_{22}(\bar{x})^T d \end{bmatrix} \quad (3.66)$$

Neste caso, como $\{\nabla g_{11}(\bar{x}), \nabla g_{12}(\bar{x})\}$ são LI, existe uma direção $d \in \mathbb{R}^n$ tal que, $\nabla g_{11}(\bar{x})^T d \neq 0$ e $\nabla g_{12}(\bar{x})^T d = 0$, portanto, sem perda de generalidade, podemos escrever

$$R_d = \begin{bmatrix} a & 0 \\ 0 & -\theta a \end{bmatrix}$$

ou seja, $\lambda_1(R_d) = a$ e $\lambda_2(R_d) = -\theta a$, isto é, $\mathcal{D}_d = \{\mathbb{I}_2\}$, assim, para $\bar{E} = \mathbb{I}_2$ os vetores $\{v_{ii}(\bar{x}, \bar{E})\} = \{\nabla g_{11}(\bar{x}), \nabla g_{22}(\bar{x})\}$ não são positivo-linearmente independentes, dado por hipótese, logo não vale Weak-Robinson's CQ. Agora, por outro lado, se $\{\nabla g_{11}(\bar{x}), \nabla g_{12}(\bar{x})\}$ são linearmente dependentes, considere então a direção

$$d = \frac{\nabla g_{11}(\bar{x})}{\|\nabla g_{11}(\bar{x})\|}$$

Assim, sem perda de generalidade teremos

$$R_d = \begin{bmatrix} 1 & \xi \\ \xi & -\theta \end{bmatrix}$$

para algum $\xi \in \mathbb{R}$. Logo, os autovalores de R_d são dados por,

$$\lambda(R_d) = \frac{(1 - \theta) \pm \sqrt{(1 - \theta)^2 + 4(\theta + \xi^2)}}{2}$$

Agora, note que, por hipótese a matriz $D_{x_i}(\bar{x}) = [\partial_{x_i} g_{ij}(\bar{x})]_{i,j=1,2}$ pode ser escrita por

$$D_{x_i}(\bar{x}) = \begin{bmatrix} \partial_{x_i} g_{11}(\bar{x}) & \xi \partial_{x_i} g_{11}(\bar{x}) \\ \xi \partial_{x_i} g_{11}(\bar{x}) & -\theta \partial_{x_i} g_{11}(\bar{x}) \end{bmatrix} = \partial_{x_i} g_{11}(\bar{x}) \begin{bmatrix} 1 & \xi \\ \xi & -\theta \end{bmatrix} \doteq \partial_{x_i} g_{11}(\bar{x}) J$$

para $i \in \{1, \dots, n\}$. Assim, os vetores $\{v_{ii}(\bar{x}, \bar{E})\}$ serão dados por

$$\begin{aligned} v_{ii}(\bar{x}, \bar{E}) &= \left[\bar{e}_i^T (\partial_{x_1} g_{11}(\bar{x}) J) \bar{e}_i, \dots, \bar{e}_i^T (\partial_{x_n} g_{11}(\bar{x}) J) \bar{e}_i \right]^T \\ &= \left[\partial_{x_1} g_{11}(\bar{x}) \bar{e}_i^T J \bar{e}_i, \dots, \partial_{x_n} g_{11}(\bar{x}) \bar{e}_i^T J \bar{e}_i \right]^T \\ &= \left[\partial_{x_1} g_{11}(\bar{x}) \bar{e}_i^T \lambda_i(R_d) \bar{e}_i, \dots, \partial_{x_n} g_{11}(\bar{x}) \bar{e}_i^T \lambda_i(R_d) \bar{e}_i \right]^T \\ &= \lambda_i(R_d) \nabla g_{11}(\bar{x}) \end{aligned}$$

Portanto $\{v_{ii}(\bar{x}, \bar{E})\}$ são positivo linearmente dependentes, logo \bar{x} não satisfaz Weak-Robinson's CQ.

O exemplo acima, prova que, para o caso \mathbb{S}_+^2 , Weak-Robinson's CQ nunca poderá ser válido quando os gradientes da diagonal principal são positivo linearmente dependentes.

Como foi visto nos exemplos acima, Nondegeneracy-CQ e Weak-Nondegeneracy não possuem equivalência direta em nenhum dos cones apresentados até então, entretanto no capítulo (2) na proposição (2.6) foi possível estabelecer, adicionando uma hipótese a mais sobre a decomposição da restrição a weak-Nondegeneracy, uma relação de equivalência entre estas sobre o contexto de (NSOCP).

Para terminarmos o capítulo, apresentaremos originalmente o próximo resultado, que complementa o trabalho ANDREANI, HAESER, MITO e RAMÍREZ, 2023, fornecendo, em análogo ao que foi feito na proposição (2.6) do capítulo anterior, uma correspondência entre Weak-Nondegeneracy e Nondegeneracy-CQ para o contexto de (NSDP).

Proposição 3.12. *Seja \bar{x} um ponto viável para (NSDP) tal que $r = \text{rank}(G(\bar{x}))$. Então \bar{x} satisfaz Nondegeneracy-CQ se, e somente se, \bar{x} satisfaz Weak-Nondegeneracy e,*

- *Para toda matriz $E \in \mathbb{R}^{m \times (m-r)}$ tal que $\text{Ker}G(\bar{x})$ é gerado por E , existe uma direção $d \in \mathbb{R}^n$ de modo que $E^T DG(\bar{x})dE$ é diagonal e simples (Uma matriz diagonal cujos elementos são diferentes entre si).*

Demonstração. Perceba que, pela proposição (3.4), fica claro que, Nondegeneracy implica em Weak-Nondegeneracy. Agora, por outro lado considere que \bar{x} satisfaz Weak-Nondegeneracy em conjunto com a hipótese adicional acima. Suponha que \bar{x} não satisfaz Nondegeneracy, ou seja, existe E tal que $\text{Ker}G(\bar{x})$ é gerado por E e o conjunto $\{v_{ii}(\bar{x}, E) : i \in \{1, \dots, m-r\}\}$ é linearmente dependente. Primeiro, mostraremos que E é uma base contínua. Seja $d \in \mathbb{R}^n$ tal que $E^T DG(\bar{x})dE$ é diagonal e simples. Considere $x^k \doteq \bar{x} + t_k, t_k \rightarrow 0_+$, deste modo, segue-se que,

$$G(x^k) = G(\bar{x}) + t_k DG(\bar{x})d + o(t_k)$$

assim

$$\begin{aligned} E^T G(x^k)E &= E^T G(\bar{x})E + t_k E^T DG(\bar{x})dE + o(t_k) \\ \frac{1}{t_k} E^T G(x^k)E &= E^T DG(\bar{x})dE + \frac{o(t_k)}{t_k} \end{aligned}$$

onde utilizamos o fato de $G(\bar{x})E = 0$. Desta forma, para k suficientemente grande, podemos assumir que $\frac{1}{t_k} E^T G(x^k)E$ é simples, visto que $E^T DG(\bar{x})dE$ é diagonal e simples por hipótese. Logo, existe uma única matriz ortogonal P^k tal que:

$$\frac{1}{t_k} E^T G(x^k)E = P^k \Lambda^k (P^k)^T \quad \text{onde,} \quad \begin{cases} P^k \rightarrow I \\ \Lambda^k \rightarrow E^T DG(\bar{x})dE. \end{cases}$$

ou seja, $\frac{1}{t_k}(P^k)^T E^T G(x^k) E P^k = \Lambda^k$ assim, escrevendo $E^k \doteq E P^k \rightarrow E$ teremos

$$\frac{1}{t_k}(E^k)^T G(x^k) E^k = \Lambda^k$$

ou seja, E^k também corresponde aos $m - r$ autovetores de $G(x^k)$, logo por hipótese, o conjunto $\{v_{ii}(\bar{x}, E)\}$ deve ser linearmente independente, o que é uma contradição, logo \bar{x} satisfaz *Nondegeneracy*.

□

Capítulo 4

Conclusão

O objetivo final deste trabalho, é realizar um levantamento teórico na literatura sobre as principais condições de qualificações para problemas não lineares semi-definidos. Contudo, estas condições, apresentam-se ser desnecessariamente generalistas para problemas de grande porte. Desta forma, buscando uma remediação a isso foi explorado, uma abordagem, que consiste em enfraquecer estas condições, de modo a preservar ao máximo a estrutura do problema. Para isto, o principal alicerce, se dá, por fazer uma análise local, e fazer análises sobre pontos limites de "auto-decompositores ativos" da restrição cônica. Como foi visto, reduzimos a análise sobre pontos limites de sequência de autovetores.

Esta abordagem, todavia, acredita-se ser(o autor), possivelmente estendida para um problema cônico qualquer, desde que o cone, tenha por si só a propriedade de simetria, visto que, a principal hipótese utilizada, foi uma espécie de decomposição espectral, sobre o cone. É conhecido na literatura que, cones simétricos, sempre admitem uma decomposição espectral veja (LOURENÇO *et al.*, 2018). Com isto, paira-nos a motivação, para generalizar os resultados obtidos para (NSOCP) e (NSDP) sobre cones simétricos quaisquer, em especial, as condições de equivalências (2.6) e (3.12) obtidas para (NSOCP) e (NSDP) com suas versões mais fracas. Acredita-se também, que há muito o que ser explorado, sobre perspectiva de se analisar estas formas fracas de condições de qualificações na existência e unicidade de Multiplicadores de Lagrange, e assim, possivelmente, construir novas condições de qualificações que exploram ao máximo a estrutura do problema.

Referências

- [ALIZADEH e GOLDFARB 2003] F. ALIZADEH e D. GOLDFARB. “Second-order cone programming”. →Mathematical Programming 95.1 (1 de jan. de 2003), pp. 3–51. ISSN: 0025-5610, 1436-4646. DOI: [10.1007/s10107-002-0339-5](https://doi.org/10.1007/s10107-002-0339-5). URL: <http://link.springer.com/10.1007/s10107-002-0339-5> (acesso em 06/03/2023) (citado na pg. 15).
- [ANDREANI, GÓMEZ *et al.* 2022] Roberto ANDREANI, Walter GÓMEZ, Gabriel HAESER, Leonardo M MITO e Alberto RAMOS. “On optimality conditions for nonlinear conic programming”. →Mathematics of Operations Research 47.3 (2022), pp. 2160–2185 (citado nas pgs. 38, 42, 43).
- [ANDREANI, HAESER e MARTÍNEZ 2011] Roberto ANDREANI, Gabriel HAESER e José Mário MARTÍNEZ. “On sequential optimality conditions for smooth constrained optimization”. →Optimization 60.5 (2011), pp. 627–641 (citado na pg. 38).
- [ANDREANI, HAESER, MITO e RAMÍREZ 2023] Roberto ANDREANI, Gabriel HAESER, Leonardo M. MITO e Héctor RAMÍREZ. “Weak notions of nondegeneracy in nonlinear semidefinite programming”. →Mathematical Programming (mai. de 2023). ISSN: 1436-4646. DOI: [10.1007/s10107-023-01970-4](https://doi.org/10.1007/s10107-023-01970-4). URL: <https://doi.org/10.1007/s10107-023-01970-4> (citado nas pgs. 2, 14, 33, 39–42, 44, 48).
- [ANDREANI, HAESER, MITO, RAMÍREZ e SILVEIRA 2022] Roberto ANDREANI, Gabriel HAESER, Leonardo M. MITO, Héctor RAMÍREZ e Thiago P. SILVEIRA. “Global Convergence of Algorithms Under Constant Rank Conditions for Nonlinear Second-Order Cone Programming”. →Journal of Optimization Theory and Applications 195.1 (out. de 2022), pp. 42–78. ISSN: 1573-2878. DOI: [10.1007/s10957-022-02056-5](https://doi.org/10.1007/s10957-022-02056-5). URL: <https://doi.org/10.1007/s10957-022-02056-5> (citado nas pgs. 2, 14, 15, 21–26).
- [ANDREANI, HAESER e VIANA 2020] Roberto ANDREANI, Gabriel HAESER e Daiana S. VIANA. “Optimality conditions and global convergence for nonlinear semidefinite programming”. en. →Mathematical Programming 180.1-2 (mar. de 2020), pp. 203–235. ISSN: 0025-5610, 1436-4646. DOI: [10.1007/s10107-018-1354-5](https://doi.org/10.1007/s10107-018-1354-5). URL: <http://link.springer.com/10.1007/s10107-018-1354-5> (acesso em 01/08/2023) (citado nas pgs. 2, 37, 38).

- [BONNANS e SHAPIRO 2000] J. Frédéric BONNANS e Alexander SHAPIRO. →Perturbation Analysis of Optimization Problems. New York, NY: Springer New York, 2000. ISBN: 978-1-4612-7129-1 978-1-4612-1394-9. DOI: [10.1007/978-1-4612-1394-9](https://doi.org/10.1007/978-1-4612-1394-9). URL: <http://link.springer.com/10.1007/978-1-4612-1394-9> (acesso em 12/04/2023) (citado nas pgs. 19, 31, 40).
- [FARAUT 1994] Jacques FARAUT. “Analysis on symmetric cones”. →Oxford mathematical monographs (1994) (citado na pg. 18).
- [HORN e JOHNSON 2012] Roger A HORN e Charles R JOHNSON. →Matrix analysis. Cambridge university press, 2012 (citado nas pgs. 7, 8).
- [IZMAILOV e SOLODOV 2005] Alexey IZMAILOV e Mikhail SOLODOV. →Otimização, volume 1: condições de otimalidade, elementos de análise convexa e de dualidade. Impa, 2005 (citado na pg. 10).
- [LEWIS 1996] Adrian S LEWIS. “Derivatives of spectral functions”. →Mathematics of Operations Research 21.3 (1996), pp. 576–588 (citado na pg. 11).
- [LOURENÇO *et al.* 2018] Bruno F LOURENÇO, Ellen H FUKUDA e Masao FUKUSHIMA. →Optimality conditions for problems over symmetric cones and a simple augmented Lagrangian method. 2018 (citado nas pgs. 10, 51).
- [ROBINSON 1976a] Stephen M. ROBINSON. “First Order Conditions for General Nonlinear Optimization”. en. →SIAM Journal on Applied Mathematics 30.4 (jun. de 1976), pp. 597–607. ISSN: 0036-1399, 1095-712X. DOI: [10.1137/0130053](https://doi.org/10.1137/0130053). URL: <http://epubs.siam.org/doi/10.1137/0130053> (acesso em 31/07/2023) (citado na pg. 1).
- [ROBINSON 1976b] Stephen M. ROBINSON. “Stability Theory for Systems of Inequalities, Part II: Differentiable Nonlinear Systems”. en. →SIAM Journal on Numerical Analysis 13.4 (set. de 1976), pp. 497–513. ISSN: 0036-1429, 1095-7170. DOI: [10.1137/0713043](https://doi.org/10.1137/0713043). URL: <http://epubs.siam.org/doi/10.1137/0713043> (acesso em 03/08/2023) (citado na pg. 13).
- [SHAPIRO 1997] Alexander SHAPIRO. “First and second order analysis of nonlinear semidefinite programs”. →Mathematical programming 77 (1997), pp. 301–320 (citado nas pgs. 2, 31, 33).
- [SHAPIRO e FAN 1995] Alexander SHAPIRO e Michael K. H. FAN. “On Eigenvalue Optimization”. en. →SIAM Journal on Optimization 5.3 (ago. de 1995), pp. 552–569. ISSN: 1052-6234, 1095-7189. DOI: [10.1137/0805028](https://doi.org/10.1137/0805028). URL: <http://epubs.siam.org/doi/10.1137/0805028> (acesso em 10/08/2023) (citado na pg. 31).
- [STEWART e SUN 1990] Gilbert W STEWART e Ji-guang SUN. “Matrix perturbation theory”. →Academic Press (1990) (citado na pg. 8).

REFERÊNCIAS

- [SÜLI e MAYERS 2007] Endre SÜLI e David F. MAYERS. →An introduction to numerical analysis. en. Repr. Cambridge: Cambridge Univ. Press, 2007. ISBN: 978-0-521-81026-5 978-0-521-00794-8 (citado na pg. 7).
- [THEOBALD 1975] C. M. THEOBALD. “An inequality for the trace of the product of two symmetric matrices”. en. →Mathematical Proceedings of the Cambridge Philosophical Society 77.2 (mar. de 1975), pp. 265–267. ISSN: 0305-0041, 1469-8064. DOI: [10.1017/S0305004100051070](https://doi.org/10.1017/S0305004100051070). URL: https://www.cambridge.org/core/product/identifier/S0305004100051070/type/journal_article (acesso em 01/08/2023) (citado na pg. 9).
- [WACHSMUTH 2013] Gerd WACHSMUTH. “On LICQ and the uniqueness of Lagrange multipliers”. en. →Operations Research Letters 41.1 (jan. de 2013), pp. 78–80. ISSN: 01676377. DOI: [10.1016/j.orl.2012.11.009](https://doi.org/10.1016/j.orl.2012.11.009). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0167637712001459> (acesso em 01/08/2023) (citado na pg. 35).