Instituto Alberto Luiz Coimbra de
Pós-Graduação e Pesquisa de Engenharia
COPPE UFRJ

RELAXAÇÕES TRATÁVEIS PARA PROBLEMAS DE OTIMIZAÇÃO CÚBICOS RESTRITOS À ESFERA

Orlando Sarmiento Chumbes

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Sistemas e Computação.

Orientador: Marcia Helena Costa Fampa

Rio de Janeiro
Junho de 2019

RELAXAÇÕES TRATÁVEIS PARA PROBLEMAS DE OTIMIZAÇÃO CÚBICOS RESTRITOS À ESFERA

Orlando Sarmiento Chumbes

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Examinada por:

Prof. Marcia Helena Costa Fampa, D.Sc.

Prof. Paulo Roberto Oliveira, D.Sc.

Prof. Fernanda Maria Pereira Raupp, D.Sc.

Prof. Laura Silvia Bahiense da Silva Leite, D.Sc.

Prof. Helder Manoel Venceslau, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
JUNHO DE 2019

Chumbes, Orlando Sarmiento

Relaxações tratáveis para problemas de otimização cúbicos restritos à esfera/Orlando Sarmiento Chumbes. – Rio de Janeiro: UFRJ/COPPE, 2019.

X, 89 p. 29, 7cm.

Orientador: Marcia Helena Costa Fampa

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2019.

Referências Bibliográficas: p. 85 – 89.

1. Programação global com polinômios. 2. Problema de otimização cúbico. 3. Desigualdades de reformulação e linearização (RLT). 4. Novas abordagens de relaxação. 5. Método de decomposição Lagrangeana. I. Fampa, Marcia Helena Costa. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

Dedico este trabalho à memória de meus pais Alipio Sarmiento e Marcosa Chumbes.
Aos meus irmãos César Alberto, Marisol e José Luis, e à minha filha Maria Isabel.

Agradecimentos

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), 140408/2015-0, e à Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), E-26/200.209/2017, pelo suporte financeiro que possibilitou a realização deste trabalho.

Quero agradecer à minha família, à minha mãe Marcosa e ao meu pai Alipio que fizeram tantos sacrifícios para que eu pudesse ter bom estudo e assim ter boas oportunidades como venho tendo. Aos meus irmãos, César Alberto, José Luis e Marisol.

À Professora Marcia Helena Costa Fampa por ter me orientado durante o doutorado. Muito obrigado pela orientação, apoio, paciência, amizade e sobre tudo pela confiança depositada.

Aos membros da banca examinadora: Prof. Paulo Roberto Oliveira, Prof. Fernanda Maria Pereira Raupp, Prof. Laura Silvia Bahiense da Silva Leite e Prof. Helder Manoel Venceslau, cujas observações, críticas e sugestões enriqueceram esta tese.

Aos professores do Programa de Engenharia de Sistemas e Computação da COPPE-UFRJ, em especial aos da linha de Otimização.

Agradeço à minha noiva Liliana, muito obrigado pelo apoio, carinho, companhia e consideração.

Aos companheiros do laboratório de otimização Daniela, Miluzca, Nancy, Tanilson e Lennin por todo esse período de estar ali todos os dias e partilhar momentos de estudos, histórias e tantos outros fatos.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), 88881.187742/2018-01, pelo suporte financeiro durante meu período de Doutorado Sanduíche na Alemanha.

Agradeço ao professor Christoph Buchheim por ter me orientado na universidade TU Dortmund. Muito obrigado por suas idéias, sugestões e críticas que enriqueceram esta tese. Finalmente agradeço aos amigos que fiz na Alemanha: Austé, Federica, Sabine, Dorothee, Jonas, Felix, Stefan, Efrain, Fernando, Moritz e Nick.

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

RELAXAÇÕES TRATÁVEIS PARA PROBLEMAS DE OTIMIZAÇÃO CÚBICOS RESTRITOS À ESFERA

Orlando Sarmiento Chumbes

Junho/2019

Orientador: Marcia Helena Costa Fampa

Programa: Engenharia de Sistemas e Computação

Neste trabalho, estamos interessados em desenvolver técnicas de relaxação para problemas de otimização cúbicos restritos à esfera. Cabe ressaltar que estes problemas são NP-difíceis em contraste com suas versões quadráticas. Apesar da não convexidade o problema quadrático restrito à esfera, conhecido na literatura como subproblema de região de confiança, pode ser resolvido eficientemente já que o problema dual associado é côncavo com gap de dualidade nulo.

Na literatura, diversas técnicas de relaxação têm sido estudadas para problemas polinomiais, baseadas em desigualdades RLT e via Programação Semidefinida, obtendo-se bons resultados em programação quadrática e recentemente estendida para programação com polinômios de maior grau.

Nesta tese, apresentamos duas técnicas para encontrar limites inferiores para os problemas cúbicos restritos à esfera. A primeira técnica desenvolvida lineariza tanto a função objetivo quanto a função que define a restrição (funções cúbica e quadrática respectivamente), esta técnica é uma adaptação das metodologias propostas na literatura para programas de otimização polinomial. A segunda técnica desenvolvida utiliza quatro diferentes abordagens de decomposição para a função objetivo dos problemas cúbicos. Diferentemente da técnica de linearização, desenvolvemos abordagens que decompõem a função objetivo sem modificar a região viável, com a intenção de obter melhores limites. Resultados computacionais são apresentados nos quais mostramos a eficiência das abordagens propostas.

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

TRACTABLE RELAXATIONS FOR CUBIC OPTIMIZATION PROBLEMS SUBJECT TO A NORM CONSTRAINT

Orlando Sarmiento Chumbes

June/2019

Advisor: Marcia Helena Costa Fampa

Department: Systems Engineering and Computer Science

In this work, we are interested in developing relaxation techniques for cubic optimization problems subject to a norm constraint. It should be noted that these problems are NP-hard in constrast to their quadratic variants. In spite of the non-convexity the latter quadratic problem, which is the well-known trust region subproblem, can be solved efficiently since it has a concave dual problem with no duality gap.

In the literature, several relaxation techniques have been studied to polynomials problems, based on RLT inequalities and via Semidefinite Programming, obtaining good results in quadratic programming, and recently extended for programs with higher degree polynomials.

In this thesis, we present two techniques to find lower bounds for the cubic optimization problems subject to a norm constraint. The first technique linearizes both objective and constraint functions (cubic and quadratic functions respectively), this technique is an adaptation of approaches proposed in the literature for polynomial optimization programs. The second technique uses four different decomposition approaches for the objective functions of the cubic optimization problems. Unlike the first technique, these approaches that decompose objective function without modifying the feasible region, with the intention of obtaining better lower bounds. Computational results are presented in which we show the efficiency of the proposed approaches.

Sumário

Lista de Tabelas

Capítulo 1

Introdução

Diversos problemas encontrados em diferentes áreas da indústria podem ser formulados como um problema de otimização. Em problemas de otimização, procuramos minimizar ou maximizar uma função de variáveis em um determinado domínio normalmente definido por um conjunto de restrições nas variáveis. O enfoque das pesquisas recentemente foi orientado a estudar problemas de Otimização Não Linear Inteira Mista, os quais são problemas de otimização em que a função a otimizar é não linear e parte das variáveis de decisão é discreta (assumem valores inteiros) e as demais são contínuas (assumem valores em um intervalo de números reais). Estes problemas são usualmente difíceis de resolver (na teoria e na prática) devido à presença de dois tipos de não convexidade: primeiro, a função objetivo ou restrições podem ser não convexas, e segundo, a presença de restrições inteiras leva o conjunto de soluções viáveis a ser não convexo.

Dentre as investigações desenvolvidas em otimização não linear inteira, podemos destacar o trabalho realizado por Buchheim et al. [1] no qual os autores apresentam um eficiente algoritmo de *branch-and-bound* para resolver o problema quadrático com variáveis inteiras na forma:

$$
\begin{aligned}
\min \quad & f(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} a_i x_i \\
\text{s.a.} \quad & l_j \leq x_j \leq u_j, \; j = 1, 2, \ldots, n \\
& x \in \mathbb{Z}^n
\end{aligned}
\tag{1.1}
$$

em que $l, u \in \mathbb{Z}^n$.

Para a obtenção dos limites inferiores do problema (1.1), o primeiro passo é escolher um elipsoide $\mathcal{E}$ que contenha todas as soluções viáveis do problema (1.1),

$$[l, u] \subseteq \mathcal{E} = \{x \in \mathbb{R}^n \, : \, (x - x^0)^\top H (x - x^0) \leq 1\},$$

em que $H \succeq 0$ e x^0 denota o centro do elipsoide.

Definimos a relaxação do problema (1.1) como

$$\begin{aligned} \min \quad & f(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} a_i x_i \\ \text{s.a.} \quad & x \in \mathcal{E}, \end{aligned} \tag{1.2}$$

onde o aspecto fundamental para obter um limite inferior apertado do problema (1.1) é fazer a melhor escolha do elipsoide $\mathcal{E}$. Por meio de algumas transformações, obtem-se um problema equivalente a (1.2) na forma

$$\begin{aligned} \min \quad & f(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} a_i x_i \\ \text{s.a.} \quad & x^\top \mathcal{H} x = 1, \end{aligned} \tag{1.3}$$

para uma apropriada matriz diagonal $\mathcal{H}$.

Nesta tese estamos interessados em obter uma relaxação do problema cúbico restrito à esfera na forma:

$$\begin{aligned} \text{PC}: \quad \min \quad & f(x) := \mathcal{A}_3 x^3 + \mathcal{A}_2 x^2 + \mathcal{A}_1 x = \sum_{i,j,k=1}^{n} a_{ijk} x_i x_j x_k + \sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} a_i x_i \\ \text{s.a.} \quad & \|x\| = 1, \end{aligned} \tag{1.4}$$

em que $n \geq 2$ e $\mathcal{A}_i$ é um tensor simétrico real de i-ésima ordem para cada $i = 1, 2, 3$. Em outras palavras, $\mathcal{A}_3$ é um tensor simétrico real $n \times n \times n$ dimensional, $\mathcal{A}_2$ é uma matriz simétrica real $n \times n$ dimensional, e $\mathcal{A}_1$ é um vetor n dimensional.

No caso particular quando $\mathcal{A}_2$ é uma matriz nula e $\mathcal{A}_1$ é um vetor nulo, o problema PC é denominado problema um-esférico cúbico (PEC) e foi mostrado por Zhang et al. [2] usando um resultado de Nesterov [3], que o problema PEC é NP-difícil. Isto implica que no caso geral PC também é um problema NP-difícil, em contraste com sua versão quadrática, a qual é o bem conhecido problema de região de confiança. Apesar de não convexidade, o problema quadrático pode ser resolvido eficientemente devido a possuir um problema dual côncavo com gap de dualidade nula, veja Seção 2.3. Usaremos esses resultados para a Abordagem 3 introduzida na Seção 4.2.

Diversas aplicações para o problema PEC podem ser encontradas em processamento de sinais, onde um sinal multidimensional é tratado como um tensor, e a aproximação de *low-rank* é usada para aproximar o sinal, algumas destas aplicações são apresentadas em [2] e nas referências neles contidas [4–10].

O problema PEC, para $n = 3$, é também usado para formular problemas de processamento de sinais de ressonâncias magnéticas em tecidos biológicos. Apresentaremos uma análise sobre essa formulação na Seção 3.4.

O objetivo principal na tese é desenvolver técnicas de relaxação para o problema cúbico restrito à esfera (PC), motivados por introduzir essas técnicas dentro de um algoritmo de *branch-and-bound* e extender os resultados mostrados por Buchheim et al. [1] para a versão cúbica do problema (1.1).

A seguir descreveremos a sequência na qual serão apresentadas nossas metodologias ao longo da tese.

1.1 Sequência da apresentação

Este trabalho está organizado da seguinte forma:

O Capítulo 2 está dividido em 3 seções, na primeira seção apresentamos as notações a serem utilizadas ao longo do capítulo. Na Seção 2.2 introduzimos o problema de otimização polinomial PP(Ω) e realizamos uma revisão das relaxações via RLT e relaxações baseadas em Programação Semidefinida. Apresentamos os principais resultados de tais investigações e alguns exemplos numéricos. Na Seção 2.3 introduzimos o problema de região de confiança e apresentamos alguns resultados e propriedades importantes para o desenvolvimento da tese.

O Capítulo 3 está dividido em 4 seções, na primeira seção apresentamos as notações a serem utilizadas ao longo do capítulo. Na Seção 3.2 introduzimos o problema PEC e aplicamos as metodologias dadas na Seção 2.2 ao problema PEC mostrando sua eficiência através de exemplos numéricos. Na Seção 3.3 apresentamos novas abordagens para encontrar limites inferiores do problema PEC assim como também mostramos resultados numéricos comparando a eficiência de todas as abordagens. Na Seção 3.4 apresentamos uma aplicação do PEC em problemas de processamento de sinais.

O Capítulo 4 está dividido em 3 seções, na primeira seção apresentamos o problema cúbico geral restrito à esfera (PC), e na Seção 4.2 aplicamos nossas novas abordagens para encontrar limites inferiores do problema PC e mostramos resultados numéricos comparando todas as abordagens. Na Seção 4.3 introduzimos uma abordagem usando técnicas de decomposição Lagrangeana.

Finalmente, no Capítulo 5 são apresentadas algumas conclusões sobre os resultados e os trabalhos futuros que poderiam ser desenvolvidos.

Parte dos resultados desta tese foram divulgados no artigo científico [11].

Capítulo 2

Revisão bibliográfica

Neste capítulo faremos uma revisão das metodologias propostas na literatura para encontrar limites inferiores do problema de otimização polinomial.

O capítulo está dividido em 3 seções, na primeira seção apresentamos as notações a serem utilizadas ao longo do capítulo. Na segunda seção mostraremos as duas técnicas de relaxação mais relevantes para um problema polinomial geral, essas duas relaxações estão baseadas na técnica de reformulação e linearização (RLT) e em Programação Semidefinida. Ambas relaxações adicionam novas variáveis ao problema original substituindo os termos não lineares, porém elas diferem na forma em que as restrições sobre essas novas variáveis são colocadas. Na terceira seção apresentaremos uma revisão teórica sobre o problema de região de confiança indefinida. Esta base teórica será necessária para o desenvolvimento das novas abordagens apresentadas nos Capítulos 3 e 4.

2.1 Notações

Ao longo deste capítulo adotaremos as seguintes notações.

1. Dado uma sequência finita de números reais $\{a_j\}_{j=1}^{n} \subset \mathbb{R}$ denotemos por

$$\sum_{j=1}^{n} a_j = a_1 + a_2 + \cdots + a_n$$

$$\prod_{j=1}^{n} a_j = a_1 a_2 \cdots a_n$$

2. Dados $n, p \in \mathbb{N}$ com $p \leq n$. Uma combinação de n elementos tomados p a p é denotado por

$$\left(\begin{array}{c} n \\ p \end{array} \right) := \frac{n\,!}{(n-p)\,!\; p\,!}$$

3. Dados dois conjuntos X e Y, definimos o produto cartesiano (ou produto direto) desses conjuntos por $X \times Y := \{(x_1, x_2) \ : \ x_1 \in X, x_2 \in Y\}$.

4. Dado um conjunto J definimos o conjunto J^d como o produto cartesiano

$$J^d = \underbrace{J \times J \times \cdots \times J}_{d \text{ vezes}}.$$

5. Define-se o fecho de um conjunto $X \subset \mathbb{R}^n$ como o conjunto

$$\bar{X} := \{x \in \mathbb{R}^n \ : \ B(x, \epsilon) \cap X \neq \emptyset, \forall \epsilon > 0\},$$

onde $B(x, \epsilon) := \{x \in \mathbb{R}^n \ : \ \|x\| < \epsilon\}$.

6. Um conjunto $X \subset \mathbb{R}^n$ é denominado um conjunto fechado se e somente se $X = \bar{X}$.

7. Um conjunto $X \subset \mathbb{R}^n$ é denominado limitado se e somente se $\exists M > 0$ tal que $\|x\| \leq M$ para todo $x \in X$.

8. Um conjunto $X \subset \mathbb{R}^n$ é denominado compacto se e somente se X é fechado e limitado.

9. Dada uma matriz A, usamos $A \succeq 0$, $(A \succ 0)$ para denotar que a matriz A é semidefinida positiva (definida positiva).

10. Dadas duas matrizes $n \times n$ dimensional, $A = [a_{ij}]_{n \times n}$ e $B = [b_{ij}]_{n \times n}$, denotamos o produto interno entre A e B como $A \bullet B = \sum_{i,j=1}^{n} a_{ij} b_{ij}$.

11. Uma matriz $A = [a_{ij}]_{n \times n}$ é dita não singular se todos seus autovalores são não nulos.

12. Uma matriz $A = [a_{ij}]_{n \times n}$ é dita singular quando não admite uma inversa.

13. Define-se o piso de um número real x, denotado por $\lfloor x \rfloor$, como o único número inteiro i tal que $i \leq x < i + 1$.

14. Define-se o teto de um número real x, denotado por $\lceil x \rceil$, como o único número inteiro j tal que $j - 1 < x \leq j$.

2.2 O problema de otimização global com polinômios PP(Ω)

Nesta seção estamos interessados em analisar o problema de otimização global com polinômios. O problema, denotado por PP(Ω), consiste em encontrar um mínimo global de uma função objetivo polinomial sujeito a um conjunto de restrições definidas por funções polinomiais, todas definidas em termos de alguma variável de decisão. Não impomos condições de convexidade às funções polinomiais, porém assumimos que a região viável do problema é um conjunto compacto. Uma formulação matemática deste problema pode ser dada por:

$$\begin{array}{lll} \mathrm{PP}(\Omega): & \min & \varphi_0(x) \\ & \text{sujeito a} & \varphi_r(x) \geq \beta_r, \ \ \forall r = 1, \ldots, R_1, \\ & & \varphi_r(x) = \beta_r, \ \ \forall r = R_1 + 1, \ldots, R, \\ & & x \in \Omega, \end{array}$$

em que

$$\Omega = \{x \in \mathbb{R}^n \ : \ l_j \leq x_j \leq u_j, \ l_j, u_j \in \mathbb{R}, \ \forall j \in \{1, \ldots, n\}\}, \ \text{ e}$$

$$\varphi_r(x) = \sum_{t \in T_r} \alpha_{rt} \left[\prod_{j \in J_{rt}} x_j \right],$$

para $r = 0, \ldots, R$.

Aqui, T_r é um conjunto de índices para os termos definidos por $\varphi_r(\cdot)$ e α_{rt} são coeficientes reais para os termos polinomiais $\prod_{j \in J_{rt}} x_j, \ t \in T_r, \ r = 0, 1, \ldots, R$. Note que permite-se uma repetição de índices dentro de cada conjunto J_{rt}. Por exemplo, se $J_{rt} = \{1, 1, 2, 3\}$, então o correspondente termo polinomial é $x_1^2 x_2 x_3$. Se δ é o grau máximo das funções polinomiais definidas em PP(Ω), temos que cada $J_{rt} \in \bar{\mathcal{N}} \equiv \cup_{d=1}^{\delta} \mathcal{N}^d, \ \forall t \in T_{rt}, \ r = 0, 1, \ldots, R$.

Os problemas de programação polinomial generalizam uma ampla classe de problemas de otimização, por exemplo, problemas de programação inteira 0-1, programação linear, programação quadrática, entre outros. Aplicações para esses problemas podem ser encontradas em diversas áreas, sugerimos consultar [12, 13] e suas referências.

O problema PP(Ω) pertence à classe mais geral de problemas de otimização global com restrições, tratados por exemplo, em Horst [14], Horst e Tuy [15]. Uma abordagem bastante utilizada para resolver estes problemas até a otimalidade é a aplicação de algoritmos do tipo *branch-and-bound* e *spatial branch-and-bound*. O nosso trabalho relaciona-se com esta abordagem de solução, e como é bem sabido, a idéia geral destes métodos é decompor o conjunto viável do problema gerando subproblemas relaxados, em princípio menos difíceis, ao resolver esses subproblemas, limites inferiores ou limites superiores são obtidos (para problemas de minimização ou de maximização, respectivamente) e por meio de estratégias eficientes este procedimento é repetido até convergir para a solução ótima do problema original.

Este fato justifica a importância de formular relaxações fortes de cada subproblema que ainda é um problema polinomial. De fato, nos últimos anos relaxações baseadas em programação linear e em programação semidefinida se tornam cada vez mais comuns, em particular a técnica chamada Reformulação-Linearização (*Reformulation Linearization Technique-RLT*) que foi introduzida por Sherali & Adams [16] (veja também [17, 18]). Uma primeira idéia básica aparece em [19], e consiste em (i) multiplicação das restrições originais por uma família de polinômios (geralmente produtos das restrições originais), (ii) linearização em um espaço de maior dimensão e (iii) resolução do programa linear resultante.

A seguir descreveremos as relaxações lineares mencionadas acima.

2.2.1 Relaxação RLT para o problema PP(Ω)

Em [17] uma técnica de reformulação e linearização (RLT) é descrita para construir relaxações lineares e usá-las dentro de um algoritmo *branch-and-bound* para resolver PP(Ω).

Para gerar a correspondente relaxação RLT, denotada por PL(Ω), geramos restrições *bound-factor* usando os produtos dos fatores, $(x_j - l_j) \geq 0$ e $(u_j - x_j) \geq 0$, $l_j, u_j \in \mathbb{R}\ \forall j \in \mathcal{N} = \{1, 2, \ldots, n\}$, tomando δ vezes, como segue

$$\prod_{j \in J_1} (x_j - l_j) \prod_{j \in J_2} (u_j - x_j) \geq 0, \quad \forall (J_1 \cup J_2) \in \mathcal{N}^\delta \tag{2.1}$$

com $|J_1 \cup J_2| = \delta$. De [20] temos que o número de distintas restrições do tipo (2.1) é dado por

$$\begin{pmatrix} 2n + \delta - 1 \\ \delta \end{pmatrix}. \tag{2.2}$$

Depois de incluir as restrições (2.1) no problema PP(Ω), usamos a seguinte substituição para linearizar o problema resultante:

$$X_J = \prod_{j \in J} x_j, \quad \forall J \in \bigcup_{d=2}^{\delta} \mathcal{N}^d. \tag{2.3}$$

Aqui os índices em J são assumidos em ordem não decrescente, e assumimos também que $X_{\{j\}} \equiv x_j$, $\forall j \in \mathcal{N}$ e $X_\emptyset \equiv 1$. De [20] o número de variáveis X adicionais é dado por:

$$\begin{pmatrix} n+\delta \\ \delta \end{pmatrix} - n - 1. \tag{2.4}$$

Denotemos por $[\cdot]_L$ a versão linearizada de qualquer expressão $[\cdot]$ sob a substitução (2.3). Assim, PL(Ω) é definido como:

$$\begin{array}{lll} \mathrm{PL}(\Omega): & \min & [\varphi_0(x)]_L \\ & \text{s. a} & [\varphi_r(x)]_L \geq \beta_r, \ \ \forall r = 1, \ldots, R_1, \\ & & [\varphi_r(x)]_L = \beta_r, \ \ \forall r = R_1+1, \ldots, R, \\ & & \left[\prod_{j \in J_1} (x_j - l_j) \prod_{j \in J_2} (u_j - x_j) \right]_L \geq 0, \ \ \forall (J_1 \cup J_2) \in \mathcal{N}^\delta \\ & & x \in \Omega. \end{array}$$

A seguir, o Lema 2.1 (demonstrado em [17]) verifica que efetivamente PL(Ω) é uma relaxação do PP(Ω), e dá uma caracterização importante do PL(Ω).

Lema 2.1 *(Lema 1, [17]) Seja $v(\cdot)$ o valor ótimo do correspondente problema $(\cdot)$. Então, $v(PL(\Omega)) \leq v(PP(\Omega))$. Além disso, se a solução ótima (x^*, X^*) obtida para $PL(\Omega)$ satisfaz a equação (2.3) para todo $J \in \cup_{r=0}^{R} \cup_{t \in T_r} \{J_{rt}\}$, então x^* é uma solução ótima do problema $PP(\Omega)$.*

Observação 2.1 *Relaxações via RLT de problemas de programação quadrática têm sido exaustivamente investigadas com resultados eficientes, veja por exemplo [21, 22] e referências deles. Por isto, uma idéia natural para abordar problemas de programação polinomial é transformar o problema $PP(\Omega)$ num problema equivalente de programação quadrática. Essa transformação pode ser realizada de várias formas diferentes, como ilustraremos no exemplo a seguir.*

Exemplo 2.1 *Considere o seguinte problema de programação polinomial sem restrições.*

$$\min 2x_1^4 - 4x_1^3x_2 + 5x_1^2x_2 - 10x_1x_2^2 + 8x_2^4 \tag{2.5}$$

Um problema quadrático equivalente ao problema (2.5) é dado por:

$$\begin{array}{ll} \min & 2x_3^2 - 4x_2x_4 + 5x_2x_3 - 10x_1x_5 + 8x_5^2 \\ s.\ a & x_3 = x_1^2 \\ & x_4 = x_1x_3 \\ & x_5 = x_2^2 \end{array}$$

onde introduzimos 3 variáveis adicionais x_3, x_4 e x_5. Note que, essa não é a única forma, a seguir mostramos outro problema quadrático equivalente a (2.5):

$$
\begin{array}{ll}
\min & 2x_3 - 4x_4x_5 + 5x_2x_5 - 10x_1x_6 + 8x_6^2 \\
s.\ a & x_3 = x_5^2 \\
& x_4 = x_1x_2 \\
& x_5 = x_1^2 \\
& x_6 = x_2^2
\end{array}
$$

A seguir descrevemos uma segunda metodologia usando desigualdades RLT, a qual define uma relaxação mais apertada para o problema PP(Ω). Ambas metodologias serão analizadas por meio de exemplos numéricos na Seção 3.2.3.

2.2.2 Relaxação RLT usando restrições *bound-grid-factor*

Sherali e Dalkiran em [23] introduziram desigualdades válidas, via RLT, para PP(Ω) adicionando restrições que denominaram *bound-grid-factor* a qual descreveremos a seguir. Suponha que temos pontos de grade $\bar{x}_{jg},\ g = 1, \ldots, G_j$, para cada $j \in \mathcal{N}$, tal que

$$l_j < \bar{x}_{j1} < \bar{x}_{j2} < \cdots < \bar{x}_{jG_j} < u_j,\ \ \forall j \in \mathcal{N},$$

e colocamos os pontos da grade uniformemente dentro de cada intervalo como segue:

$$\bar{x}_{jg} = l_j + g\left(\frac{u_j - l_j}{G_j + 1}\right),\ \ \text{para}\ \ g = 1, \ldots, G_j,\ \forall j \in \mathcal{N}.$$

Seja $\mathcal{L}$ a coleção de multiconjuntos de até $\lfloor \delta/2 \rfloor$ combinações de índices (j, g) (com possíveis repetições). Para cada $J_3 \in \mathcal{L}$, seja J_3^* o multiconjunto correspondente é dado por

$$J_3^* \equiv \cup_{(j,g)\in J_3}\{j, j\}.$$

Então, como uma generalização de restrições RLT mostradas na Seção 2.2.1, se define os combinados *bound-grid-factors*:

$$\{(x_j - l_j), (u_j - x_j), (x_j - \bar{x}_{jg})^2; g = 1, \ldots, G_j\},\ \forall j \in \mathcal{N},$$

e consequentemente, se gera as seguintes restrições *bound-grid-factors* de ordem δ:

$$\left[\prod_{j\in J_1}(x_j - l_j)\prod_{j\in J_2}(u_j - x_j)\prod_{(j,g)\in J_3}(x_j - \bar{x}_{jg})^2\right]_L \geq 0,\ \forall\{J_1, J_2, J_3\} : (J_1 \cup J_2 \cup J_3^*) \in \mathcal{N}^\delta. \tag{2.6}$$

Assumindo que temos q pontos de grade para cada variável $j \in \mathcal{N}$, e denotando $\delta' = \lfloor \delta/2 \rfloor$, o número total de restrições em (2.6) é dado por:

$$\sum_{i=0}^{\delta'} \begin{pmatrix} qn+i-1 \\ i \end{pmatrix} \begin{pmatrix} 2n+\delta-2i-1 \\ \delta-2i \end{pmatrix}. \tag{2.7}$$

Notemos que o número de restrições aqui é significativamente maior que o número de restrições dado por (2.1), porém essas restrições adicionais levam a melhores limitantes do valor ótimo do PP(Ω) o qual será analisado no Capítulo 3.

Substituindo (2.6) em vez de (2.1) obtemos a seguinte relaxação do PP(Ω).

$$
\begin{aligned}
\mathrm{PL}_g(\Omega): \quad \min \quad & [\varphi_0(x)]_L \\
\text{s. a} \quad & [\varphi_r(x)]_L \geq \beta_r, \quad \forall r = 1, \ldots, R_1, \\
& [\varphi_r(x)]_L = \beta_r, \quad \forall r = R_1+1, \ldots, R, \\
& \left[\prod_{j \in J_1} (x_j - l_j) \prod_{j \in J_2} (u_j - x_j) \prod_{(j,g) \in J_3} (x_j - \bar{x}_{jg})^2 \right]_L \geq 0, \\
& \qquad \forall \{J_1, J_2, J_3\} : (J_1 \cup J_2 \cup J_3^*) \in \mathcal{N}^\delta \\
& x \in \Omega.
\end{aligned}
$$

Exemplos numéricos

Nesta seção apresentamos exemplos numéricos com o fim de comparar a eficiência das relaxações PL($\cdot$) e PL$_g$($\cdot$). Implementamos as formulações usando a linguagem de programação algébrica AMPL com o solver CPLEX para resolvê-los.

Exemplo 2.2 (Problema tomado de [20]) *Consideremos o seguinte problema polinomial:*

$$
\begin{aligned}
\textit{(PC1)}: \quad \min \quad & x_1^2 x_2/2 + (x_1^2 + x_2^2)/10 + (x_3^2 + x_4^2)/3 \\
\textit{s. a} \quad & -5x_1 + 2x_2 \leq -10, \\
& x_1 - 6x_2 \leq -15, \\
& -5x_3 + 2x_4 \leq -10, \\
& x_3 - 6x_4 \leq -15, \\
& 0.5 \leq x_j \leq 5, \quad j = 1, 2, 3, 4.
\end{aligned}
$$

O valor ótimo do problema $(PC1)$ *é* 24.15 *com* $x_1^* = x_3^* = 3.21$ *e* $x_2^* = x_4^* = 3.04$. *Para o problema acima a formulação PL(PC1) gera* 120 *restrições do tipo RLT com valor ótimo* 11.83. *Por outro lado, a formulação dado por PL$_g$(PC1) gera* 152 *restrições do tipo RLT, porém o seu valor ótimo é* 15.02, *o qual é um melhor limite inferior para o valor ótimo do problema (PC1).*

Exemplo 2.3 (Adaptado do problema 86, [24]) *Consideremos o seguinte problema polinomial:*

$$\begin{aligned} (PC2): \quad \min \quad & \sum_{j=1}^{5} e_j x_j + \sum_{i=1}^{5}\sum_{j=1}^{5} c_{ij} x_i x_j + \sum_{j=1}^{5} d_j x_j^3 \\ s.\ a \quad & \sum_{j=1}^{5} a_{sj} x_j \geq b_s, \quad s = 1, \ldots, 10, \\ & 0 \leq x_j \leq 0.5 \ \ j = 1, 2, 3, 4, 5. \end{aligned}$$

onde

$$e = [-15, -27, -36, -18, -12],$$

$$d = [4, 8, 10, 6, 2],$$

$$b = [-40, -2, -0.25, -4, -4, -1, -40, -60, 5, 1],$$

$$c = \begin{bmatrix} 30 & -20 & -10 & 32 & -10 \\ -20 & 39 & -6 & -31 & 32 \\ -10 & -6 & 10 & -6 & -10 \\ 32 & -31 & -6 & 39 & -20 \\ -10 & 32 & -10 & -20 & 30 \end{bmatrix},$$

$$a = \begin{bmatrix} -16 & 2 & 0 & 1 & 0 \\ 0 & -2 & 0 & 4 & 2 \\ -3.5 & 0 & 2 & 0 & 0 \\ 0 & -2 & 0 & -4 & -1 \\ 0 & -9 & -2 & 1 & -2.8 \\ 2 & 0 & -4 & 0 & 0 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -2 & -3 & -2 & -1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

O valor ótimo do problema (PC2) é -32.35 *com* $x_1^* = x_2^* = 0.3$ *e* $x_3^* = x_4^* = 0.4$, $x_5^* = 0.2$.

Para o problema acima a formulação PL(PC2) gera 220 *restrições do tipo RLT com valor ótimo* -42.62. *Por outro lado, a formulação dado por* PL_g*(PC2) gera* 270 *restrições do tipo RLT, porém o seu valor ótimo é* -38.38*, o qual é um melhor limite inferior para o valor ótimo do problema (PC2).*

Exemplo 2.4 (Adaptado do problema 117, [24]) *Consideremos o seguinte* *problema polinomial:*

$$
\begin{aligned}
\textit{(PC3)}: \quad \min \qquad & -\sum_{s=1}^{10} b_s x_s + \sum_{i=1}^{5}\sum_{j=1}^{5} c_{ij} x_{10+i} x_{10+j} + 2\sum_{j=1}^{5} d_j x_{10+j}^3 \\
s.\ a \quad & 2\sum_{i=1}^{5} c_{ij} x_{10+i} + 3 d_j x_{10+j}^2 + e_j - \sum_{s=1}^{10} a_{sj} x_s \geq 0, \quad j = 1, \ldots, 5, \\
& 0 \leq x_1 \leq 0.5, \\
& 0 \leq x_2 \leq 0.5, \\
& 0 \leq x_3 \leq 5.5, \\
& 0 \leq x_4 \leq 0.5, \\
& 0 \leq x_5 \leq 3.5, \\
& 0 \leq x_6 \leq 12, \\
& 0 \leq x_q \leq 0.5 \quad q = 7, \ldots, 15.
\end{aligned}
$$

onde

$$e = [-15, -27, -36, -18, -12],$$

$$d = [4, 8, 10, 6, 2],$$

$$b = [-40, -2, -0.25, -4, -4, -1, -40, -60, 5, 1],$$

$$c = \begin{bmatrix} 30 & -20 & -10 & 32 & -10 \\ -20 & 39 & -6 & -31 & 32 \\ -10 & -6 & 10 & -6 & -10 \\ 32 & -31 & -6 & 39 & -20 \\ -10 & 32 & -10 & -20 & 30 \end{bmatrix},$$

$$a = \begin{bmatrix} -16 & 2 & 0 & 1 & 0 \\ 0 & -2 & 0 & 4 & 2 \\ -3.5 & 0 & 2 & 0 & 0 \\ 0 & -2 & 0 & -4 & -1 \\ 0 & -9 & -2 & 1 & -2.8 \\ 2 & 0 & -4 & 0 & 0 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -2 & -3 & -2 & -1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

O valor ótimo do problema $(PC3)$ *é* 32.34.

Para o problema acima a formulação PL(PC3) gera $4\,960$ *restrições do tipo RLT com valor ótimo* 20.13. *Por outro lado, a formulação dado por* PL_g*(PC3) gera* $5\,410$ *restrições do tipo RLT, porém o seu valor ótimo é* 25.40, *o qual é um melhor limite* *inferior para o valor ótimo do problema (PC3).*

A seguir descreveremos a segunda metodologia de relaxação para problemas polinomiais.

2.2.3 Relaxação semidefinida para o problema PP(Ω)

A seguinte metodologia foi introduzida por Laserre [25, 26] e surgiu de maneira natural como uma generalização do procedimento para o caso quadrático. Por esse motivo, para ilustrar os conceitos, inicialmente consideremos o problema de programação polinomial quadrático da forma:

$$\begin{array}{lll} \text{PPQ}: & \min & x^\top Q_0 x + a_0^\top x \\ & \text{s. a} & x^\top Q_i x + a_i^\top x \leq b_i, \ \ i \in \mathcal{I} \\ & & x^\top Q_i x + a_i^\top x = b_i, \ \ i \in \mathcal{E} \\ & & l \leq x \leq u. \end{array}$$

em que $x, l, u \in \mathbb{R}^n$, $\mathcal{I} \cup \mathcal{E} = \{1, \ldots, m\}$ e cada Q_i é uma matriz simétrica.

Note que, dada uma matriz Q podemos escrever

$$x^\top Q x = \sum_{i,j=1}^{n} x_i x_j Q_{ij} = Tr(Q \cdot x x^\top) = Q \bullet X$$

com $X = x x^\top$.

Assim, PPQ é equivalente a

$$\begin{array}{ll} \min & Q_0 \bullet X + a_0^\top x \\ \text{s. a} & Q_i \bullet X + a_i^\top x \leq b_i, \ \ i \in \mathcal{I} \\ & Q_i \bullet X + a_i^\top x = b_i, \ \ i \in \mathcal{E} \\ & l \leq x \leq u, \quad X = x x^\top \end{array}$$

Portanto, uma relaxação semidefinida positiva do PPQ pode ser representada por

$$\begin{array}{lll} \text{SDP(PPQ)}: & \min & Q_0 \bullet X + a_0^\top x \\ & \text{s. a} & Q_i \bullet X + a_i^\top x \leq b_i, \ \ i \in \mathcal{I} \\ & & Q_i \bullet X + a_i^\top x = b_i, \ \ i \in \mathcal{E} \\ & & l \leq x \leq u, \quad X - x x^\top \succeq 0. \end{array}$$

Além disso, devido ao bem conhecido resultado (Teorema Complemento de Schur)

$$X - x x^\top \succeq 0 \Leftrightarrow \begin{bmatrix} 1 & x^\top \\ x & X \end{bmatrix} \succeq 0,$$

e considerando

$$x_{(1)} = \begin{bmatrix} 1 \\ x \end{bmatrix} \text{ e definindo a matriz } M_1 \equiv [x_{(1)}x_{(1)}^\top] = \begin{bmatrix} 1 & x^\top \\ x & X \end{bmatrix}, \tag{2.8}$$

com $M_1 \succeq 0$, obtemos a formulação para o SDP(PPQ) é dada por:

$$\begin{array}{ll} \min & Q_0 \bullet X + a_0^\top x \\ \text{s. a} & Q_i \bullet X + a_i^\top x \leq b_i, \quad i \in \mathcal{I} \\ & Q_i \bullet X + a_i^\top x = b_i, \quad i \in \mathcal{E} \\ & l \leq x \leq u, \quad M_1 \succeq 0. \end{array}$$

Agora, vejamos uma extensão da idéia dada em (2.8) para o caso polinomial geral. Para isso, defina-se $x_{(m)}$ como o vetor que tem todas as componentes de $x_{(1)}$ com todos os termos quadráticos envolvendo as variáveis em x, depois todos os termos cúbicos, e assim por diante, até todos os possíveis monômios de ordem m. Isto é,

$$x_{(m)} := [1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_1x_n, x_2^2, \ldots, x_n^2, \ldots, x_1^m, \ldots, x_n^m]^\top.$$

Podemos definir a matriz momento linearizada

$$M_m \equiv [x_{(m)}x_{(m)}^\top]_L.$$

onde $[\cdot]_L$ representa a linearização da expresão $[\cdot]$ sob a substitução:

$$x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n} \to y_\alpha \in \mathbb{R}, \alpha := (\alpha_1, \alpha_2, \ldots, \alpha_n) \tag{2.9}$$

Por exemplo, em $\mathbb{R}^2$ temos que $x_{(2)}$ é definido por

$$x_{(2)} := [1, x_1, x_2, x_1^2, x_1x_2, x_2^2]^\top,$$

e assim podemos obter a matriz $x_{(2)}x_{(2)}^\top$ como

$$x_{(2)}x_{(2)}^\top = \begin{bmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \\ x_1 & x_1^2 & x_1x_2 & x_1^3 & x_1^2x_2 & x_1x_2^2 \\ x_2 & x_1x_2 & x_2^2 & x_1^2x_2 & x_1x_2^2 & x_2^3 \\ x_1^2 & x_1^3 & x_1^2x_2 & x_1^4 & x_1^3x_2 & x_1^2x_2^2 \\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^3x_2 & x_1^2x_2^2 & x_1x_2^3 \\ x_2^2 & x_1x_2^2 & x_2^3 & x_1^2x_2^2 & x_1x_2^3 & x_2^4 \end{bmatrix}$$

e a sua linearização, usando 2.9, é

$$M_2 = \begin{bmatrix} 1 & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{bmatrix}.$$

Então, para o caso de um problema de programação polinomial do tipo:

$$\min\{p(x) : x \in \mathbb{R}^n\},$$

em que $p(x)$ é um polinômio de grau $2m$ definido por

$$p(x) := \sum_{\alpha} p_\alpha x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \quad \text{e} \quad \sum_i \alpha_i \leq 2m \tag{2.10}$$

com p_α sendo o vetor de coeficientes dos monomios associados.

Em [25], para cada $N \geq m$ considera-se a relaxação

$$\begin{aligned} \min \quad & [p(x)]_L := \sum_\alpha p_\alpha y_\alpha \\ \text{s. a} \quad & M_N := [x_{(N)} x_{(N)}^\top]_L \succeq 0, \\ & [\theta(x) x_{(N-1)} x_{(N-1)}^\top]_L \succeq 0. \end{aligned} \tag{2.11}$$

onde $[\cdot]_L$ representa a linearização da expresão $[\cdot]$ sob a substitução (2.9), $m \geq \lceil \delta/2 \rceil$ e $\theta(x) := a^2 - \|x\|^2$, com $a > 0$, sendo o raio da esfera que contém uma solução ótima.

O teorema seguinte (demonstrado em [25]) mostra que a relaxação (2.11) é assintoticamente exata [25], isto é, quando $N \to \infty$, o valor ótimo da relaxação se aproxima do valor ótimo do problema polinomial associado.

Teorema 2.1 (Teorema 3.4 em [25]) *Dado $p(x) : \mathbb{R}^n \to \mathbb{R}$ um polinômio de grau $2m$ definido como em (2.10) e com mínimo global $p^* = \min_{x \in \mathbb{R}^n} p(x)$ tal que $\|x^*\| \leq a$ para algum $a > 0$ em algum minimizador global x^*. Denotando o problema (2.11) por Q_a^N. Então temos que*

$$\lim_{N \to \infty} \inf Q_a^N = p^*.$$

Por outro lado, considere o problema de programação polinomial com restrições

$$\min\{p(x) : \varphi_r(x) \geq \beta_r, \ \forall r = 1, \dots, R\}$$

onde $p(x)$ é um polinômio de grau m e para cada $r = 1, \dots, R$ o grau de $\varphi_r(x)$ é δ_r, uma relaxação similar a (2.11) foi gerada como segue.

Para $N \geq \max\{\lceil \delta_r/2 \rceil, \ r = 1, 2, \ldots n\}$ e $N \geq \lceil m/2 \rceil$

$$\begin{array}{ll} \min & [p(x)]_L \\ \text{s. a} & M_N \succeq 0, \\ & [(\varphi_r(x) - \beta_r) x_{(N - \lceil \delta_r/2 \rceil)} x_{(N - \lceil \delta_r/2 \rceil)}^\top]_L \succeq 0, \ \forall r = 1, \ldots, R \end{array} \tag{2.12}$$

O seguinte resultado (demonstrado em [25]) garante que a relaxação (2.12) é assintoticamente exata quando $N \to \infty$, [25, 26], sob certas condições impostas aos polinômios que definem cada região viável.

Teorema 2.2 (Teorema 4.2 em [25]) *Dado $p(x) : \mathbb{R}^n \to \mathbb{R}$ um polinômio de grau m definido como em (2.10), $K := \{\varphi_r(x) \geq 0, r = 1, \ldots, R\}$ um conjunto não vazio, compacto e seja $p_K^* := \min_{x \in K} p(x)$. Denotando ao problema (2.12) por Q_K^N. Então se verifica*

$$\lim_{N \to \infty} \inf Q_K^N = p_K^*$$

Observação 2.2 *Podemos perceber que as relaxações via programação semidefinida, formulações (2.11) e (2.12), possuem uma importante desvantagem computacional, visto que o tamanho das matrizes e número de variáveis são enormes ainda para pequenos valores de n e m.*

Tamanho das matrizes:

$$\binom{n+m}{m} \times \binom{n+m}{m}$$

Número de variáveis:

$$\binom{n+2m}{2m}.$$

2.2.4 Relaxação semidefinida vs relaxação RLT no caso quadrático

Nesta seção mostraremos alguns resultados obtidos por Anstreicher [21], em programas quadráticos, comparando as duas metodologias: relaxação via Programação Semidefinida e relaxação via RLT.

Considere o problema quadrático

$$(\text{QCQP}) \quad \begin{array}{lrcl} max & x^\top Q_0 x + a_0^\top x & & \\ s.a. & x^\top Q_i x + a_i^\top x & \leq & b_i, \quad i \in \mathcal{I} \\ & x^\top Q_i x + a_i^\top x & = & b_i, \quad i \in \mathcal{E} \\ & l \leq x \leq u, & & \end{array}$$

em que $x \in \mathbb{R}^n$ e $\mathcal{I} \cup \mathcal{E} = \{1, \dots, m\}$. Assumimos que $l, u \in \mathbb{R}^n$ e as matrices Q_i são todas simétricas.

Temos, da Seção 2.2.3, que $x^\top Q_0 x = \sum_{i,j=1}^n x_i x_j (Q_0)_{ij} = Tr(Q_0 \cdot xx^\top) = Q_0 \bullet X$ onde $X = xx^\top$. Então a relaxação Semidefinida de QCQP pode ser escrita como

$$
(\mathrm{SDP}) \qquad \begin{array}{lllll} max & Q_0 \bullet X + a_0^\top x & & & \\ s.a. & Q_i \bullet X + a_i^\top x & \leq & b_i, & i \in \mathcal{I} \\ & Q_i \bullet X + a_i^\top x & = & b_i, & i \in \mathcal{E} \\ & l \leq x \leq u, & & & \\ & X - xx^\top \succeq 0. & & & \end{array}
$$

Observação 2.3 *Dadas as formulações acima podemos mencionar algumas importantes características.*

1. *Quando o problema original QCQP é convexo ($Q_0 \preceq 0, Q_i \succeq 0$ para $i \in \mathcal{I}$, e $Q_i = 0$ para $i \in \mathcal{E}$), SDP é equivalente a QCQP.*

   De fato, é claro que toda solução ótima de SDP é solução ótima de QCQP. Por isto, é suficiente mostrar que se $\bar{x}$ é uma solução ótima do QCQP então $\bar{X} = \bar{x}\bar{x}^\top$ é solução ótima do SDP.

   Sejam x, X pontos viáveis para o problema SDP deste modo, $X - xx^\top \succeq 0$ e $Q_0 \preceq 0$ então

   $$Q_0 \bullet X \leq Q_0 \bullet xx^\top,$$

   e portanto,

   $$\begin{array}{rcl} Q_0 \bullet X + a_0^\top x & \leq & Q_0 \bullet xx^\top + a_0^\top x \\ & = & x^\top Q_0 x + a_0^\top x \\ & \leq & \bar{x}^\top Q_0 \bar{x} + a_0^\top \bar{x} \\ & = & Q_0 \bullet \bar{x}\bar{x}^\top + a_0^\top \bar{x} \\ & = & Q_0 \bullet \bar{X} + a_0^\top \bar{x}, \end{array}$$

   onde a segunda desigualdade é válida porque $\bar{x}$ é uma solução ótima de QCQP.

2. *Note que se QCQP é não convexo então SDP poderia ser não limitado, devido que não se satisfaz $Q_0 \bullet X \leq Q_0 \bullet xx^\top$. Mas isto poderia ser remediado adicionando limites superiores às componentes da diagonal de X. De fato, note que $X - xx^\top \succeq 0$,*

   $$\begin{bmatrix} X_{11} - x_1^2 & * & \cdots & * \\ * & X_{22} - x_2^2 & & * \\ \vdots & & \ddots & \vdots \\ * & * & \cdots & X_{nn} - x_n^2 \end{bmatrix} \succeq 0.$$

$$X_{11} - x_1^2 + X_{22} - x_2^2 + \cdots + X_{nn} - x_n^2 \geq 0$$

$$X_{11} + X_{22} + \cdots + X_{nn} \geq x_1^2 + \cdots + x_n^2 \geq 0$$

Então para limitar o problema SDP basta limitar superiormente as componentes diagonais de X.

Por outro lado, usando a metodologia RLT, para as variáveis x_i e x_j temos as restrições $x_i - l_i \geq 0$, $u_i - x_i \geq 0$, $x_j - l_j \geq 0$, $u_j - x_j \geq 0$. Multiplicando cada restrição envolvendo x_i por restrições envolvendo x_j e substituindo o termo $x_i x_j$ com uma nova variável X_{ij}, obtemos

$$\begin{aligned} X_{ij} - l_j x_j - l_j x_i &\geq -l_i l_j, \\ X_{ij} - u_i x_j - u_j x_i &\geq -u_i u_j, \\ X_{ij} - l_i x_j - u_j x_i &\leq -l_i u_j, \\ X_{ij} - l_j x_i - u_i x_j &\leq -l_j u_i, \end{aligned}$$

$i, j = 1, \ldots, n.$

As ultimas duas restrições são iguais se impomos $X = X^\top$. A relaxação resultante de QCQP pode ser escrito como:

$$(RLT) \quad \begin{array}{lllll} max & Q_0 \bullet X + a_0^\top x & & & \\ s.a. & Q_i \bullet X + a_i^\top x & \leq & b_i, & i \in \mathcal{I} \\ & Q_i \bullet X + a_i^\top x & = & b_i, & i \in \mathcal{E} \\ & X - lx^\top - xl^\top & \geq & -ll^\top, & \\ & X - ux^\top - xu^\top & \geq & -uu^\top, & \\ & X - lx^\top - xu^\top & \leq & -lu^\top, & \\ & X = X^\top & & & \end{array}$$

A seguir examinaremos o efeito de adicionar a condição $X \succeq xx^\top$ à relaxação RLT do QCQP. Sem perda de generalidade, assumamos $l = 0, u = e$ e considere duas variáveis x_1, x_2 tal que $0 \leq x_1 \leq x_2 \leq 0.5$. Sob essas hipóteses as restrições RLT em X_{11}, X_{22}, X_{12} tornam-se

$$0 \leq \quad X_{11} \quad \leq x_1 \tag{2.13}$$

$$0 \leq \quad X_{22} \quad \leq x_2 \tag{2.14}$$

$$0 \leq \quad X_{12} \quad \leq x_1. \tag{2.15}$$

Agora considere a condição $X - xx^\top \succeq 0$ a qual é equivalente a

$$\begin{pmatrix} X_{11} - x_1^2 & X_{12} - x_1x_2 \\ X_{12} - x_2x_1 & X_{22} - x_2^2 \end{pmatrix} \succeq 0, \tag{2.16}$$

e equivalente às restrições

$$X_{11} - x_1^2 \geq 0$$

$$(X_{11} - x_1^2)(X_{22} - x_2^2) - (X_{12} - x_1x_2)^2 \geq 0$$

Portanto, (2.16) é equivalente a

$$\begin{aligned} X_{11} &\geq x_1^2, & (2.17)\\ X_{22} &\geq x_2^2, & (2.18)\\ X_{12} &\leq x_1x_2 + \sqrt{(X_{11} - x_1^2)(X_{22} - x_2^2)}, & (2.19)\\ X_{12} &\geq x_1x_2 - \sqrt{(X_{11} - x_1^2)(X_{22} - x_2^2)}. & (2.20) \end{aligned}$$

Nosso seguinte objetivo é comparar, fixando valores de x_1, x_2, as regiões viáveis tridimensionais para (X_{11}, X_{22}, X_{12}) correspondentes à condição RLT (2.13)-(2.15) antes e depois de ser adicionada a condição SDP (2.17)-(2.20).

Observação 2.4 *1. Assumindo $x_1 > 0$, $x_2 > 0$ e a condição SDP, é claro que os limites inferiores de X_{ii} melhoram de $X_{ii} \geq 0$ para $X_{ii} \geq x_i^2$ e os limites superiores de X_{ii} não mudam.*

Portanto, para quaisquer valores X_{ii} satisfazendo $x_i^2 \leq X_{ii} \leq x_i$, $i = 1, 2$ os valores de X_{12}, para que (X_{11}, X_{22}, X_{12}) seja viável, devem satisfazer as restantes restrições de SDP e RLT.

2. Note que para $x_1 = x_2 = 0.5$, as restrições (2.19) e (2.20) dominam à restrição RLT (2.15). De fato,

$$X_{12} \leq \frac{1}{4} + \sqrt{(X_{11} - \frac{1}{4})(X_{22} - \frac{1}{4})} \leq \frac{1}{2} = x_1$$

$$X_{12} \geq \frac{1}{4} - \sqrt{(X_{11} - \frac{1}{4})(X_{22} - \frac{1}{4})} \geq 0.$$

Para valores mais gerais de x_1, x_2 essa situação é mais complexa.

No seguinte Teorema descrevemos o volume tridimensional da região SDP + RLT para todos os valores relevantes de x_1, x_2.

Teorema 2.3 *Suponha que $l = 0, u = e, 0 < x_1 \leq x_2 \leq 0.5$. Então o volume de $\{(X_{11}, X_{22}, X_{12})\}$ viável em RLT (2.13)-(2.15) é $x_1^2 x_2$, e o volume de $\{(X_{11}, X_{22}, X_{12})\}$ viável via RLT + SDP (2.13)-(2.15) e (2.17)-(2.20) é*

$$x_1^2 x_2(1-x_2) - \frac{1}{9}x_1^3(6x_2^2 - 6x_2 + 5) + \frac{1}{3}x_1^3((1-x_2)^3 - x_2^3)ln\left(\frac{1-x_2}{x_2}\right)$$

$$-\frac{1}{3}x_1^3((1-x_2)^3 + x_2^3)ln\left(\frac{1-x_1}{x_1}\right).$$

Demonstração. É claro que para encontrar o volume (V) de $\{(X_{11}, X_{22}, X_{12})\}$ viável em (2.13)-(2.15) devemos resolver a integral $\int_0^{x_1}\int_0^{x_2} x_1 dX_{22} dX_{11}$, então

$$V := \int_0^{x_1}\int_0^{x_2} x_1 dX_{22} dX_{11} = \int_0^{x_1} x_1 x_2 dX_{11} = x_1^2 x_2.$$

Por outro lado, para calcular o volume de $\{(X_{11}, X_{22}, X_{12})\}$ viável via RLT + SDP é conveniente considerar as regiões com $X_{12} \leq x_1x_2$ e $X_{12} \geq x_1x_2$ separadamente.

Primeiro consideremos o caso $X_{12} \leq x_1x_2$.

Temos que

$$x_1^2 \leq X_{ii} \leq x_i \;\; i = 1, 2. \tag{2.21}$$

Vejamos os seguintes dois subcasos:

1. Note que o limite inferior de (2.15) dominará o limite inferior de (2.20) (ou seja $x_1x_2 - \sqrt{(X_{11} - x_1^2)(X_{22} - x_2^2)} \leq 0$) se o somente se

$$\frac{x_2^2 X_{11}}{X_{11} - x_1^2} \leq X_{22}, \tag{2.22}$$

   e como $X_{12} \leq x_2$ então

$$\frac{x_2^2 X_{11}}{X_{11} - x_1^2} \leq x_2$$

   isto implica que

$$\frac{x_1^2}{1 - x_2} < X_{11}. \tag{2.23}$$

   Portanto, de (2.21), (2.22) e (2.23) obtemos

$$\begin{array}{ccccc} \frac{x_1^2}{1-x_2} & < & X_{11} & \leq & x_1 \\ \frac{x_2^2 X_{11}}{X_{11}-x_1^2} & \leq & X_{22} & \leq & x_2 \\ 0 & \leq & X_{12} & \leq & x_1x_2. \end{array} \tag{2.24}$$

2. O limite inferior de (2.20) dominará o limite inferior de (2.15) (ou seja $x_1x_2 -$

$\sqrt{(X_{11} - x_1^2)(X_{22} - x_2^2)} \geq 0)$ se o somente se

$$X_{22} \leq \frac{x_2^2 X_{11}}{X_{11} - x_1^2}, \tag{2.25}$$

Note que $X_{22} \leq x_2$ também é satisfeito. Então considere primeiro $\frac{x_2^2 X_{11}}{X_{11} - x_1^2} \leq x_2$ o qual é equivalente a

$$\frac{x_1^2}{1 - x_2} < X_{11}.$$

Portanto, usando a desigualdade acima, junto com (2.21) e (2.25) obtemos

$$\begin{array}{rcccl} \frac{x_1^2}{1-x_2} & < & X_{11} & \leq & x_1 \\ x_2^2 & \leq & X_{22} & \leq & \frac{x_2^2 X_{11}}{X_{11}-x_1^2} \\ x_1x_2 - \sqrt{(X_{11} - x_1^2)(X_{22} - x_2^2)} & \leq & X_{12} & \leq & x_1x_2. \end{array} \tag{2.26}$$

No caso $X_{22} \leq x_2 \leq \frac{x_2^2 X_{11}}{X_{11} - x_1^2}$ obtemos

$$X_{11} \leq \frac{x_1^2}{1 - x_2} \quad (x_2 \neq 0),$$

deste modo, junto com (2.21) obtemos

$$\begin{array}{rcccl} x_1^2 & < & X_{11} & \leq & \frac{x_1^2}{1-x_2} \\ x_2^2 & \leq & X_{22} & \leq & x_2 \\ x_1x_2 - \sqrt{(X_{11} - x_1^2)(X_{22} - x_2^2)} & \leq & X_{12} & \leq & x_1x_2. \end{array} \tag{2.27}$$

Note também que desde a hipótese $0 < x_1 \leq 0.5$ e $0 < x_1 \leq 0.5$ temos,

$$x_1 + x_2 \leq 1 \text{ o que implica que } \frac{x_1^2}{1 - x_2} \leq x_1.$$

Deste modo, $x_1^2 \leq X_{11} \leq \frac{x_1^2}{1-x_2} \leq x_1$. Portanto, de (2.24), (2.26) e (2.27), o volume de $\{(X_{11}, X_{22}, X_{12})\}$ viável via RLT + SDP considerando a região $X_{12} \leq x_1x_2$ é dado por

$$\begin{aligned} V_1 \; := & \int_{\frac{x_1^2}{1-x_2}}^{x_1} \int_{\frac{x_2^2 X_{11}}{X_{11}-x_1^2}}^{x_2} x_1x_2 dX_{22} dX_{11} \\ & + \int_{\frac{x_1^2}{1-x_2}}^{x_1} \int_{x_2^2}^{\frac{x_2^2 X_{11}}{X_{11}-x_1^2}} (X_{11} - x_1^2)^{1/2} (X_{22} - x_2^2)^{1/2} dX_{22} dX_{11} \\ & + \int_{x_1^2}^{\frac{x_1^2}{1-x_2}} \int_{x_2^2}^{x_2} (X_{11} - x_1^2)^{1/2} (X_{22} - x_2^2)^{1/2} dX_{22} dX_{11} \end{aligned}$$

Integrando cada parcela obtemos,

$$V_1 = x_1^2 x_2^2 (1 - x_1 - x_2) + \frac{4}{9} x_1^3 x_2^3 - \frac{1}{3} x_1^3 x_2^3 \ln \left(\frac{1 - x_1}{x_1} \frac{1 - x_2}{x_2} \right) \tag{2.28}$$

Analogamente podemos obter o volume V_2 no caso $X_{12} \geq x_1 x_2$.

Neste caso, temos

$$x_1 x_2 \leq X_{12} \leq \sqrt{(X_{11} - x_1^2)(X_{22} - x_2^2)} + x_1 x_2$$

$$x_i \leq X_{ii} \leq x_i, \quad i = 1, 2.$$

$$0 \leq X_{12} \leq x_1$$

Primeiro notemos que $X_{12} \leq x_1$ domina $X_{12} \leq \sqrt{(X_{11} - x_1^2)(X_{22} - x_2^2)} + x_1 x_2$ se e somente se $X_{11} \geq x_1^2 + \frac{x_1^2(1-x_2)^2}{X_{22} - x_2^2}$.

Além disso, como $x_1 \geq X_{11} \geq x_1^2 + \frac{x_1^2(1-x_2)^2}{X_{22} - x_2^2}$ então

$$X_{22} \geq x_1 + \frac{(x_1 - x_2)^2}{1 - x_1}.$$

Assim, obtemos

$$\begin{array}{rcccl} x_1^2 + \frac{x_1^2(1-x_2)^2}{X_{22}-x_2^2} & < & X_{11} & \leq & x_1 \\ x_1 + \frac{(x_1-x_2)^2}{1-x_1} & \leq & X_{22} & \leq & x_2 \\ x_1 x_2 & \leq & X_{12} & \leq & x_1. \end{array} \tag{2.29}$$

Por outro lado, se $X_{12} \leq x_1$ é dominado por $X_{12} \leq \sqrt{(X_{11} - x_1^2)(X_{22} - x_2^2)} + x_1 x_2$ então

$$\begin{array}{rcccl} x_1^2 & < & X_{11} & \leq & x_1^2 + \frac{x_1^2(1-x_2)^2}{X_{22}-x_2^2} \\ x_1 + \frac{(x_1-x_2)^2}{1-x_1} & \leq & X_{22} & \leq & x_2 \\ x_1 x_2 & \leq & X_{12} & \leq & \sqrt{(X_{11} - x_1^2)(X_{22} - x_2^2)} + x_1 x_2. \end{array} \tag{2.30}$$

Além disso, quando $x_1 \leq \frac{x_1^2(1-x_2)^2}{X_{22}-x_2^2}$ então $X_{22} \leq x_1 + \frac{(x_1-x_2)^2}{1-x_1}$ e portanto

$$\begin{array}{rcccl} x_1^2 & < & X_{11} & \leq & x_1 \\ x_2^2 & \leq & X_{22} & \leq & x_1 + \frac{(x_1-x_2)^2}{1-x_1} \\ x_1 x_2 & \leq & X_{12} & \leq & \sqrt{(X_{11} - x_1^2)(X_{22} - x_2^2)} + x_1 x_2. \end{array} \tag{2.31}$$

Assim, de (2.29)-(2.31) o volume V_2 no caso $X_{12} \geq x_1 x_2$ é dado por

$$
\begin{aligned}
V_2 \quad := & \int_{x_1+\frac{(x_1-x_2)^2}{1-x_1}}^{x_2} \int_{x_1^2+\frac{x_1^2(1-x_2)^2}{X_{22}-x_2^2}}^{x_1} (x_1 - x_1x_2)dX_{11}dX_{22} \\
+ & \int_{x_1+\frac{(x_1-x_2)^2}{1-x_1}}^{x_2} \int_{x_1^2}^{x_1^2+\frac{x_1^2(1-x_2)^2}{X_{22}-x_2^2}} (X_{11} - x_1^2)^{1/2}(X_{22} - x_2^2)^{1/2}dX_{11}dX_{22} \\
+ & \int_{x_2^2}^{x_1+\frac{(x_1-x_2)^2}{1-x_1}} \int_{x_1^2}^{x_1} (X_{11} - x_1^2)^{1/2}(X_{22} - x_2^2)^{1/2}dX_{11}dX_{22}
\end{aligned}
$$

Integrando cada parcela obtemos

$$
V_2 = x_1^2(1-x_2)^2(x_2-x_1) + \frac{4}{9}x_1^3(1-x_2)^3 - \frac{1}{3}x_1^3(1-x_2)^3 ln\left(\frac{1-x_1}{x_1}\frac{x_2}{1-x_2}\right) \quad (2.32)
$$

Finalmente, adicionando (2.28) com (2.32) obtemos o resultado. ■

Comparando os volumes das regiões viáveis de RLT e SDP + RLT para valores fixos de x_1 e x_2, o resultado anterior pode ser usado para derivar o volume 5-dimensional das regiões correspondentes.

Teorema 2.4 *Suponha que* $0 \leq x_i \leq 1, i = 1, 2$. *Então o volume de* $\{(x_1, x_2, X_{11}, X_{22}, X_{12})\}$ *que são viáveis para as restrições RLT é* $1/60$ *e o volume de* $\{(x_1, x_2, X_{11}, X_{22}, X_{12})\}$ *que são viáveis via RLT+SDP é* $1/240$.

Demonstração. Do Teorema 2.3, o volume de $\{(x_1, x_2, X_{11}, X_{22}, X_{12})\}$ viável para restrições RLT com $0 \leq x_1 \leq x_2 \leq 0.5$ é

$$
\int_0^{1/2} \int_0^{x_2} x_1^2 x_2 dx_1 dx_2,
$$

o qual dá como resultado $\frac{1}{480}$. Para encontrar o volume de $\{(x_1, x_2, X_{11}, X_{22}, X_{12})\}$ que também satisfaça a condição SDP (2.16), requer calcular

$$
\int_0^{1/2} \int_0^{x_2} v(x_1, x_2) dx_1 dx_2
$$

onde $v(x_1, x_2)$ é o volume tridimensional dado no Teorema 2.3. Usando Maple para resolver a integral, o resultado dá $\frac{1}{1920}$.

Além disso, a região $0 \leq x_1 \leq x_2 \leq 0.5$ representa $\frac{1}{8}$ da região viável $0 \leq x_1 \leq 1$, $i = 1, 2$, Portanto, obtemos o resultado. ■

Temos como resultado então que adicionando-se a condição SDP à relaxação RLT remove-se exatamente 75% da região viável determinada por duas das variáveis originais.

Observação 2.5 *Em resumo podemos dizer que teoricamente a relaxação via programação semidefinida (SDP) dá melhores resultados; e relaxações via desigualdades RLT dá bons resultados na prática devido a disponibilidade e eficiência dos softwares para programação linear que estão capacitados para resolver problemas de grande porte. Os resultados computacionais mostrados em [21], indicam que o melhor gap de otimalidade foi dado quando ambas relaxações foram usadas em conjunto, embora tenha ocasionado um retardo de 200 segundos de CPU em media, comparado a 1 segundo de CPU obtido quando as relaxações RLT e SDP são usadas em separado. Para mais detalhes veja [21].*

Observação 2.6 *Para o caso de polinômios de maior grau, Laserre em [26] mostrou que existe um quadro natural comum para a comparação da relaxação SDP e a relaxação via linearização (no sentido RLT), no campo da teoria dos momentos e a teoria de representação de polinômios não negativos definidos em um conjunto semi-algébrico compacto. O foco da análise foi teórico concluindo que ambas relaxações obtém bons limites inferiores (ou até uma solução ótima), porém na prática dependendo da ordem, a relaxação SDP não é uma boa escolha para problemas de tamanho pequeno e médio. Veja [26] para mais detalhes.*

Na seguinte seção faremos uma revisão da base teórica usada nas abordagens que serão apresentadas no Capítulos 3 e 4.

2.3 O problema de região de confiança indefinido

Em uma importante classe de algoritmos de minimização chamados métodos de região de confiança (veja, for exemplo, [27–29]), o cálculo dos passos iterativos requer a minimização de uma função quadrática restrita a uma esfera. Este subproblema é da forma

$$\begin{array}{lll} \text{(P)} & \min & \mu(y) = y^\top B y - 2\psi^\top y \\ & \text{s. a} & \beta \leq y^\top C y \leq \alpha, \\ & & y \in \mathbb{R}^n \end{array}$$

onde B e C são matrizes simétricas indefinidas, e $-\infty \leq \beta \leq \alpha \leq \infty$. Os problemas de região de confiança são importantes tanto em otimização irrestrita como restrita. A teoria, algoritmos e aplicações têm sido descritos em vários artigos e textos, veja por exemplo [30–33]. A seguir apresentaremos alguns resultados mostrados por R. Stern e H. Wolkowicz em [34], os quais usaremos em nossas abordagens nos Capítulos 3 e 4.

2.3.1 Condições de otimalidade

O seguinte resultado (demonstrado em [34]) é de grande importância porque caracteriza uma solução ótima global do problema (P).

Teorema 2.5 *(Teorema 2.1 em [34]) Seja y um ponto viável para o problema (P). Então y é o mínimo global para (P) se e somente se existe um multiplicador de Lagrange $\lambda \in \mathbb{R}$ tal que*

$$(B - \lambda C)y = \psi \tag{2.33}$$

$$B - \lambda C \succeq 0 \tag{2.34}$$

e

$$\lambda(\beta - y^\top Cy) \geq 0 \geq \lambda(y^\top Cy - \alpha). \tag{2.35}$$

Além disso, se

$$B - \lambda C \succ 0, \tag{2.36}$$

então y é o único ponto de minimização. Ademais, suponha que a seguinte restrição de qualificação satisfaz

$$Cy = 0 \;\; \textit{implica que} \;\; \beta < 0 < \alpha. \tag{2.37}$$

Então y resolve (P) se e somente se as condições (2.33)-(2.35) são satisfeitas, para algum $\lambda \in \mathbb{R}$.

Como consequência do teorema acima, obtemos a seguinte condição à matriz $B - \lambda C$ (demonstrado em [34]) para que (P) possua um ponto de minimização.

Corolário 2.1 *(Corolário 2.2 em [34]) Suponha que C é não singular e* $\max\{|\alpha|, |\beta|\} > 0\}$*. Se y resolve (P), então*

$$\exists \hat{\lambda} \in \mathbb{R} \;\; \textit{tal que} \;\; B - \hat{\lambda} C \succeq 0. \tag{2.38}$$

Antes de mostrar o principal resultado de existência, distinguimos o resultado de (2.38) em dois casos, o caso regular e o caso irregular.

1. No caso regular consideramos que (P) é viável e tem-se que $\exists \hat{\lambda} \in \mathbb{R}$ tal que $B - \hat{\lambda} C \succ 0$.

2. No caso irregular consideramos que (P) é viável porém não existe $\lambda \in \mathbb{R}$ tal que $B - \lambda C \succ 0$.

Agora podemos descrever o resultado principal (demonstrado em [34]) de existência de um ponto minimizador para o problema (P).

Teorema 2.6 *(Teorema 2.4 em [34]) Considere o problema (P) com C não singular.*

1. *Se (P) possui um ponto minimizador e* $\max\{|\alpha|, |\beta|\} > 0$*, então a condição (2.38) é satisfeita.*

2. *Reciprocamente, assumindo que (P) é viável, (2.38) se verifica, e ambos α e β são finitos. Então temos os seguintes casos.*

   (a) *regular. (P) possui um ponto minimizador.*

   (b) *irregular. (P) possui um ponto minimizador se e somente se (2.33)-(2.35) são consistentes, neste caso y é o ponto minimizador com $\hat{\lambda}$ multiplicador de Lagrange.*

2.3.2 Dualidade

Nesta seção derivamos o problema dual para (P) o qual é um problema de maximização côncava. Isto ilustra que (P) é um problema convexo implicito e mostra porque o mínimo global pode ser caracterizado e encontrado. De fato, será também mostrado que o Lagrangeano dual possui gap de dualidade nulo (demonstrado em [34]).

Teorema 2.7 *(Teorema 5.1 em [34]) Suponha que y^* resolve (P) com valor ótimo $\mu^* = \mu(y^*)$ e multiplicador de Lagrange λ^*. Seja*

$$L(y, \nu, \omega) = \mu(y) + \nu(\alpha - y^\top Cy) + \omega(y \top Cy - \beta) \tag{2.39}$$

a função Lagrangeana para (P). Seja

$$\phi(\nu, \omega) = \inf_y L(y, \nu, \omega) \tag{2.40}$$

o funcional dual de Lagrange, e seja

$$h(\nu, \omega) = \nu\alpha - \omega\beta - \psi^\top (B - \nu C + \omega C)^{-1}\psi \tag{2.41}$$

o funcional dual quadrático. Então o valor ótimo de (P) satisfaz

$$\mu^* = \max_{\nu \leq 0, \omega \leq 0} \phi(\nu, \omega). \tag{2.42}$$

Se o caso regular vale, então adicionalmente temos

$$\mu^* = \sup_{\substack{B - \nu C + \omega C \succ 0 \\ \nu \leq 0, \omega \leq 0}} h(\nu, \omega). \tag{2.43}$$

Além disso, o máximo em (2.42) é obtido para

$$\nu^* = -(-\lambda^*)_+ \ \text{ e } \ \omega^* = -(\lambda^*)_+, \tag{2.44}$$

onde

$$(\lambda)_+ = \begin{cases} \lambda & \text{se } \lambda \geq 0, \\ 0 & \text{caso contrário} \end{cases} \tag{2.45}$$

Note que no resultado acima (2.42) fornece o programa dual Lagrangeano e (2.43) fornece o programa quadrático dual. Ambos problemas duais não possuim gap de dualidade e ambos problemas maximizam uma função côncava sobre um conjunto convexo isto indica que os problemas de região de confiança são implicitamente programas convexos.

Podemos obter o resultado (demonstrado em [34]) de dualidade com somente um multiplicador.

Corolário 2.2 *(Corolário 5.2 em [34]) Suponha que estamos no caso regular e y resolve (P) com λ como multiplicador de Lagrange. Define-se*

$$\bar{h}(\lambda) = -(-\lambda)_+\alpha + (\lambda)_+\beta - \psi^\top (B - \lambda C)^+ \psi. \tag{2.46}$$

Então o valor ótimo de (P) satisfaz

$$\mu^* = \sup_{B - \lambda C \succ 0} \bar{h}(\lambda). \tag{2.47}$$

Além disso, no caso regular, o máximo é atingido, enquanto no caso irregular é atingido para λ com $B - \lambda C$ semidefinido positiva e possivelmente singular.

Capítulo 3

O problema um-esférico cúbico

Neste capítulo introduzimos o problema de otimização um-esférico cúbico e apresentamos diferentes técnicas para construir relaxações tratáveis do problema mencionado. O capítulo está dividido como segue: Na primeira seção introduziremos as notações que usaremos ao longo do capítulo. Na Seção 3.2 definiremos o problema um-esférico cúbico ao qual denotaremos por PEC e adaptaremos ao caso cúbico as metodologias apresentadas nas Seções 2.2.1 e 2.2.2. Na Seção 3.3 apresentaremos nossas propostas para encontrar limites inferiores do PEC e mostraremos resultados numéricos comparando todas as abordagens. Finalmente, na Seção 3.4 mostraremos uma aplicação da formulação PEC dentro de um problema de processamento de sinais biológicos.

3.1 Notações

Uma matriz $\mathcal{A} = [a_{ijk}]$ de dimensão $(n \times n \times n)$ é denominada tensor. Ao longo da tese um tensor $\mathcal{A}$ é considerado simétrico no sentido que seu elemento a_{ijk} é invariante sob qualquer permutação do seus indices (i, j, k), isto é, se por exemplo $a_{143} = 7$ então

$$a_{134} = a_{314} = a_{341} = a_{413} = a_{431} = 7.$$

Denotemos por $S^3(\mathbb{R}^n)$ o espaço de tensores simétricos reais de terceira ordem $(n \times n \times n)$-dimensional. Dado um tensor simétrico $\mathcal{A} \in S^3(\mathbb{R}^n)$ para algum $\ell \in \{1, \ldots, n\}$, definimos A_ℓ como a matriz simétrica composta por elementos $a_{\ell jk}$ de $\mathcal{A}$, para todo $j, k = 1, \ldots, n$. Além disso, denotamos $\tilde{A}_\ell$ pela submatriz de A_ℓ onde a ℓ-ésima linha e coluna são eliminadas.

Por exemplo, para $n = 3$, temos

$$\mathcal{A} := \begin{pmatrix} a_{111} & a_{112} & a_{113} & a_{211} & a_{212} & a_{213} & a_{311} & a_{312} & a_{313} \\ a_{121} & a_{122} & a_{123} & a_{221} & a_{222} & a_{223} & a_{321} & a_{322} & a_{323} \\ a_{131} & a_{132} & a_{133} & a_{231} & a_{232} & a_{233} & a_{331} & a_{332} & a_{333} \end{pmatrix} = (A_1, A_2, A_3)$$

com

$$A_1 = \begin{pmatrix} a_{111} & a_{112} & a_{113} \\ a_{121} & a_{122} & a_{123} \\ a_{131} & a_{132} & a_{133} \end{pmatrix}, A_2 = \begin{pmatrix} a_{211} & a_{212} & a_{213} \\ a_{221} & a_{222} & a_{223} \\ a_{231} & a_{232} & a_{233} \end{pmatrix}, A_3 = \begin{pmatrix} a_{311} & a_{312} & a_{313} \\ a_{321} & a_{322} & a_{323} \\ a_{331} & a_{332} & a_{333} \end{pmatrix},$$

e,

$$\tilde{A}_1 := \begin{pmatrix} a_{122} & a_{123} \\ a_{132} & a_{133} \end{pmatrix}, \tilde{A}_2 := \begin{pmatrix} a_{211} & a_{213} \\ a_{231} & a_{233} \end{pmatrix}, \tilde{A}_3 := \begin{pmatrix} a_{311} & a_{312} \\ a_{321} & a_{322} \end{pmatrix}.$$

Ao longo do capítulo, para uma matriz real simétrica X, denotamos por $\lambda_{\min}(X)$ e $\lambda_{\max}(X)$ o menor autovalor e o maior autovalor de X respectivamente.

Dado um vetor $x \in \mathbb{R}^n$ e $\ell \in \{1, \ldots, n\}$, definimos o vetor $x_{\hat{\ell}} \in \mathbb{R}^{n-1}$ como $x_{\hat{\ell}} := (x_1, \ldots, x_{\ell-1}, x_{\ell+1}, \ldots, x_n)$, isto é, o vetor x onde a ℓ-ésima componente é eliminada.

3.2 O problema PEC

O problema de otimização um-esférico cúbico tem a seguinte formulação matemática:

$$\begin{aligned} \text{PEC}: \quad \min \quad & f(x) := \mathcal{A}x^3 = \sum_{i,j,k=1}^{n} a_{ijk} x_i x_j x_k \\ \text{s. a} \quad & \|x\| = 1 \\ & x \in \mathbb{R}^n, \end{aligned} \tag{3.1}$$

onde $n \geq 2$ e $\mathcal{A} \in S^3(\mathbb{R}^n)$.

Para simplificar as notações, consideremos ao somatório $\sum_{i,j,k=1}^{n}$ como sendo

$$\sum_{i,j,k=1}^{n} a_{ijk} x_i x_j x_k = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} a_{ijk} x_i x_j x_k.$$

Como foi mostrado por Zhang et al. [2], usando um resultado de Nesterov [3], o problema (PEC) é NP-difícil.

Diversas aplicações para o problema de otimização um-esférico cúbico podem ser encontradas em processamento de sinais [2] , e em referências nele contidas [4–10], onde um sinal multidimensional é tratado como um tensor a ser aproximado de forma *low-rank*.

O problema PEC, para $n = 3$, é também usado para formular problemas em sinais de ressonâncias magnéticas em tecidos biológicos. Apresentaremos uma análise sobre essa formulação na Seção 3.4.

A seguir analisaremos as relaxações RLT, apresentadas no Capítulo 2, aplicadas ao problema PEC.

3.2.1 Limite inferior do PEC via RLT

Notemos que o conjunto de restrição do PEC pode ser relaxado considerando variáveis no intervalo $[l_j, u_j]$ com $l_j = -1$ e $u_j = 1$, $\forall j \in \mathcal{N} = \{1, 2, \ldots, n\}$. De fato, para gerar a correspondente relaxação RLT, geramos restrições usando os produtos dos fatores $(x_j - l_j) \geq 0$ e $(u_j - x_j) \geq 0$, $\forall j \in \mathcal{N}$ tomado no máximo até $\delta = 3$ termos, como segue

$$\prod_{j \in J_1} (x_j - l_j) \prod_{j \in J_2} (u_j - x_j) \geq 0, \quad \forall (J_1 \cup J_2) \in \mathcal{N}^3, \tag{3.2}$$

onde $|J_1 \cup J_3| = 3$. Usando (2.2), o número de restrições do tipo (3.2) é dado por:

$$\binom{2n+2}{3}.$$

Incluimos as restrições (3.2) ao problema PEC e substituimos

$$X_J = \prod_{j \in J} x_j, \quad \forall J \in \mathcal{N}^2 \cup \mathcal{N}^3, \tag{3.3}$$

onde os índices em J são assumidos em ordem não decrescente.

Desta forma a relaxação linear básica via RLT do PEC, é formulada como:

$$\begin{aligned}
\text{PL(PEC)}: \quad \min \quad & \sum_{i,j,k=1}^{n} a_{ijk} X_{ijk} \\
\text{s. a} \quad & \sum_{i=1}^{n} X_{ii} = 1, \\
& \left[\prod_{j \in J_1} (x_j - l_j) \prod_{j \in J_2} (u_j - x_j) \right]_L \geq 0, \ \forall (J_1 \cup J_2) \in \mathcal{N}^3, \\
& x \in \Omega := \{x \in \mathbb{R}^n : l_i \leq x_i \leq u_i, \ \ i = 1, \ldots, n\}.
\end{aligned}$$

onde $[\cdot]_L$ denota a linearização de $[\cdot]$ sob a substitução (3.3). Além disso, note que as variáveis de decisão deste problema estão definidas por:

$$(x_1, \ldots, x_n, X_{11}, \ldots, X_{1n}, X_{22}, \ldots, X_{2n}, \ldots, X_{n-1\,n-1}, X_{n-1\,n}, X_{nn}, X_{111}, \ldots, X_{nnn}).$$

Usando a fórmula dada em (2.4) obtemos o número total de variáveis como:

$$\begin{pmatrix} n+3 \\ 3 \end{pmatrix} - 1.$$

Ilustremos por meio do seguinte exemplo as definições dadas acima.

Exemplo 3.1 *Considere o problema um-esférico cúbico*

$$(P_1) \quad \begin{array}{ll} \min & \displaystyle\sum_{i,j,k=1}^{2} a_{ijk}x_ix_jx_k = a_{111}x_1^3 + 3a_{112}x_1^2x_2 + 3a_{122}x_1x_2^2 + a_{222}x_2^3 \\ s.\ a & x_1^2 + x_2^2 = 1 \end{array}$$

Considerando $l_j = -1$ e $u_j = 1$ para todo $j \in \mathcal{N} := \{1, 2\}$ para gerar a correspondente relaxação RLT, geramos restrições usando os produtos dos fatores $(x_j - l_j) \geq 0$ e $(u_j - x_j) \geq 0$, $\forall j \in \mathcal{N}$. Portanto, as restrições dadas em (3.2) são:

$$\begin{array}{rcl} (x_1 - l_1)(u_1 - x_1)(u_2 - x_2) & \geq & 0, \quad onde\ J_1 = \{1\},\ J_2 = \{1, 2\} \\ (x_2 - l_2)(u_1 - x_1)(u_2 - x_2) & \geq & 0, \quad onde\ J_1 = \{2\},\ J_2 = \{1, 2\} \\ (x_1 - l_1)(x_2 - l_2)(u_1 - x_1) & \geq & 0, \quad onde\ J_1 = \{1, 2\},\ J_2 = \{1\} \\ (x_1 - l_1)(x_2 - l_2)(u_2 - x_2) & \geq & 0, \quad onde\ J_1 = \{1, 2\},\ J_2 = \{2\} \end{array}$$

$$\begin{array}{rcl} (x_1 - l_1)^2(u_1 - x_1) & \geq & 0, \quad onde\ J_1 = \{1, 1\},\ J_2 = \{1\} \\ (x_1 - l_1)^2(u_2 - x_2) & \geq & 0, \quad onde\ J_1 = \{1, 1\},\ J_2 = \{2\} \\ (x_2 - l_2)^2(u_1 - x_1) & \geq & 0, \quad onde\ J_1 = \{2, 2\},\ J_2 = \{1\} \\ (x_2 - l_2)^2(u_2 - x_2) & \geq & 0, \quad onde\ J_1 = \{2, 2\},\ J_2 = \{2\} \\ (x_1 - l_1)^2(x_2 - l_2) & \geq & 0, \quad onde\ J_1 = \{1, 1, 2\},\ J_2 = \emptyset \\ (x_2 - l_2)^2(x_1 - l_1) & \geq & 0, \quad onde\ J_1 = \{1, 2, 2\},\ J_2 = \emptyset \end{array}$$

$$\begin{array}{rcl} (u_1 - x_1)^2(x_1 - l_1) & \geq & 0, \quad onde\ J_1 = \{1\},\ J_2 = \{1, 1\} \\ (u_1 - x_1)^2(x_2 - l_2) & \geq & 0, \quad onde\ J_1 = \{2\},\ J_2 = \{1, 1\} \\ (u_2 - x_2)^2(x_1 - l_1) & \geq & 0, \quad onde\ J_1 = \{1\},\ J_2 = \{2, 2\} \\ (u_2 - x_2)^2(x_2 - l_2) & \geq & 0, \quad onde\ J_1 = \{2\},\ J_2 = \{2, 2\} \\ (u_1 - x_1)^2(u_2 - x_2) & \geq & 0, \quad onde\ J_1 = \emptyset,\ J_2 = \{1, 1, 2\} \\ (u_2 - x_2)^2(u_1 - x_1) & \geq & 0, \quad onde\ J_1 = \emptyset,\ J_2 = \{1, 2, 2\} \end{array}$$

$$\begin{array}{rcl} (x_1 - l_1)^3 & \geq & 0, \quad \textit{onde } J_1 = \{1,1,1\},\ J_2 = \emptyset \\ (x_2 - l_2)^3 & \geq & 0, \quad \textit{onde } J_1 = \{2,2,2\},\ J_2 = \emptyset \\ (u_1 - x_1)^3 & \geq & 0, \quad \textit{onde } J_1 = \emptyset,\ J_2 = \{1,1,1\} \\ (u_2 - x_2)^3 & \geq & 0, \quad \textit{onde } J_1 = \emptyset,\ J_2 = \{2,2,2\}, \end{array}$$

fazendo as respectivas multiplicações e substituindo (veja (3.3))

$$X_{11} := x_1^2, X_{22} := x_2^2, X_{12} := x_1x_2, X_{111} = x_1^3, X_{112} := x_1^2x_2, X_{122} := x_1x_2^2, X_{222} := x_2^3,$$

obtemos a relaxação linear PL(P_1):

$$(PL(P_1)) \quad \begin{array}{ll} \min & \displaystyle\sum_{i,j,k=1}^{2} a_{ijk}X_{ijk} = a_{111}X_{111} + 3a_{112}X_{112} + 3a_{122}X_{122} + a_{222}X_{222} \\ s.\ a & X_{11} + X_{22} = 1 \\ & -x_2 - X_{11} - X_{112} + 1 \geq 0 \\ & -x_1 - X_{22} + X_{122} + 1 \geq 0 \\ & x_2 - X_{11} - X_{112} + 1 \geq 0 \\ & x_1 - X_{22} - X_{122} + 1 \geq 0 \\ & x_1 - X_{11} - X_{111} + 1 \geq 0 \\ & 2x_1 - x_2 + X_{11} - 2X_{12} - X_{112} + 1 \geq 0 \\ & -x_1 + 2x_2 + X_{22} - 2X_{12} - X_{122} + 1 \geq 0 \\ & x_2 - X_{22} - X_{222} + 1 \geq 0 \\ & 2x_1 + x_2 + X_{11} + 2X_{12} + X_{112} + 1 \geq 0 \\ & x_1 + 2x_2 + X_{22} + 2X_{12} + X_{122} + 1 \geq 0 \\ & -x_1 - X_{11} + X_{111} + 1 \geq 0 \\ & -2x_1 + x_2 + X_{11} - 2X_{12} + X_{112} + 1 \geq 0 \\ & x_1 - 2x_2 + X_{22} - 2X_{12} + X_{122} + 1 \geq 0 \\ & -x_2 - X_{22} + X_{222} + 1 \geq 0 \\ & -2x_1 - x_2 + X_{11} + 2X_{12} - X_{112} + 1 \geq 0 \\ & -x_1 - 2x_2 + X_{22} + 2X_{12} - X_{122} + 1 \geq 0 \\ & 3x_1 + 3X_{11} + X_{111} + 1 \geq 0 \\ & 3x_2 + 3X_{22} + X_{222} + 1 \geq 0 \\ & -3x_1 + 3X_{11} - X_{111} + 1 \geq 0 \\ & -3x_2 + 3X_{22} - X_{222} + 1 \geq 0 \\ & -1 \leq x_i \leq 1, \quad i = 1,2. \end{array}$$

Note que o número de restrições RLT, é dado por

$$\left(\begin{array}{c} 2(2)+2 \\ 3 \end{array}\right) = \frac{6\,!}{3\,!\,3\,!} = 20,$$

e o número total de variáveis é

$$\begin{pmatrix} 2+3 \\ 3 \end{pmatrix} - 1 = \frac{5\,!}{3\,!\,2\,!} - 1 = 9.$$

3.2.2 Limite inferior do PEC usando *bound-grid-factor*

Nesta seção formularemos a relaxação linear RLT usando restrições *bound-grid-factor* para o problema PEC. Como na Seção 2.2.2, suponha que temos pontos de grade $\bar{x}_{jg},\ g = 1, \ldots, G_j$, para cada $j \in \mathcal{N}$, tal que

$$l_j < \bar{x}_{j1} < \bar{x}_{j2} < \cdots < \bar{x}_{jG_j} < u_j, \ \ \forall j \in \mathcal{N},$$

e colocamos os pontos da grade uniformemente dentro de cada intervalo como segue:

$$\bar{x}_{jg} = l_j + g\left(\frac{u_j - l_j}{G_j + 1}\right), \ \text{ para } \ g = 1, \ldots, G_j, \ \forall j \in \mathcal{N} \tag{3.4}$$

Então a nova formulação será obtida substituindo (3.2) por

$$\left[\prod_{j \in J_1}(x_j - l_j)\prod_{j \in J_2}(u_j - x_j)\prod_{(j,g) \in J_3}(x_j - \bar{x}_{jg})^2\right]_L \geq 0, \ \forall \{J_1, J_2, J_3\} : (J_1 \cup J_2 \cup J_3^*) \in \mathcal{N}^3 \tag{3.5}$$

dentro da formulação acima.

Isto é, obtemos a seguinte formulação:

$$\begin{aligned}
\mathrm{PL}_g(\mathrm{PEC}):\quad \min \quad & \sum_{i,j,k=1}^{n} a_{ijk} X_{ijk} \\
\text{s. a} \quad & \sum_{i=1}^{n} X_{ii} = 1, \\
& \left[\prod_{j \in J_1}(x_j - l_j)\prod_{j \in J_2}(u_j - x_j)\prod_{(j,g) \in J_3}(x_j - \bar{x}_{jg})^2\right]_L \geq 0, \\
& \qquad\qquad \forall \{J_1, J_2, J_3\} : (J_1 \cup J_2 \cup J_3^*) \in \mathcal{N}^3 \\
& x \in \Omega := \{x \in \mathbb{R}^n : l_i \leq x_i \leq u_i, \ \ i = 1, \ldots, n\}.
\end{aligned}$$

Além disso, como nosso problema é cúbico, de (2.7), temos $q = 1$ ponto de grade para cada variável x_j, $j \in \mathcal{N}$, e $\delta = 3$, assim $\delta' := \lfloor \delta/2 \rfloor = 1$ e portanto o número de restrições em (3.5) é dado por:

$$\begin{pmatrix} n-1 \\ 0 \end{pmatrix}\begin{pmatrix} 2n+2 \\ 3 \end{pmatrix} + \begin{pmatrix} n \\ 1 \end{pmatrix}\begin{pmatrix} 2n \\ 1 \end{pmatrix}.$$

Exemplo 3.2 *Considerando o Exemplo 3.1, como* $l_1 = l_2 = -1$ *e* $u_1 = u_2 = 1$*, notemos que se* $J_3 = \emptyset$ *então (3.5) é exatamente (3.2). Então para formular o problema* $PL_g(P_1)$ *devemos considerar todas as restrições geradas por (3.2) e acrescentar as seguintes restrições de grade:*

$$\begin{array}{rcl}
(x_1+1)(x_1-0)^2 & \geq & 0, \quad \textit{onde } J_1 = \{1\},\ J_2 = \emptyset,\ J_3 = \{(1,1)\} \\
(x_1+1)(x_2-0)^2 & \geq & 0, \quad \textit{onde } J_1 = \{1\},\ J_2 = \emptyset,\ J_3 = \{(2,1)\} \\
(x_2+1)(x_1-0)^2 & \geq & 0, \quad \textit{onde } J_1 = \{2\},\ J_2 = \emptyset,\ J_3 = \{(1,1)\} \\
(x_2+1)(x_2-0)^2 & \geq & 0, \quad \textit{onde } J_1 = \{2\},\ J_2 = \emptyset,\ J_3 = \{(2,1)\} \\
(1-x_1)(x_1-0)^2 & \geq & 0, \quad \textit{onde } J_1 = \emptyset,\ J_2 = \{1\},\ J_3 = \{(1,1)\} \\
(1-x_1)(x_2-0)^2 & \geq & 0, \quad \textit{onde } J_1 = \emptyset,\ J_2 = \{1\},\ J_3 = \{(2,1)\} \\
(1-x_2)(x_1-0)^2 & \geq & 0, \quad \textit{onde } J_1 = \emptyset,\ J_2 = \{2\},\ J_3 = \{(1,1)\} \\
(1-x_2)(x_2-0)^2 & \geq & 0, \quad \textit{onde } J_1 = \emptyset,\ J_2 = \{2\},\ J_3 = \{(2,1)\}
\end{array}$$

onde o ponto de grade $\bar{x}_{j1}$ *, veja (3.4), é dado por:*

$$\bar{x}_{j1} = l_j + 1 \cdot \left(\frac{u_j - l_j}{1+1}\right) = \frac{l_j + u_j}{2} = \frac{-1+1}{2} = 0, \quad \forall j = 1, 2$$

fazendo as respectivas multiplicações e substituindo (veja (3.3))

$$X_{11} := x_1^2, X_{22} := x_2^2, X_{111} = x_1^3, X_{112} := x_1^2 x_2, X_{122} := x_1 x_2^2, X_{222} := x_2^3 \qquad (3.6)$$

obtemos a formulação $PL_g(P_1)$

$$
PL_g(P_1)\quad
\begin{array}{ll}
\min & \displaystyle\sum_{i,j,k=1}^{2} a_{ijk}X_{ijk} = a_{111}X_{111} + 3a_{112}X_{112} + 3a_{122}X_{122} + a_{222}X_{222} \\
s.\ a & X_{11} + X_{22} = 1 \\
& x_2 - X_{11} - X_{112} + 1 \geq 0 \\
& x_1 - X_{22} - X_{122} + 1 \geq 0 \\
& -x_2 - X_{11} - X_{112} + 1 \geq 0 \\
& -x_1 - X_{22} + X_{122} + 1 \geq 0 \\
& 2x_1 + x_2 + X_{11} + 2X_{12} + X_{112} + 1 \geq 0 \\
& x_1 - X_{11} - X_{111} + 1 \geq 0 \\
& 2x_1 - x_2 + X_{11} - 2X_{12} - X_{112} + 1 \geq 0 \\
& x_1 + 2x_2 + X_{22} + 2X_{12} + X_{122} + 1 \geq 0 \\
& -x_1 + 2x_2 + X_{22} - 2X_{12} - X_{122} + 1 \geq 0 \\
& x_2 - X_{22} - X_{222} + 1 \geq 0 \\
& -2x_1 - x_2 + X_{11} + 2X_{12} - X_{112} + 1 \geq 0 \\
& -x_1 - X_{11} + X_{111} + 1 \geq 0 \\
& -2x_1 + x_2 + X_{11} - 2X_{12} + X_{112} + 1 \geq 0 \\
& -x_1 - 2x_2 + X_{22} + 2X_{12} - X_{122} + 1 \geq 0 \\
& x_1 - 2x_2 + X_{22} - 2X_{12} + X_{122} + 1 \geq 0 \\
& -x_2 - X_{22} + X_{222} + 1 \geq 0 \\
& 3x_1 + 3X_{11} + X_{111} + 1 \geq 0 \\
& 3x_2 + 3X_{22} + X_{222} + 1 \geq 0 \\
& -3x_1 + 3X_{11} - X_{111} + 1 \geq 0 \\
& -3x_2 + 3X_{22} - X_{222} + 1 \geq 0 \\
& X_{111} + X_{11} \geq 0 \\
& X_{122} + X_{22} \geq 0 \\
& X_{112} + X_{11} \geq 0 \\
& X_{222} + X_{22} \geq 0 \\
& X_{11} + X_{111} \geq 0 \\
& X_{22} - X122 \geq 0 \\
& X_{11} - X_{112} \geq 0 \\
& X_{22} - X_{222} \geq 0 \\
& -1 \leq x_i \leq 1, \quad i = 1, 2.
\end{array}
$$

onde o número de restrições é

$$
\begin{pmatrix} 2-1 \\ 0 \end{pmatrix}\begin{pmatrix} 4+2 \\ 3 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix}\begin{pmatrix} 2(2) \\ 1 \end{pmatrix} = 20 + 8 = 28.
$$

Note que as variáveis em (3.6) já foram consideradas na formulação PL(P_1), assim o número de variáveis de PL(P_1) e PL$_g$(P_1) é o mesmo.

Na tabela a seguir, mostramos uma comparação entre o número de variáveis e restrições das formulações PL(PEC) e PL$_g$(PEC).

	N. variáveis adicionais	N. restrições adicionais	
n	PL(PEC) / PL$_g$(PEC)	PL(PEC)	PL$_g$(PEC)
3	16	56	74
10	275	1540	1740
50	23 375	171 700	176 700
75	76 076	573 800	585 050
100	176 851	1 353 400	1 373 400

Tabela 3.1: Comparação de tamanhos dos problemas PL(PEC) e PL$_g$(PEC).

Os resultados mostrados na Tabela 3.1 indicam que ambas relaxações PL(PEC) e PL$_g$(PEC) aumentam consideravelmente o tamanho do problema original. Por exemplo, quando tem-se $n = 100$ variáveis no problema cúbico original, os problemas relaxados têm 176 851 variáveis adicionais e 1 353 400, e 1 373 400 restrições adicionais para PL(PEC) e PL$_g$(PEC), respectivamente. Apesar de se aumentar o tamanho do problema uma vantagem de ambas relaxações é que os problemas resultantes são lineares, e na literatura existem diversos softwares que resolvem eficazmente problemas lineares de grande porte.

3.2.3 Experimentos numéricos

Nesta seção descreveremos testes computacionais usando as metodologias apresentadas nas seções 3.2.1 e 3.2.2. Os testes foram executados em um computador com as seguintes características: Intel(R) Xeon(R), com 3 GHz, 12GB de memoria RAM e sistema operacional Ubuntu 14.04. Usamos a linguagem de programação algébrica AMPL com o solver CPLEX para resolver os problemas.
O objetivo dos testes computacionais é comparar os limites inferiores encontrados ao resolver os problemas relaxados PL(PEC) e PL$_g$(PEC).
Note que o grupo de restrições definidas em (3.2) para formular o problema PL(PEC) é dado por:

Para $i < j < k$:

$$\begin{aligned}
[(x_i - l_i)(x_j - l_j)(x_k - l_k)]_L &\geq 0 \\
[(x_i - l_i)(x_j - l_j)(u_k - u_k)]_L &\geq 0 \\
[(u_i - x_i)(u_j - x_j)(x_k - l_k)]_L &\geq 0 \\
[(u_i - x_i)(u_j - x_j)(u_k - x_k)]_L &\geq 0 \\
[(x_i - l_i)(u_j - x_j)(x_k - l_k)]_L &\geq 0 \\
[(u_i - x_i)(x_j - l_j)(x_k - l_k)]_L &\geq 0 \\
[(x_i - l_i)(u_j - x_j)(u_k - x_k)]_L &\geq 0 \\
[(u_i - x_i)(x_j - l_j)(u_k - x_k)]_L &\geq 0,
\end{aligned}$$

para $i = j = k$:

$$\begin{aligned}
[(x_i - l_i)^3]_L &\geq 0 \\
[(x_i - l_i)(u_i - x_i)^2]_L &\geq 0 \\
[(x_i - l_i)^2(u_i - x_i)]_L &\geq 0 \\
[(u_i - x_i)^3]_L &\geq 0,
\end{aligned}$$

para $i = j < k$:

$$\begin{aligned}
[(x_i - l_i)^2(x_k - l_k)]_L &\geq 0 \\
[(x_i - l_i)^2(u_k - u_k)]_L &\geq 0 \\
[(u_i - x_i)^2(x_k - l_k)]_L &\geq 0 \\
[(u_i - x_i)^2(u_k - x_k)]_L &\geq 0 \\
[(u_i - x_i)(x_i - l_i)(x_k - l_k)]_L &\geq 0 \\
[(x_i - l_i)(u_i - x_i)(u_k - x_k)]_L &\geq 0,
\end{aligned}$$

para $i < j = k$:

$$\begin{aligned}
[(x_i - l_i)(x_j - l_j)^2]_L &\geq 0 \\
[(u_i - x_i)(u_j - x_j)^2]_L &\geq 0 \\
[(u_i - x_i)(x_j - l_j)^2]_L &\geq 0 \\
[(x_i - l_i)(u_j - x_j)^2]_L &\geq 0 \\
[(x_i - l_i)(x_j - l_j)(u_j - u_j)]_L &\geq 0 \\
[(u_i - x_i)(u_j - x_j)(x_j - l_j)]_L &\geq 0.
\end{aligned}$$

Além dessas restrições, (veja Seção 3.2.2 equação (3.5)), para formular o problema $\text{PL}_g(\text{PEC})$ temos que adicionar as seguintes restrições de grade:
Para $i \leq j$:

$$\begin{aligned} [(x_i - l_i)(x_j - \bar{x}_{j1})^2]_L &\geq 0 \\ [(u_i - x_i)(x_j - \bar{x}_{j1})^2]_L &\geq 0, \end{aligned}$$

para $i < j$:

$$\begin{aligned} [(x_j - l_j)(x_i - \bar{x}_{j1})^2]_L &\geq 0 \\ [(u_j - x_j)(x_i - \bar{x}_{j1})^2]_L &\geq 0, \end{aligned}$$

sendo o ponto de grade $\bar{x}_{j1}$ dado por:

$$\bar{x}_{j1} = l_j + 1 \cdot \left(\frac{u_j - l_j}{1+1} \right) = \frac{l_j + u_j}{2}, \ \ \forall j \in \mathcal{N}.$$

A seguir apresentamos alguns resultados numéricos ao aplicar as duas relaxações lineares PL(PEC) e $\text{PL}_g(\text{PEC})$ do problema PEC. Nas Tabelas 3.2-3.5 apresentamos os resultados de instâncias com diferentes densidades da matriz tensor $\mathcal{A}$. Nós referimos por densidade de uma matriz tensor simplesmente como a quantidade de elementos diferentes de zero que a matriz tensor contém. A primeira coluna indica o porcentagem de elementos não nulas que possui a matriz tensor $\mathcal{A}$, nas outras colunas $v(\cdot)$ denota o valor ótimo do problema $(\cdot)$ e o tempo é mostrado em segundos.

Experimento 3.1 *Considere o problema PEC quando $n = 3$, onde a matriz $\mathcal{A}$ que representa a matriz tensor é composta de elementos zeros e uns, isto é, $a_{ijk} \in \{0,1\}, \ \forall i,j,k = 1,2,3$.*

Densidade $\mathcal{A}$ (%)	v(PL(PEC))	v(PL_g(PEC))	tempo(s)
11	-3.00	-1.00^*	0.009 / 0.010
23	-6.00	-3.25	0.009 / 0.010
35	-9.00	-5.29	0.009 / 0.010
46	-12.00	-6.00	0.009 / 0.011
54	-14.00	-10.29	0.009 / 0.011
100	-24.67	-12.00	0.009 / 0.011

Tabela 3.2: Comparação entre os valores ótimos dos problemas PL(PEC) e $\text{PL}_g(\text{PEC})$, $(n = 3)$.

Experimento 3.2 *Os resultados apresentados na Tabela 3.3, consideram o problema PEC com $n = 5$ e $a_{ijk} \in \{0,1\}, \ \forall i,j,k = 1,2,3,4,5$.*

Densidade $\mathcal{A}$ (%)	v(PL(PEC))	v(PL$_g$(PEC))	tempo(s)
4.0	-5	-1.00^*	0.010 / 0.010
7.2	-17	-9.31	0.011 / 0.012
10.4	-29	-13.00	0.012 / 0.012
13.6	-43	-14.33	0.012 / 0.012
100.0	-125	-73.00	0.011 / 0.013

Tabela 3.3: Comparação entre os valores ótimos dos problemas PL(PEC) e PL$_g$(PEC), ($n = 5$).

Experimento 3.3 *Neste experimento considere o problema PEC com* $n = 10$ *variáveis e as mesmas condições dadas nos experimentos anteriores. Assim, pelas fórmulas dadas no capítulo 2 os problemas relaxados PL(PEC) e PL$_g$(PEC) têm* 285 *variavéis com* 1 541 *e* 1 741 *restrições respectivamente. Se a matriz tensor* $\mathcal{A}$ *é* 100% *densa, ou seja, se* $a_{ijk} = 1$ *para todo* $i, j, k \in \{1, 2, \ldots, 10\}$*, obtemos v(PL(PEC))=* -1000 *e v(PL$_g$(PEC))=* -748*, o qual mostra que a relaxação dada por PL$_g$(PEC) é superior em relação a PL(PEC).*

Note que, quando a matriz tensor $\mathcal{A}$ é muito esparsa, veja * nas Tabelas 3.2 e 3.3, usando a formulação PL$_g$(PEC) o gap de dualidade é zero, ou seja, o limite inferior obtido é igual ao limite superior -1. Dos resultados, também podemos perceber que a medida que a matriz tensor $\mathcal{A}$ é mais densa os limites inferiores se afastam do limite superior conhecido -1.

A seguir, nas Tabelas 3.4 e 3.5 comparamos os resultados obtidos pela formulação PL$_g$(PEC) com as soluções obtidas usando o solver BMIBNB do pacote de programação YALMIP em MATLAB.

Usamos o solver BMIBNB para nossas comparações pois este solver é uma implementação do algoritmo *branch-and-bound* espacial para problemas não convexos baseados em relaxações lineares e isso nós permite realizar uma comparação entre os limites inferiores.

Este solver depende de solvers lineares, quadráticos e de programação semidefinida para resolver os subproblemas relaxados e usa solvers não lineares para calcular limites superiores. O BMIBNB é um solver de otimização global com a desvantagem de não ser eficiente para problemas com mais de 10 variáveis. Para mais detalhes veja [35].

Nas Tabelas 3.4 e 3.5 denotamos:

- $lim\ inf$: representa o valor ótimo do problema raiz no método *Branch-and-bound*, o qual foi obtido usando o solver BMIBNB do YALMIP.

- BMIBNB: representa o valor ótimo do problema original obtido pelo solver BMIBNB.

- gap: representa o gap de otimalidade obtido pelo solver BMIBNB. O critério de parada do algoritmo implementado no solver BMIBNB é o numero máximo de iterações conseguindo os melhores limites superiores e melhores limites inferiores possíveis.

Densidade $\mathcal{A}$ (%)	v(PL$_g$(PEC))	lim inf	BMIBNB	gap(%)
11.1*	$-\mathbf{1.00}^*$	−1.00	−1.00	0.00
22.2	$-\mathbf{3.25}^*$	−3.25	−1.90	0.22
33.3	−5.29	−5.20	−2.33	0.52
44.4	−6.00	−5.87	−2.98	0.49
55.6	$-\mathbf{10.29}^*$	−11.20	−3.19	0.53
100.0	$-\mathbf{12.00}^*$	−13.00	−5.20	0.86

Tabela 3.4: Comparação entre os valores ótimos dos problemas PL$_g$(PEC) e os valores obtidos com BMIBNB, ($n = 3$).

Densidade $\mathcal{A}$ (%)	v(PL$_g$(PEC))	lim inf	BMIBNB	gap(%)
4.0*	$-\mathbf{1.00}^*$	−1.00	−1.00	0.00
13.6	−9.31	−9.00	−2.97	0.95
52.0	−46.23	−46.00	−6.94	20.00
80.8	$-\mathbf{67.00}^*$	−67.00	−9.41	44.40
100.0	$-\mathbf{73.00}^*$	−73.00	−11.18	47.50

Tabela 3.5: Comparação entre os valores ótimos dos problemas PL$_g$(PEC) e os valores obtidos com BMIBNB, ($n = 5$).

Os resultados da segunda coluna das Tabelas 3.4 e 3.5 indicados com o simbolo $*$ mostram que a formulação PL$_g$(PEC), para suas correspondentes instâncias, equipara ou melhora os limites inferiores obtidos pelo solver BMIBNB.

Os resultados na coluna *gap* indicam que para problemas com $n = 3$ variáveis o algoritmo BMIBNB consegue resolver os problemas com um gap razoável, veja Tabela 3.4. Por outro lado, note que pra $n = 5$, Tabela 3.5, quando os tensores são mais densos, por exemplo para tensores 52%, 80.8% e 100% densos o solver BMIBNB não obtém gap bons. Isto é devido à dificuldade que existe para encontrar ótimos globais de problemas não convexos.

Em recentes investigações tem-se introduzido novas técnicas de relaxações, [23, 36, 37], as quais acrescentam outras desigualdades válidas às formulações dos problemas relaxados PL(PEC) e PL$_g$(PEC), por exemplo, desigualdades válidas baseadas em programação semidefinida, e/ou cortes disjuntivos, as quais passarão a ser tema de trabalhos de pesquisa futura.

A seguir descreveremos nossas propostas para obter limites inferiores para o problema PEC.

3.3 Novas abordagens

Diferentemente da seção anterior, nesta seção apresentaremos novas propostas de relaxações para o problema PEC mantendo a região viável $\|x\| = 1$. Apresentaremos 3 diferentes abordagens para obter limites inferiores do problema PEC e faremos comparações por meio de experimentos numéricos.

Nas próximas seções usaremos os seguintes resultados.

Lema 3.1 *Dada uma matriz simétrica A e um número real $a > 0$. Então, tem-se que*

$$a\lambda_{min}(A) = \min_{\substack{\|x\|^2=a \\ x\in\mathbb{R}^n}} x^\top Ax$$

e

$$a\lambda_{max}(A) = \max_{\substack{\|x\|^2=a \\ x\in\mathbb{R}^n}} x^\top Ax\,.$$

Demonstração. Sendo A uma matriz simétrica, usando o teorema espectral, existe uma matriz P ortogonal ($P^\top = P^{-1}$) que diagonaliza A.

Considerando $\lambda_{min} := \lambda_1, \lambda_2, \ldots, \lambda_{\max} := \lambda_n$ os autovalores da matriz A ordenados em forma não decrescente, obtemos

$$P^\top AP = \Lambda \;\text{ com }\; \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n, \tag{3.7}$$

onde $\Lambda := \mathrm{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.

Note que P sendo ortogonal preserva comprimentos, assim fazendo $y = P^\top x$ temos que $x = Py$ e

$$\min_{\substack{\|x\|^2=a \\ x\in\mathbb{R}^n}} x^\top Ax \equiv \min_{\substack{\|y\|^2=a \\ y\in\mathbb{R}^n}} y^\top \Lambda y. \tag{3.8}$$

$$\max_{\substack{\|x\|^2=a \\ x\in\mathbb{R}^n}} x^\top Ax \equiv \max_{\substack{\|y\|^2=a \\ y\in\mathbb{R}^n}} y^\top \Lambda y. \tag{3.9}$$

devido que $\|x\|^2 = x^\top x = (Py)^\top Py = yP^\top Py = \|y\|^2$.

Sendo $\lambda_1 := \lambda_{\min}$ e $\lambda_n := \lambda_{\max}$ a menor e a maior das entradas da diagonal de Λ, respectivamente e $\|y\|^2 = a$, temos

$$a\lambda_1 = \lambda_1(y_1^2 + y_2^2 + \cdots + y_n^2) \leq \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 = y^\top \Lambda y \tag{3.10}$$

e

$$y^\top \Lambda y = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \leq \lambda_n(y_1^2 + y_2^2 + \cdots + y_n^2) = a\lambda_n \tag{3.11}$$

logo $a\lambda_1$ e $a\lambda_n$ são cotas inferior e superior de $y^\top \Lambda y$, respectivamente.

Por outro lado, para $y = \sqrt{a}e_1$ sendo $e_1 := (1, 0, \ldots, 0, 0) \in \mathbb{R}^n$, obtemos

$$\min_{\substack{\|y\|^2=a \\ y\in\mathbb{R}^n}} y^\top \Lambda y \leq (\sqrt{a}e_1)^\top \Lambda(\sqrt{a}e_1) = a\lambda_1, \tag{3.12}$$

e para $y = \sqrt{a}e_n$ sendo $e_n := (0, 0, \ldots, 0, 1) \in \mathbb{R}^n$, obtemos

$$\max_{\substack{\|y\|^2=a \\ y\in\mathbb{R}^n}} y^\top \Lambda y \geq (\sqrt{a}e_n)^\top \Lambda(\sqrt{a}e_n) = a\lambda_n, \tag{3.13}$$

deste modo, de (3.10) e (3.12) temos

$$a\lambda_1 = \min_{\substack{\|y\|^2=a \\ y\in\mathbb{R}^n}} y^\top \Lambda y$$

e de (3.11) e (3.13) temos

$$a\lambda_n = \max_{\substack{\|y\|^2=a \\ y\in\mathbb{R}^n}} y^\top \Lambda y$$

e portanto de (3.8) e (3.9) obtemos

$$a\lambda_1 = \min_{\substack{\|x\|^2=a \\ x\in\mathbb{R}^n}} x^\top Ax,$$

$$a\lambda_n = \max_{\substack{\|x\|^2=a \\ x\in\mathbb{R}^n}} x^\top Ax,$$

o que demonstra o resultado. ∎

Lema 3.2 *Dados $c \in \mathbb{R}^n$, e um número real $a > 0$ tem-se que*

$$-\sqrt{a}\|c\| = \min_{\substack{\|x\|^2=a \\ x\in\mathbb{R}^n}} c^\top x \tag{3.14}$$

Demonstração. Usando o Teorema 2.5, para $\alpha = \beta = a$, C a matriz identidade, B a matriz nula e $\varphi = \frac{-c}{2}$, obtemos que x^* é uma solução do problema definido em (3.14) se e somente existe um multiplicador de Lagrange λ tal que

$$-\lambda x^* = \frac{-c}{2}, \text{ e } \|x^*\|^2 = a \tag{3.15}$$

com $\lambda \leq 0$.

Assim, como $x^* = \frac{c}{2\lambda}$ e $\|x^*\|^2 = a$, obtemos

$$\|c\|^2 = 4a\lambda^2,$$

como $\lambda < 0$, obtemos

$$\lambda = \frac{-\|c\|}{2\sqrt{a}}. \tag{3.16}$$

Finalmente, de (3.15) e (3.16), obtemos

$$c^\top x^* = (2\lambda x^*)^\top x^* = 2\lambda a = \frac{-2a\|c\|}{2\sqrt{a}} = -\sqrt{a}\|c\|.$$

O que demonstra o resultado. ■.

3.3.1 Limite inferior via autovalores: Abordagem 1

Com objetivo de encontrar diferentes metodologias para obter limites inferiores do problema PEC introduzimos uma primeira abordagem usando os autovalores das matrizes definidas na função objetivo do problema. Esta ideia é uma adaptação da metodologia usada por Zhang, et al. [2].

Usando a notação introduzida na Seção 3.1 note que a função objetivo do problema PEC pode ser escrita como

$$\mathcal{A}x^3 = \sum_{k=1}^{n}(x^T A_k x)x_k. \tag{3.17}$$

Por outro lado, dados n números reais a_k, $k = 1, \dots, n$ têm-se as seguintes propriedades

$$(\sum_{k=1}^{n} a_k)^2 = \sum_{i,j=1}^{n} a_i a_j \tag{3.18}$$

e

$$\sum_{i,j=1}^{n} a_i a_j \leq n \sum_{i=1}^{n} a_i^2, \tag{3.19}$$

deste modo, considerando $a_k = (x^T A_k x)x_k$, de (3.18) e (3.19) obtemos

$$(\sum_{k=1}^{n}(x^T A_k x)x_k)^2 = \sum_{i,j=1}^{n}(x^T A_i x)(x^T A_j x)x_i x_j \tag{3.20}$$

$$\sum_{i,j=1}^{n}(x^T A_i x)(x^T A_j x)x_i x_j \leq n \sum_{i=1}^{n}(x^T A_i x)^2 x_i^2. \tag{3.21}$$

assim de (3.17), (3.20) e (3.21), obtemos

$$
\begin{aligned}
(\mathcal{A}x^3)^2 &= (\sum_{k=1}^{n}(x^T A_k x)x_k)^2 \\
&= \sum_{i,j=1}^{n}(x^T A_i x)(x^T A_j x)x_i x_j \\
&\leq n\sum_{i=1}^{n}(x^T A_i x)^2 x_i^2.
\end{aligned} \tag{3.22}
$$

Sejam $x \in \mathbb{R}^n$ tal que $\|x\| = 1$, e

$$
\gamma_i = \max\{\lambda^2 \ : \ \lambda \text{ é um autovalor da matriz } A_i\}, \tag{3.23}
$$

para cada $i = 1, 2, \ldots, n$.

De (3.22), (3.23) e o Lema 3.1, obtemos

$$
\begin{aligned}
(\mathcal{A}x^3)^2 &\leq n\sum_{i=1}^{n}\gamma_i x_i^2 \\
&\leq n\gamma^*\sum_{i=1}^{n}x_i^2 \\
&= n\gamma^*
\end{aligned}
$$

onde a segunda desigualdade segue porque definimos $\gamma^* := \max\{\gamma_i \ : \ i = 1, 2, \ldots, n\}$.

Finalmente,

$$
-\sqrt{n\gamma^*} \leq \mathcal{A}x^3 \leq \sqrt{n\gamma^*}.
$$

O tempo para obter este limite inferior é dominado pelo cálculo de todos os autovalores das $n \times n$-matrizes simétricas $A_1, \ldots, A_n$.

3.3.2 Limite inferior por decomposição: Abordagem 2

Para desenvolver a segunda abordagem decompomos a função objetivo de PEC para o primeiro índice, como segue:

$$
\sum_{i,j,k=1}^{n} a_{ijk}x_i x_j x_k = \sum_{i=1}^{n}\Big(x_i \sum_{\substack{j,k=1\\ j,k\neq i}}^{n} a_{ijk}x_j x_k + 2x_i^2\sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}x_j + a_{iii}x_i^3\Big) \tag{3.24}
$$

Então concluimos que

$$\min_{\|x\|=1} \sum_{i,j,k=1}^{n} a_{ijk}x_ix_jx_k \geq$$
$$\sum_{i=1}^{n} \min_{\|x\|=1} \Big(x_i \sum_{\substack{j,k=1 \\ j,k\neq i}}^{n} a_{ijk}x_jx_k + 2x_i^2 \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{iij}x_j + a_{iii}x_i^3 \Big) . \tag{3.25}$$

Com mais uma decomposição, para cada $i = 1, 2, \ldots, n$, temos

$$\min_{\|x\|=1} \Big(x_i \sum_{\substack{j,k=1 \\ j,k\neq i}}^{n} a_{ijk}x_jx_k + 2x_i^2 \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{iij}x_j + a_{iii}x_i^3 \Big) \geq$$
$$\min_{\|x\|=1} \; x_i \sum_{\substack{j,k=1 \\ j,k\neq i}}^{n} a_{ijk}x_jx_k + \min_{\|x\|=1} \; 2x_i^2 \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{iij}x_j \tag{3.26}$$
$$+ \min_{\|x\|=1} a_{iii}x_i^3.$$

Lema 3.3 *A seguinte igualdade é válida*

$$\min_{\|x\|=1} \; x_i \sum_{\substack{j,k=1 \\ j,k\neq i}}^{n} a_{ijk}x_jx_k = \min_{x_i\in[-1,1]} \; x_i \Big(\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j,k=1 \\ j,k\neq i}}^{n} a_{ijk}x_jx_k \Big). \tag{3.27}$$

Demonstração. Seja $x \in \mathbb{R}^n$ tal que $\|x\| = 1$, então

$$-1 \leq x_i \leq 1, \;\; \text{e} \;\; \|x_{\hat{i}}\| := \sqrt{\sum_{\substack{j=1 \\ j\neq i}}^{n} x_j^2} = \sqrt{1 - x_i^2},$$

logo

$$\min_{x_i\in[-1,1]} x_i \min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j,k=1 \\ j,k\neq i}}^{n} a_{ijk}x_jx_k \leq x_i \sum_{\substack{j,k=1 \\ j,k\neq i}}^{n} a_{ijk}x_jx_k$$

ou seja, o termo do lado esquerdo é um limite inferior para os elementos $x_i \sum_{\substack{j,k=1 \\ j,k\neq i}}^{n} a_{ijk}x_jx_k$. Portanto, por definição de infimo, obtemos

$$\min_{\|x\|=1} \; x_i \sum_{\substack{j,k=1 \\ j,k\neq i}}^{n} a_{ijk}x_jx_k \geq \min_{x_i\in[-1,1]} \; x_i \Big(\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j,k=1 \\ j,k\neq i}}^{n} a_{ijk}x_jx_k \Big). \tag{3.28}$$

Para demonstrar a outra desigualdade, seja $x_i \in [-1, 1]$ e $\|x_{\hat{i}}\| = \sqrt{1 - x_i^2}$, então

$\|x\| = 1$ e existem

$$\bar{x}_i = \min_{x_i \in [-1,1]} x_i \quad \text{e} \quad \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} \bar{x}_j \bar{x}_k = \min_{\|x_{\hat{i}}\| = \sqrt{1-x_i^2}} \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} x_j x_k,$$

deste modo,

$$\bar{x}_i \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} \bar{x}_j \bar{x}_k \geq \min_{\|x\|=1} \ x_i \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} x_j x_k$$

e portanto,

$$\min_{\|x\|=1} \ x_i \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} x_j x_k \leq \min_{x_i \in [-1,1]} \ x_i \Big(\min_{\|x_{\hat{i}}\| = \sqrt{1-x_i^2}} \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} x_j x_k \Big). \tag{3.29}$$

Finalmente, de (3.28) e (3.29), obtemos (3.27). ■

Analogamente ao Lema acima, pode ser demonstrado que

$$\min_{\|x\|=1} 2x_i^2 \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{iij} x_j = \min_{x_i \in [-1,1]} \ 2x_i^2 \Big(\min_{\|x_{\hat{i}}\| = \sqrt{1-x_i^2}} \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{iij} x_j \Big). \tag{3.30}$$

Substituindo (3.27) e (4.10), em (3.26), obtemos

$$\begin{aligned} \min_{\|x\|=1} \Big(x_i \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} x_j x_k + 2x_i^2 \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{iij} x_j + a_{iii} x_i^3 \Big) \geq & \\ \min_{x_i \in [-1,1]} \ x_i \Big(\min_{\|x_{\hat{i}}\| = \sqrt{1-x_i^2}} \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} x_j x_k \Big) & \\ + \min_{x_i \in [-1,1]} \ 2x_i^2 \Big(\min_{\|x_{\hat{i}}\| = \sqrt{1-x_i^2}} \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{iij} x_j \Big) & \\ + \min_{\|x\|=1} a_{iii} x_i^3. & \end{aligned} \tag{3.31}$$

Agora consideremos independentemente cada problema do lado direito de (3.31). Primeiro, usando o Lema 3.1, obtemos

$$\min_{\|x_{\hat{i}}\| = \sqrt{1-x_i^2}} \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} x_j x_k = (1 - x_i^2) \lambda_{\min}(\tilde{A}_i).$$

Multiplicando o lado direito por x_i e tomando o mínimo sobre $x_i \in [-1, 1]$, obtemos

$$\min_{x_i \in [-1,1]} \ x_i \Big(\min_{\|x_{\hat{i}}\| = \sqrt{1-x_i^2}} \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} x_j x_k \Big) = -\frac{2}{9} \sqrt{3} \ |\lambda_{\min}(\tilde{A}_i)| \ . \tag{3.32}$$

Além disso, usando o Lema (3.2), obtemos

$$\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{iij}x_j = (-1)\sqrt{\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{iij}^2}\ \sqrt{1-x_i^2}$$

multiplicando o lado direito por $2x_i^2$ e tomando o mínimo sobre $x_i \in [-1,1]$, obtemos

$$\min_{x_i\in[-1,1]}\ 2x_i^2\Big(\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{iij}x_j\Big) = -\frac{4}{9}\sqrt{3}\sqrt{\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{iij}^2}\ . \tag{3.33}$$

Finalmente,

$$\min_{\|x\|=1} a_{iii}x_i^3 = -|a_{iii}|\ . \tag{3.34}$$

de (3.25), (3.31), e (3.32)–(3.34) obtemos

$$\min_{\|x\|=1} \sum_{i,j,k=1}^{n} a_{ijk}x_ix_jx_k \geq \\ -\sum_{i=1}^{n}\left(\frac{2}{9}\sqrt{3}\ |\lambda_{\min}(\tilde{A}_i)| + \frac{4}{9}\sqrt{3}\sqrt{\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{iij}^2} + |a_{iii}|\right).$$

O tempo para obter este limite inferior é dominado pelo cálculo do menor autovalor das n simétricas $(n-1)\times(n-1)$-matrizes $\tilde{A}_1,\ldots,\tilde{A}_n$.

3.3.3 Limite inferior por dualidade: Abordagem 3

Nesta terceira abordagem, nosso objetivo é decompor a função objetivo do PEC em poucos termos, esperando obter um limite inferior mais apertado. Para isto, usamos

$$\min_{\|x\|=1} \sum_{i,j,k=1}^{n} a_{ijk}x_ix_jx_k \geq \sum_{i=1}^{n} \min_{\|x\|=1} \sum_{j,k=1}^{n} a_{ijk}x_ix_jx_k\ . \tag{3.35}$$

Note que, para cada $i = 1, 2, \ldots, n$ fixado, e de (3.27) obtemos

$$\min_{\|x\|=1} \sum_{j,k=1}^{n} a_{ijk}x_ix_jx_k = \min_{x_i\in[-1,1]}\ \min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{j,k=1}^{n} a_{ijk}x_ix_jx_k$$

assim,

$$\begin{aligned}
&\min_{\|x\|=1} \sum_{j,k=1}^{n} a_{ijk} x_i x_j x_k \\
=\ &\min_{x_i \in [-1,1]} \min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{j,k=1}^{n} a_{ijk} x_i x_j x_k \\
=\ &\min_{x_i \in [-1,1]} \min_{\|y\|=1} \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} x_i (1-x_i^2) y_j y_k + 2 \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{iij} x_i^2 \sqrt{1-x_i^2}\ y_j + a_{iii} x_i^3 \\
=\ &\min_{x_i \in [-1,1]} \min_{\|y\|=1} \Big(\ x_i(1-x_i^2) y^\top \tilde{A}_i y + x_i^2 \sqrt{1-x_i^2}\ a_i^\top y + a_{iii} x_i^3 \Big), \qquad (3.36)
\end{aligned}$$

onde definimos

$$\begin{aligned}
y &:= \tfrac{1}{\sqrt{1-x_i^2}} (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)^\top = \tfrac{1}{\sqrt{1-x_i^2}} x_{\hat{i}} \in \mathbb{R}^{n-1}\ , \\
a_i &:= 2(a_{ii1}, a_{ii2}, \ldots, a_{ii(i-1)}, a_{ii(i+1)}, \ldots, a_{ii(n-1)}, a_{iin})^\top \in \mathbb{R}^{n-1}\ .
\end{aligned}$$

A seguir, tomaremos proveito da decomposição espectral de $\tilde{A}_i$. Para isto, sejam $\lambda_{\min}(\tilde{A}_i) = \lambda_{i_1} \leq \ldots \leq \lambda_{i_n}$ os autovalores de $\tilde{A}_i$, e $v_{i_1}, \ldots, v_{i_n}$ seus correspondentes autovetores normalizados. Temos $\tilde{A}_i = V_i \Lambda_i V_i^\top$, onde $V_i := (v_{i_1}, \ldots, v_{i_n}) \in \mathbb{R}^{(n-1)\times(n-1)}$, com $V_i^\top V_i = V_i V_i^\top = I_{n-1}$, e $\Lambda_i := \mathrm{Diag}(\lambda_{i_1}, \ldots, \lambda_{i_n})$. Assim, para cada $i = 1, \ldots, n$ temos

$$\begin{aligned}
&\min_{\|y\|=1} \left(x_i(1-x_i^2) y^\top \tilde{A}_i y + x_i^2 \sqrt{1-x_i^2}\ a_i^\top y + a_{iii} x_i^3 \right) = \\
&\qquad \min_{\|z\|=1} \left(x_i(1-x_i^2) z^\top \Lambda_i z + x_i^2 \sqrt{1-x_i^2}\ b_i^\top z + a_{iii} x_i^3 \right),
\end{aligned} \qquad (3.37)$$

onde substituímos $z := V_i^\top y$ e $b_i := V_i^\top a_i$. Reparemos que x_i é uma constante neste contexto.

Note que o Problema (3.37) minimiza uma função quadrática sobre uma esfera unitária, o qual é uma instância do problema conhecido como subproblema da região de confiança (veja Seção 2.3), o qual pode ser resolvido eficientemente. O problema dual de (3.37), com *gap* de dualidade zero, é apresentado em [34] (veja também Seção 2.3.2). Para sua solução, três casos devem ser discriminados. Em todos os casos, um limite inferior de (3.37) é dado por

$$\max_{x_i(1-x_i^2)\Lambda_i - \mu I \succ 0} \left(\mu - \frac{1}{4} x_i^4 (1-x_i^2) b_i^\top \left(x_i(1-x_i^2)\Lambda_i - \mu I \right)^{-1} b_i + a_{iii} x_i^3 \right)\ . \qquad (3.38)$$

Assumindo que b_i tem entradas não nulas, temos que distinguir entre os casos $x_i \in \{-1, 0, 1\}$ e $x_i \in (-1, 0) \cup (0, 1)$. No último caso, temos $x_i^4(1-x_i^2) \neq 0$ e portanto o maximizador de (3.38) será um ponto interior do conjunto viável (chamado na literatura de "caso fácil"). No primeiro caso, (3.37) torna-se trivial, com valor ótimo $a_{iii}x_i^3$. Porém, nosso objetivo é derivar um limite sem saber x_i com antecedência, assim não podemos aplicar este caso a priori, o que significa que temos que usar o limite inferior dado por (3.38) em todos os casos.

Vamos agora dividir ainda mais o caso $x_i \in (-1,0)\cup(0,1)$ em dois subproblemas, isto é, x_i sendo negativo ou positivo. Assumindo $x_i \in (-1,0)$, definimos $\alpha = -\frac{\mu}{x_i(1-x_i^2)}$, substituimos μ por $-x_i(1-x_i^2)\alpha$ em (3.38), obtendo

$$\max_{-\Lambda_i-\alpha I\succ 0}\left(-x_i(1-x_i^2)\alpha - \frac{1}{4}x_i^4(1-x_i^2)b_i^\top\left(x_i(1-x_i^2)(\Lambda_i+\alpha I)\right)^{-1}b_i + a_{iii}x_i^3\right). \tag{3.39}$$

Portanto, para $x_i \in (-1,0)$, considerando a dualidade fraca, obtemos o limite inferior

$$\begin{aligned}\min_{\|y\|=1}\left(x_i(1-x_i^2)y^\top\tilde{A}_iy + x_i^2\sqrt{1-x_i^2}\;a_i^\top y + a_{iii}x_i^3\right) \geq \\ \max_{\alpha<-\lambda_{i_n}}\left(-x_i(1-x_i^2)\alpha - \frac{1}{4}x_i^3b_i^\top(\Lambda_i+\alpha I)^{-1}b_i + a_{iii}x_i^3\right).\end{aligned} \tag{3.40}$$

Analogamente, quando $x_i \in (0,1)$, definimos $\alpha = \frac{\mu}{x_i(1-x_i^2)}$, substituimos μ por $x_i(1-x_i^2)\alpha$ em (3.38), obtemos

$$\begin{aligned}\min_{\|y\|=1}\left(x_i(1-x_i^2)y^\top\tilde{A}_iy + x_i^2\sqrt{1-x_i^2}\;a_i^\top y + a_{iii}x_i^3\right) \geq \\ \max_{\alpha<\lambda_{i_1}}\left(x_i(1-x_i^2)\alpha - \frac{1}{4}x_i^3b_i^\top(\Lambda_i-\alpha I)^{-1}b_i + a_{iii}x_i^3\right).\end{aligned} \tag{3.41}$$

Levando em conta que nosso objetivo é encontrar limites inferiores para o PEC, nossa estratégia é fixar $\alpha = -\lambda_{i_n} - \epsilon$ em (3.40) e $\alpha = \lambda_{i_1} - \epsilon$ em (3.41), com $\epsilon > 0$, para cada $i = 1,\ldots,n$. Para cada valor de ϵ escolhido, obtemos um limite inferior definido por

$$\min_{\|x\|=1}\sum_{i,j,k=1}^{n} a_{ijk}x_ix_jx_k \geq \sum_{i=1}^{n}\min\{\omega_i,\nu_i,-|a_{iii}|\} \tag{3.42}$$

onde

$$\begin{aligned}\omega_i &:= \min_{x_i\in(-1,0)}\left(x_i(1-x_i^2)(\lambda_{i_n}+\epsilon) - \sum_{j=1}^{n}\frac{x_i^3(v_{i_j}^\top a_i)^2}{4(\lambda_{i_j}-\lambda_{i_n}-\epsilon)} + a_{iii}x_i^3\right),\\ \nu_i &:= \min_{x_i\in(0,1)}\left(x_i(1-x_i^2)(\lambda_{i_1}-\epsilon) - \sum_{j=1}^{n}\frac{x_i^3(v_{i_j}^\top a_i)^2}{4(\lambda_{i_j}-\lambda_{i_1}+\epsilon)} + a_{iii}x_i^3\right).\end{aligned} \tag{3.43}$$

O valor $-|a_{iii}|$ em (3.42) corresponde ao caso $x_i \in \{-1,0,1\}$. Note que os problemas de minimização (3.43) são problemas de otimização polinomial unidimensional de grau 3 e portanto podem ser facilmente resolvidos usando uma fórmula fechada.

O esforço computacional para obter este limite inferior é dominado pela diagonalização das $(n-1)\times(n-1)$-matrizes $\tilde{A}_1, \dots, \tilde{A}_n$ simétricas.

Observação 3.1 *Note que, para obter o limite inferior dado por (3.42), a desigualdade*

$$\min_{x_i} \max_{\alpha} f(x_i, \alpha) \geq \max_{\alpha} \min_{x_i} f(x_i, \alpha) \tag{3.44}$$

é fundamental. De fato, ao fixar $\alpha = -\lambda_{i_n} - \epsilon$ *em* (3.40) *e* $\alpha = \lambda_{i_1} - \epsilon$ *em* (3.41) *estamos discretizando o problema dual com o objetivo de nos aproximarmos da desigualdade (3.44). Analisaremos esses detalhes na Seção 3.3.5.*

3.3.4 Discretização no problema primal

Com o objetivo de avaliar o potencial de nossa abordagem de decomposição, adicionalmente consideramos desigualdades (3.35) para cada $i = 1, 2, \dots, n$ e resolvemos cada problema quadrático

$$\min_{\|y\|=1} \left(x_i(1-x_i^2) y^\top \tilde{A}_i y + x_i^2 \sqrt{1-x_i^2}\; a_i^\top y + a_{iii} x_i^3 \right) \tag{3.45}$$

usando o algoritmo introduzido em [38]. Isto, no entanto, não produz uma solução de forma fechada, de modo que, diferentemente das seções anteriores, não podemos minimizar a expressão resultante exatamente em $x_i \in [-1, 1]$. Em vez disso, nós discretizamos o intervalo $[-1, 1]$ e usamos os menores valores obtidos para qualquer ponto da grade. Note, no entanto, que esta abordagem não produz um limite inferior seguro em geral, uma vez que não podemos estimar o erro incorrido pela discretização.

3.3.5 Experimentos numéricos

Entre os primeiros objetivos alcançados podemos mencionar a introdução de novas abordagens para encontrar limites inferiores para a relaxação contínua de (PEC). Resultados computacionais são apresentados nos quais mostramos a eficiência das relaxações propostas.

Baseadas nas abordagens dadas, implementamos os programas para calcular os limites inferiores para os problemas de otimização um-esférico cúbico (PEC) em MATLAB R2017b. Nossos experimentos foram realizados em um computador com as seguintes caracteristicas: cluster of 64-bit Intel(R) Xeon(R) E5-4620 processador rodando em 2.20GHz com 252.4 GB de memoria.

Cabe ressaltar que resolvemos ambos problemas em (3.43) pela fórmula fechada, para 100 diferentes valores de ϵ igualmente distribuidos no intervalo $[0.01, 10]$, e reportamos o melhor limite obtido (isto é, o maior de todos).

Instâncias obtidas das referências

Consideramos primeiro 3 exemplos introduzidos em [10].

Exemplo 3.3 (Exemplo 3.2 em [10]) *Consideremos* $\mathcal{A} \in S^3(\mathbb{R}^3)$ *definido por*

$$\mathcal{A}_{111} = -0.1281, \mathcal{A}_{112} = 0.0516, \mathcal{A}_{113} = -0.0954,$$
$$\mathcal{A}_{122} = -0.1958, \mathcal{A}_{123} = -0.1790, \mathcal{A}_{133} = -0.2679,$$
$$\mathcal{A}_{222} = 0.3251, \mathcal{A}_{223} = 0.2513, \mathcal{A}_{233} = 0.1773, \mathcal{A}_{333} = 0.0338.$$

Exemplo 3.4 (Exemplo 3.3 em [10]) *Consideremos* $\mathcal{A} \in S^3(\mathbb{R}^3)$ *definido por*

$$\mathcal{A}_{111} = 0.0517, \mathcal{A}_{112} = 0.3579, \mathcal{A}_{113} = 0.5298, \mathcal{A}_{122} = 0.7544, \mathcal{A}_{123} = 0.2156,$$
$$\mathcal{A}_{133} = 0.3612, \mathcal{A}_{222} = 0.3943, \mathcal{A}_{223} = 0.0146, \mathcal{A}_{233} = 0.6718, \mathcal{A}_{333} = 0.9723.$$

Exemplo 3.5 (Exemplo 3.5 em [10]) *Consideremos* $\mathcal{A} \in S^3(\mathbb{R}^5)$ *definido por*

$$\mathcal{A}_{i_1,i_2,i_3} = \tfrac{(-1)^{i_1}}{i_1} + \tfrac{(-1)^{i_2}}{i_2} + \tfrac{(-1)^{i_3}}{i_3}.$$

Na Tabela 3.6, mostramos limites inferiores obtidos pelas 3 abordagens para estas instâncias. A melhor solução conhecida indicada na tabela é dada em [10] e é calculada mediante uma heurística apresentada no artigo.

Problema	Abordagem 1	Abordagem 2	Abordagem 3	Discretização	Melhor solução
Ex 3.3	−1.0967	−1.3172	−1.2683	−1.0849	−0.8730
Ex 3.4	−2.5984	−3.3009	−3.1877	−2.9730	−2.1110
Ex 3.5	−16.692	−20.9114	−18.5364	−11.3319	−9.9779

Tabela 3.6: Resultados para instâncias dadas nas referências.

Como mencionado na seção anterior, a Discretização não fornece um limite inferior rigoroso, já que em vez de resolver globalmente o problema (3.45) para cada i, obtemos a melhor solução ótima entre os problemas em que x_i é fixado dentro de um conjunto de discretização no intervalo $[-1, 1]$. Para ter uma ideia melhor da qualidade dessas soluções, fizemos um experimento onde conseguimos uma solução para (3.45), para cada i, 10 vezes. Primeiro, temos somente 5 pontos discretizados. Depois, em cada iteração $k = 2, \dots, 10$, adicionamos $50k$ pontos ao conjunto de discretização. Os pontos adicionados são sempre equidistantes. Na última iteração, consideramos 2255 pontos. Os limites inferiores obtidos para os diferentes números de pontos de discretização (npontos) são apresentados na Tabela 3.7. Observamos que quando incrementamos o número de pontos na discretização depois de 155 pontos as soluções obtidas são muito similares umas das outras. A diferença percentual relativa mostrada na última coluna, é dada por

$$dif.rel := (limite.inferior(k-1) - limite.inferior(k))/|limite.inferior(k-1)| * 100,$$

sempre é menor do que 0.04% quando $k > 2$, para os dois primeiros exemplos, e é menor do que 0.16% para o Exemplo 3.5. Estes resultados sugerem que os limites inferiores obtidos pela Discretização convergem rapidamente para um limite inferior válido quando o número de pontos na discretização aumenta.

Instâncias aleatórias

Finalmente, geramos tensores aleatórios $(n \times n \times n)$-dimensional simétricos $\mathcal{A}$, com entradas uniformemente distribuidas no intervalo $(0, 1)$. Os tensores foram gerados usando o software Tensor Toolbox for MATLAB [39]. As tabelas 3.8-3.9 reportam os valores médios para 20 instâncias para cada $n = 3, 5, 10, 30, 50, 100, 200$.

Calculamos os limites superiores de cada instância por meio de uma heurística simples escolhendo a melhor solução primal entre

(a) $x_i = 1$ e todas as outras componentes de x iguais a zero, para cada $i = 1, \ldots, n$.

(b) $x_i = -1$ e todas as outras componentes de x iguais a zero, para cada $i = 1, \ldots, n$.

(c) $x_i = 1/\sqrt{n}$ para todo $i = 1, \ldots, n$.

(d) $x_i = -1/\sqrt{n}$ para todo $i = 1, \ldots, n$.

Para a Discretização no problema primal, para cada $i = 1, \ldots, n$, resolvemos o problema quadrático (3.45), para 200 pontos igualmente espaçados x_i no intervalo $[-1, 1]$, usando o algoritmo descrito em [38]. O tempo computacional necessário para esta abordagem aumenta significativamente com n. Portanto, somente aplicaremos esta abordagem para instâncias menores, na Tabela 3.8. Para instâncias de maior porte na Tabela 3.9, aplicamos nossas três abordagens realmente destinadas a gerar limites inferiores para o PEC.

Ressaltamos mais uma vez que o objetivo principal da aplicação da Discretização é ter uma avaliação da qualidade dos limites inferiores calculados pelas outras abordagens. Note que, no caso do número de pontos discretos na Discretização se aproximar do infinito, devemos ter sua solução convergindo para o melhor limite possível dado pela Abordagem 1.

Notamos também que o tempo computacional da Abordagem 2 ainda pode ser reduzido, porque os tempos computacionais relatados para a Abordagem 2 levam em consideração a solução dos problemas de minimização em (3.43) para 100 valores diferentes de ϵ. Esta estratégia pode tornar-se mais prática, através de uma melhor análise do problema com o objetivo de reduzir o número de soluções duais consideradas, sem comprometer em demasia a qualidade dos limites. A melhoria nesses cálculos faz parte de nossa pesquisa futura.

3.4 Uma aplicação do PEC em processamento de sinais

Nesta seção descreveremos como o problema PEC pode modelar um problema em sinais de ressonância magnética em tecidos biológicos. Para isso, introduzimos alguns termos e conceitos na área de Ciências Biológicas. Para mais detalhes veja [40–42].

3.4.1 Coeficientes de aparente assimetria em um modelo de imagem de tensor de difusão

O modelo de difusão tensora (DTI) é um modelo de ressonância magnética (MRI) usado para estudar as propriedades da difusão da molécula de água em fibras de matéria branca no cérebro. Fibras de matéria branca conectam diferentes partes do cérebro e devem ser protegidas durante uma cirurgia, o que torna o estudo muito relevante. Uma perfeita distribuição Gaussiana para o movimento da molécula de água é assumida no modelo DTI [8]. No entanto, a água frequentemente mostra um comportamento de difusão não-gaussiano em estruturas biológicas. A fim de superar essa desvantagem do modelo DTI, Liu et. al [9] propôs a chamada imagem de tensores de difusão generalizada (GDTI) para caracterizar a difusão não gaussiana das moléculas de água nos tecidos. Nesta seção, consideramos a seguinte aproximação do GDTI:

$$\ln\left(\frac{S(b)}{S(0)}\right) = -\sum_{i_1,i_2=1}^{3} D^{(2)}_{i_1 i_2} b^{(2)}_{i_1 i_2} + \sum_{i_1,\cdots,i_4=1}^{3} D^{(4)}_{i_1\cdots i_4} b^{(4)}_{i_1\cdots i_4} - j \sum_{i_1,i_2,i_3=1}^{3} D^{(3)}_{i_1 i_2 i_3} b^{(3)}_{i_1 i_2 i_3}, \quad (3.46)$$

onde os tensores de segunda, terceira e quarta ordem ($D^{(2)}$, $D^{(3)}$, e $D^{(4)}$) são respectivamente, o tensor de difusão, o tensor de assimetria de difusão (DS) e o tensor de curtose de difusão (DK), e $b^{(n)} (n = 2, 3, 4)$ são funções da direção, a magnitude e a temporização dos gradientes codificadores da difusão.

Para entender o significado biológico e clínico dos tensores, é importante medir e calcular algumas grandezas e parâmetros relacionados a eles. Neste contexto, o objetivo final neste procedimento é calcular os maiores e menores coeficientes de aparente assimetria (ASC) que estão associados com o tensor DS. Como mostrado em [43], isso pode ser feito maximizando e minimizando

$$\sum_{i_1,i_2,i_3=1}^{3} D^{(3)}_{i_1 i_2 i_3} b^{(3)}_{i_1 i_2 i_3},$$

que pode finalmente ser modelado como os problemas de otimização não convexos:

$$\max(\min) \left\{ Px^3 \equiv \sum_{i,j,k=1}^{3} P_{ijk} x_i x_j x_k : ||x||^2 = 1 \right\},$$

onde P é um tensor dado 3-dimensional.

Na literatura, existem muitos métodos numéricos que podem ser usados para calcular os valores de ASC, veja [2, 25, 43].

3.4.2 O modelo de imagem de tensor de difusão generalizado (DTI)

Liu et. al [9] propôs o seguinte modelo generalizado de tensores de difusão (GDTI)

$$\begin{aligned} \ln\left(\frac{S(b)}{S(0)}\right) &= -\sum_{i_1,i_2=1}^{3} D^{(2)}_{i_1 i_2} b^{(2)}_{i_1 i_2} + \sum_{i_1,\cdots,i_4=1}^{3} D^{(4)}_{i_1\cdots i_4} b^{(4)}_{i_1\cdots i_4} + \cdots \\ &\quad +(-1)^n \sum_{i_1,\cdots,i_{2n}=1}^{3} D^{(2n)}_{i_1\cdots i_{2n}} b^{(2n)}_{i_1\cdots i_{2n}} + \cdots \\ &\quad +j\left(-\sum_{i_1,i_2,i_3=1}^{3} D^{(3)}_{i_1 i_2 i_3} b^{(3)}_{i_1 i_2 i_3} + \sum_{i_1,\cdots i_5=1}^{3} D^{(5)}_{i_1\cdots i_5} b^{(5)}_{i_1\cdots i_5} + \cdots \right. \\ &\quad \left. +(-1)^n \sum_{i_1,\cdots,i_{2n+1}=1}^{3} D^{(2n+1)}_{i_1\cdots i_{2n+1}} b^{(2n+1)}_{i_1\cdots i_{2n+1}} + \cdots \right), \end{aligned} \tag{3.47}$$

para caracterizar a difusão não gaussiana das moléculas de água nos tecidos, onde $S(0)$ e $S(b)$ são a magnetização transversal medida (TE) na ausência e presença de gradiente de difusão, respectivamente, $j = \sqrt{-1}$, e $D^{(n)}$, para cada $n \geq 2$, é o tensor do coeficiente de ordem n, que pode ser determinado usando métodos estatísticos. Se o gradiente do campo magnético é um vetor constante sobre o tempo considerado, por [9], o elemento $b^{(n)}_{i_1 i_2 \ldots i_n}$ de tensor $b^{(n)}$ pode ser escrito como

$$b^{(n)}_{i_1 i_2 \ldots i_n} = (\gamma g \delta)^n \left(\Delta - \frac{n-1}{n+1}\delta\right) x_{i_1} x_{i_2} \cdots x_{i_n}, \quad i_1, i_2, \cdots, i_n = 1, 2, 3, \tag{3.48}$$

onde $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ representa a direção, Δ é o tempo de separação dos dois gradientes de difusão, δ é a duração de cada lóbulo de gradiente e g, γ é um número positivo apropriado.

Neste trabalho, consideramos a aproximação de (3.47), dada por (3.46), que pode ser obtida aproximando-se (3.47) pelo tensor de quarta ordem que contém informação útil do sinal magnético. Os dois primeiros termos em (3.46) estão relacionados com a magnitude do sinal e o último termo está relacionado com a fase do sinal.

3.4.3 Os valores ASC

Para definir nossos problemas de otimização, podemos usar a formulação proposta por Zhang et. al. em [43]. Escreveremos

$$P = (\gamma g \delta)^3 \left(\Delta - \frac{2}{4}\delta \right) D^{(3)}, \tag{3.49}$$

que é um tensor simétrico tridimensional de terceira ordem. O tensor P é simétrico, no sentido de que seu elemento P_{ijk} é invariante sob qualquer permutação de seus índices(i, j, k). Isto é, se dissermos $P_{112} = 5$, isso implica que $P_{121} = P_{211} = 5$. Portanto, os elementos independentes de P são $P_{111}; P_{112}; P_{113}; P_{122}; P_{123}; P_{133}; P_{222}; P_{223}; P_{233}; P_{333}$. Denotemos por $S_{app}(x)$, o coeficiente de aparente assimetria na direção x, dado por

$$S_{app}(x) = \frac{Px^3}{\|x\|^2}, \tag{3.50}$$

onde $Px^3 \equiv \sum_{i,j,k=1}^{3} P_{ijk} x_i x_j x_k$.

Denote os maiores e menores valores de ASC como S_{max} e S_{min}, respectivamente. Então, os problemas para resolver são:

$$\begin{aligned} S_{max} &= \max Px^3 \\ & \quad s.a. \ \|x\|^2 = 1, \end{aligned} \tag{3.51}$$

$$\begin{aligned} S_{min} &= \min Px^3 \\ & \quad s.a. \ \|x\|^2 = 1. \end{aligned} \tag{3.52}$$

3.4.4 Exemplo numérico

Exploramos a metodologia descrita na Seção 2.2 e aplicamos nossas abordagens para obter limites inferiores do problema em (3.52). Assim, o problema linear relaxado usando pontos de grade via RLT é formulado como

$$
\begin{array}{ll}
\min & \sum_{i,j,k=1}^{3} P_{ijk} X_{ijk} \\
s.t. & \sum_{i=1}^{3} X_{ii} = 1, \\
& \left[\prod_{j \in J_1} (x_j - l_j) \prod_{j \in J_2} (u_j - x_j) \prod_{(j,g) \in J_3} (x_j - \bar{x}_{jg})^2 \right]_L \geq 0, \\
& x \in \Omega := \{x \in \mathbb{R}^3 : -1 \leq x_j \leq 1, \ \ \forall j \in N\}
\end{array}
\tag{3.53}
$$

onde $[\cdot]_L$ denota a linearização de $[\cdot]$ sob a substituições $X_{ijk} = x_i x_j x_k$ e $X_{ij} = x_i x_j$, para todo $i, j, k \in N := \{1, 2, 3\}$, J_1, J_2, J_3 são conjuntos de indices em $N^3 := N \times N \times N$ e $\bar{x}_{jg}$ é igual a zero para todo $(j, g) \in N \times N$.

Exemplo 3.6 *Os dados deste exemplo foram retirados de [43]. Para o teste, definimos os parâmetros em (3.48) como:*

$$\Delta = 1, \quad \delta = 0.5, \quad g = 1, \quad \gamma = 1.$$

Então o tensor P em (3.49) tornar-se $P = \frac{3}{32} D^{(3)}$ e os dez elementos independentes de $D^{(3)}$ são

$$D^{(3)}_{111} = -2.36, D^{(3)}_{112} = 47.9, D^{(3)}_{113} = 0, D^{(3)}_{122} = -0.773, D^{(3)}_{123} = -0.575,$$

$$D^{(3)}_{133} = 0.282, D^{(3)}_{222} = -28.7, D^{(3)}_{223} = 0, D^{(3)}_{233} = 3.61, D^{(3)}_{333} = 0.488$$

em unidades de $10^{-8} mm^3/s$. De [43], temos que $\mathbf{S_{min} = -0.4922 \times 10^{-7}}$, atingido em $(0.8514, -0.5244, -0.0097)^T$. Na tabela 3.10 mostramos os resultados obtidos nas diferentes abordagens.

Exemplo 3.3

it	npontos	limite inferior	dif rel
1	5	-7.661209e-001	-
2	55	-1.082002e+000	41.2312600
3	155	-1.083636e+000	0.1510226
4	305	-1.083790e+000	0.0142051
5	505	-1.084534e+000	0.0686056
6	755	-1.084610e+000	0.0070751
7	1055	-1.084610e+000	0.0000000
8	1405	-1.084610e+000	0.0000000
9	1805	-1.084610e+000	0.0000000
10	2255	-1.084688e+000	0.0071920

Exemplo 3.4

it	npontos	limite inferior	dif rel
1	5	-2.117380e+000	-
2	55	-2.972419e+000	40.3819833
3	155	-2.973344e+000	0.0311089
4	305	-2.973650e+000	0.0102988
5	505	-2.973650e+000	0.0000000
6	755	-2.973650e+000	0.0000000
7	1055	-2.973688e+000	0.0012582
8	1405	-2.973770e+000	0.0027572
9	1805	-2.973770e+000	0.0000000
10	2255	-2.973770e+000	0.0000000

Exemplo 3.5

it	npontos	limite inferior	dif rel
1	5	-7.405235e+000	-
2	55	-1.051328e+001	41.9709771
3	155	-1.131087e+001	7.5864276
4	305	-1.131287e+001	0.0176827
5	505	-1.131301e+001	0.0012360
6	755	-1.131359e+001	0.0051961
7	1055	-1.131359e+001	0.0000000
8	1405	-1.131405e+001	0.0039842
9	1805	-1.133194e+001	0.1581602
10	2255	-1.133194e+001	0.0000000

Tabela 3.7: Discretização – Resultados para diferentes pontos de discretização.

n=3 (limite superior = -2.65210)		
Abordagem	limite inferior	tempo(seg.)
1	$\mathbf{-3.00017}$	0.001
2	-3.48020	0.002
3	-3.45158	0.043
Disc.	-3.44801	113.790

Tabela 3.8: Resultados para as instâncias randômicas, $n = 3$.

n=5 (limite superior = -5.6204)		
Abordagem	limite inferior	tempo(seg.)
1	$\mathbf{-6.1264}$	0.0006
2	-7.1952	0.0014
3	-8.2862	0.0807
n=10 (limite superior = -15.7504)		
Abordagem	limite inferior	tempo(seg.)
1	$\mathbf{-16.6472}$	0.0020
2	-19.1051	0.0016
3	-28.4392	0.1656
n=30 (limite superior = -82.1696)		
Abordagem	limite inferior	tempo(seg.)
1	$\mathbf{-84.2250}$	0.0174
2	-93.9775	0.0097
3	-220.1732	0.5579
n=50 (limite superior = -176.7865)		
Abordagem	limite inferior	tempo(seg.)
1	$\mathbf{-179.6662}$	0.1004
2	-197.1711	0.0587
3	-581.0479	1.1046
n=100 (limite superior = -499.9933)		
Abordagem	limite inferior	tempo(seg.)
1	$\mathbf{-504.4121}$	0.6789
2	-540.4582	0.6257
3	-2203.4445	5.8779
n=200 (limite superior = -1414.1020)		
Abordagem	limite inferior	tempo(seg.)
1	$\mathbf{-1421.0862}$	11.0411
2	-1495.4581	9.9779
3	-8483.9468	36.8817

Tabela 3.9: Resultados para as instâncias randômicas, $n = 5, 10, 30, 50, 100, 200$.

solução ótima $= -0.4922 \times 10^{-7}$		
Abordagem	limite inferior	tempo(seg.)
RLT	-1.1711×10^{-7}	0.0090
1	-0.8034×10^{-7}	7.0310e-04
2	-0.6919×10^{-7}	0.0180
3	-0.6576×10^{-7}	4.2310
Disc.	-0.6549×10^{-7}	34.1326

Tabela 3.10: Resultados para o exemplo em processamento de sinais.

Capítulo 4

O problema de otimização cúbico restrito à esfera

Neste capítulo estenderemos os resultados mostrados no capítulo anterior para encontrar limites inferiores para uma função cúbica geral restrita à esfera. Também introduziremos uma diferente metodologia para encontrar limites inferiores via Decomposição Lagrangeana.

4.1 O problema PC

Consideramos o problema de otimização cúbico restrito à esfera na forma

$$\begin{aligned} \text{PC}: \quad & \min_{x\in\mathbb{R}^n} \quad f(x) := \mathcal{A}_3x^3 + \mathcal{A}_2x^2 + \mathcal{A}_1x = \sum_{i,j,k=1}^{n} a_{ijk}x_ix_jx_k + \sum_{i,j=1}^{n} a_{ij}x_ix_j + \sum_{i=1}^{n} a_ix_i \\ & \text{s.a.} \quad \|x\| = 1, \end{aligned} \tag{4.1}$$

onde $n \geq 2$ e $\mathcal{A}_i$ é um tensor simétrico real de i-ésima ordem para cada $i = 1, 2, 3$. Em outras palavras, $\mathcal{A}_3$ é um tensor simétrico real $n \times n \times n$ dimensional, $\mathcal{A}_2$ é uma matriz simétrica real $n \times n$ dimensional, e $\mathcal{A}_1$ é um vetor n dimensional.

No caso particular quando $\mathcal{A}_2$ é uma matriz nula e $\mathcal{A}_1$ é um vetor nulo foi mostrado, por Zhang et al. [2] usando um resultado de Nesterov [3], que o problema (PC) é NP-difícil. Isto implica que no caso geral (PC) também é NP-difícil, em contraste com sua versão quadrática, a qual é o bem conhecido problema de região de confiança. Apesar da não convexidade, o problema quadrático pode ser resolvido devido a possuir um problema dual côncavo com gap de dualidade nula, veja Seção 2.3. Usaremos esses resultados para uma de nossas abordagens na Seção 4.2.

Na seguinte seção apresentaremos as extensões naturais das abordagens introduzidas no capítulo anterior para o problema (PC).

4.2 Relaxações para o PC

Nesta seção, o nosso objetivo é obter limites inferiores para o problema PC. Para isto, relaxamos o problema de diferentes maneiras. A ideia é decompor a função objetivo do PC em partes que possam ser minimizadas sobre a restrição $||x|| = 1$. Combinando todos os mínimos obtemos um limite inferior para o PC.

4.2.1 Limite inferior via autovalores: Abordagem 1

Primeiro notemos que

$$\begin{aligned} \min_{\|x\|=1} \mathcal{A}_3 x^3 + \mathcal{A}_2 x^2 + \mathcal{A}_1 x &= \min_{\|x\|=1} \Big(\sum_{i,j,k=1}^{n} a_{ijk} x_i x_j x_k + \sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} a_i x_i \Big) \\ &\geq \min_{\|x\|=1} \sum_{i,j,k=1}^{n} a_{ijk} x_i x_j x_k + \min_{\|x\|=1} \sum_{i,j=1}^{n} a_{ij} x_i x_j + \min_{\|x\|=1} \sum_{i=1}^{n} a_i x_i \end{aligned} \tag{4.2}$$

Agora vamos limitar o lado direito da desigualdade em (4.2). Note que para a primeira parcela, usando o resultado obtido na Seção 3.3.1, obtemos

$$\min_{\|x\|=1} \sum_{i,j,k=1}^{n} a_{ijk} x_i x_j x_k \geq -\sqrt{n\gamma^*} \tag{4.3}$$

onde $\gamma^* = \max\{\gamma_i : i = 1, 2, \ldots, n\}$ e $\gamma_\ell = \max\{\lambda^2 : \lambda \text{ é um autovalor da matriz } A_\ell\}$.

Por outro lado, do Lema 3.1, obtemos

$$\min_{\|x\|=1} \sum_{i,j=1}^{n} a_{ij} x_i x_j = \lambda_{\min}(\mathcal{A}_2) \tag{4.4}$$

e do Lema 3.2, obtemos

$$\min_{\|x\|=1} \sum_{i=1}^{n} a_i x_i = -\|\mathcal{A}_1\|, \tag{4.5}$$

portanto, de (4.3)–(4.5) em (4.2), obtemos

$$\min_{\|x\|=1} \mathcal{A}_3 x^3 + \mathcal{A}_2 x^2 + \mathcal{A}_1 x \geq -\sqrt{n\gamma^*} + \lambda_{\min}(\mathcal{A}_2) - \|\mathcal{A}_1\| \tag{4.6}$$

O tempo para obter este limite inferior é dominado pelo cálculo do todos os autovalores das $n \times n$-matrizes simétricas $A_1, \ldots, A_n$.

4.2.2 Limite inferior por decomposição: Abordagem 2

Primeiro decompomos a função objetivo de PC para o primeiro índice, como segue:

$$\begin{aligned}&\sum_{i,j,k=1}^{n} a_{ijk}x_ix_jx_k + \sum_{i,j=1}^{n} a_{ij}x_ix_j + \sum_{i=1}^{n} a_ix_i = \\ &\quad \sum_{i=1}^{n}\Big(x_i\sum_{\substack{j,k=1\\ j,k\neq i}}^{n} a_{ijk}x_jx_k + 2x_i^2\sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}x_j + x_i\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j + a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i\Big)\end{aligned} \tag{4.7}$$

Então concluimos que

$$\begin{aligned}&\min_{\|x\|=1}\sum_{i,j,k=1}^{n} a_{ijk}x_ix_jx_k + \sum_{i,j=1}^{n} a_{ij}x_ix_j + \sum_{i=1}^{n} a_ix_i \geq \\ &\quad \sum_{i=1}^{n}\min_{\|x\|=1}\Big(x_i\sum_{\substack{j,k=1\\ j,k\neq i}}^{n} a_{ijk}x_jx_k + 2x_i^2\sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}x_j + x_i\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j + a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i\Big).\end{aligned} \tag{4.8}$$

Com mais uma decomposição, para cada $i = 1, 2, \ldots, n$, obtemos

$$\begin{aligned}&\min_{\|x\|=1}\Big(x_i\sum_{\substack{j,k=1\\ j,k\neq i}}^{n} a_{ijk}x_jx_k + 2x_i^2\sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}x_j + x_i\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j + a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i\Big) \geq \\ &\quad \min_{\|x\|=1} x_i\sum_{\substack{j,k=1\\ j,k\neq i}}^{n} a_{ijk}x_jx_k + \min_{\|x\|=1} 2x_i^2\sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}x_j + \min_{\|x\|=1} x_i\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j \\ &\quad + \min_{\|x\|=1} a_{iii}x_i^3 + \min_{\|x\|=1} a_{ii}x_i^2 + \min_{\|x\|=1} a_ix_i.\end{aligned} \tag{4.9}$$

Pelo demonstrado em (3.27) não é dificil mostrar que

$$\begin{aligned}\min_{\|x\|=1} x_i\sum_{\substack{j,k=1\\ j,k\neq i}}^{n} a_{ijk}x_jx_k &= \min_{x_i\in[-1,1]} x_i\Big(\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}}\sum_{\substack{j,k=1\\ j,k\neq i}}^{n} a_{ijk}x_jx_k\Big)\\ \min_{\|x\|=1} 2x_i^2\sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}x_j &= \min_{x_i\in[-1,1]} 2x_i^2\Big(\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}}\sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}x_j\Big)\\ \min_{\|x\|=1} x_i\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j &= \min_{x_i\in[-1,1]} x_i\Big(\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}}\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j\Big)\end{aligned} \tag{4.10}$$

assim, substituindo (4.10) em (4.9), obtemos

$$\min_{\|x\|=1}\Big(x_i \sum_{\substack{j,k=1\\ j,k\neq i}}^{n} a_{ijk}x_jx_k + 2x_i^2\sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}x_j + x_i\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j + a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i\Big) \geq$$
$$\min_{x_i\in[-1,1]} x_i\Big(\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j,k=1\\ j,k\neq i}}^{n} a_{ijk}x_jx_k\Big) + \min_{x_i\in[-1,1]} 2x_i^2\Big(\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}x_j\Big)$$
$$+ \min_{x_i\in[-1,1]} x_i\Big(\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j\Big) + \min_{\|x\|=1} a_{iii}x_i^3 + \min_{\|x\|=1} a_{ii}x_i^2 + \min_{\|x\|=1} a_ix_i. \tag{4.11}$$

Agora consideremos independentemente cada problema do lado direito de (4.11). Primeiro, usando o Lema 3.1, note que

$$\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j,k=1\\ j,k\neq i}}^{n} a_{ijk}x_jx_k = (1-x_i^2)\lambda_{\min}(\tilde{A}_i),$$

multiplicando o lado direito por x_i e tomando o mínimo sobre $x_i \in [-1,1]$, obtemos

$$\min_{x_i\in[-1,1]} x_i\Big(\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j,k=1\\ j,k\neq i}}^{n} a_{ijk}x_jx_k\Big) = -\frac{2}{9}\sqrt{3}\,\left|\lambda_{\min}(\tilde{A}_i)\right|. \tag{4.12}$$

Além disso, do Lema 3.2, temos

$$\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}x_j = (-1)\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}^2}\,\sqrt{1-x_i^2}$$

e

$$\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j = (-1)\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}^2}\,\sqrt{1-x_i^2}$$

multiplicando o lado direito das equações acima por $2x_i^2$ e x_i respectivamente e tomando o mínimo sobre $x_i \in [-1,1]$, obtemos

$$\min_{x_i\in[-1,1]} 2x_i^2\Big(\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}x_j\Big) = -\frac{4}{9}\sqrt{3}\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}^2} \tag{4.13}$$

$$\min_{x_i\in[-1,1]} x_i\Big(\min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j\Big) = -\frac{1}{2}\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}^2} \tag{4.14}$$

Finalmente,

$$\min_{\|x\|=1} a_{iii}x_i^3 = -|a_{iii}| \ . \tag{4.15}$$

$$\min_{\|x\|=1} a_{ii}x_i^2 = \min\{0, a_{ii}\} \ . \tag{4.16}$$

$$\min_{\|x\|=1} a_i x_i = -|a_i| \ . \tag{4.17}$$

Adicionando, de (4.8), (4.11), e (4.12)–(4.17) obtemos

$$\begin{aligned}\min_{\|x\|=1} \mathcal{A}_3 x^3 + \mathcal{A}_2 x^2 + \mathcal{A}_1 x \geq \\ -\sum_{i=1}^{n} \left(\frac{2}{9}\sqrt{3}\,|\lambda_{\min}(\tilde{A}_i)| + \frac{4}{9}\sqrt{3}\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}^2} + \frac{1}{2}\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}^2} + |a_{iii}| + \max\{0, -a_{ii}\} + |a_i| \right).\end{aligned}$$

O tempo para obter este limite inferior é dominado pelo cálculo do menor autovalor das n, $(n-1)\times(n-1)$-matrizes simétricas $\tilde{A}_1, \dots, \tilde{A}_n$.

4.2.3 Limite inferior por dualidade: Abordagem 3

Na terceira abordagem, nosso objetivo é decompor a função objetivo do PC em poucos termos, esperando obter um limite inferior mais apertado. Para isto, usamos

$$\begin{aligned}\min_{\|x\|=1} \sum_{i,j,k=1}^{n} a_{ijk}x_i x_j x_k + \sum_{i,j=1}^{n} a_{ij}x_i x_j + \sum_{i=1}^{n} a_i x_i \geq \\ \sum_{i=1}^{n} \min_{\|x\|=1} \Big(\sum_{j,k=1}^{n} a_{ijk}x_i x_j x_k + \sum_{j=1}^{n} a_{ij}x_i x_j + a_i x_i \Big).\end{aligned} \tag{4.18}$$

Note que, para cada $i = 1, 2, \dots, n$ fixado, de (3.27) obtemos

$$\begin{aligned}\min_{\|x\|=1} \sum_{j,k=1}^{n} a_{ijk}x_i x_j x_k + \sum_{i,j=1}^{n} a_{ij}x_i x_j + a_i x_i = \\ \min_{x_i\in[-1,1]} \min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{j,k=1}^{n} a_{ijk}x_i x_j x_k + \sum_{i,j=1}^{n} a_{ij}x_i x_j + a_i x_i\end{aligned}$$

assim,

$$
\begin{aligned}
&\min_{\|x\|=1} \sum_{j,k=1}^{n} a_{ijk}x_ix_jx_k + \sum_{j=1}^{n} a_{ij}x_ix_j + a_ix_i \\
= &\min_{x_i\in[-1,1]} \min_{\|x_{\hat{i}}\|=\sqrt{1-x_i^2}} \sum_{j,k=1}^{n} a_{ijk}x_ix_jx_k + \sum_{j=1}^{n} a_{ij}x_ix_j + a_ix_i \\
= &\min_{x_i\in[-1,1]} \min_{\|y\|=1} \sum_{\substack{j,k=1\\ j,k\neq i}}^{n} a_{ijk}x_i(1-x_i^2)y_jy_k + 2\sum_{\substack{j=1\\ j\neq i}}^{n} a_{iij}x_i^2\sqrt{1-x_i^2}\; y_j + a_{iii}x_i^3 \\
&\qquad + \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_i\sqrt{1-x_i^2}\; y_j + a_{ii}x_i^2 + a_ix_i \\
= &\min_{x_i\in[-1,1]} \min_{\|y\|=1} x_i(1-x_i^2)y^\top \tilde{A}_i y + \sum_{\substack{j=1\\ j\neq i}}^{n} \big(2a_{iij}x_i^2\sqrt{1-x_i^2} + a_{ij}x_i\sqrt{1-x_i^2}\big)y_j \\
&\qquad + a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i,
\end{aligned} \tag{4.19}
$$

onde definimos

$$
y := \tfrac{1}{\sqrt{1-x_i^2}}(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)^\top = \tfrac{1}{\sqrt{1-x_i^2}} x_{\hat{i}} \in \mathbb{R}^{n-1}
$$

A seguir, tomamos proveito da decomposição espectral de $\tilde{A}_i$. Para isto, sejam $\lambda_{\min}(\tilde{A}_i) = \lambda_{i_1} \leq \ldots \leq \lambda_{i_n}$ os autovalores de $\tilde{A}_i$, e $v_{i_1}, \ldots, v_{i_n}$ seus correspondentes autovetores normalizados. Temos $\tilde{A}_i = V_i\Lambda_iV_i^\top$, onde $V_i := (v_{i_1}, \ldots, v_{i_n}) \in \mathbb{R}^{(n-1)\times(n-1)}$, com $V_i^\top V_i = V_iV_i^\top = I_{n-1}$, e $\Lambda_i := \mathrm{Diag}(\lambda_{i_1}, \ldots, \lambda_{i_n})$. Assim, para cada $i = 1, \ldots, n$ temos

$$
\begin{aligned}
&\min_{\|y\|=1} x_i(1-x_i^2)y^\top \tilde{A}_i y + \sum_{\substack{j=1\\ j\neq i}}^{n} \big(2a_{iij}x_i^2\sqrt{1-x_i^2} + a_{ij}x_i\sqrt{1-x_i^2}\big)y_j + a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i = \\
&\qquad \min_{\|z\|=1} \left(x_i(1-x_i^2)z^\top\Lambda_i z + (x_i^2\sqrt{1-x_i^2}\; b_i + x_i\sqrt{1-x_i^2}\; \tilde{b}_i)^\top z + a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i \right),
\end{aligned} \tag{4.20}
$$

onde substituímos $z := V_i^\top y$, $b_i := 2V_i^\top \hat{a}_i$, e $\tilde{b}_i := V_i^\top \tilde{a}_i$.

$$
\hat{a}_i := (a_{ii1}, a_{ii2}, \ldots, a_{ii(i-1)}, a_{ii(i+1)}, \ldots, a_{iin})^\top
$$

$$
\tilde{a}_i := (a_{i1}, a_{i2}, \ldots, a_{i(i-1)}, a_{i(i+1)}, \ldots, a_{in})^\top.
$$

Repare que x_i é uma constante neste contexto e portanto o Problema (4.20) minimiza uma função quadrática sobre uma esfera unitária, sendo uma instância do problema conhecido como subproblema da região de confiança.. O problema dual de (4.20), com *gap* de dualidade nulo, é apresentado em [34] (veja Seção 2.3.2).

Para sua solução, três casos devem ser discriminados. Em todos os casos, um limite inferior de (4.20) é dado por

$$\max_{x_i(1-x_i^2)\Lambda_i-\mu I\succ 0} \mu - \frac{1}{4}x_i^2(1-x_i^2)(x_ib_i+\tilde{b}_i)^\top \left(x_i(1-x_i^2)\Lambda_i - \mu I\right)^{-1}(x_ib_i+\tilde{b}_i)$$
$$+a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i \ . \quad (4.21)$$

Assumindo que b_i tem entradas não nulas, temos que distinguir entre os casos $x_i \in \{-1,0,1\}$ e $x_i \in (-1,0)\cup(0,1)$.

Note que se $x_i \in \{-1,0,1\}$, (4.21) é trivial, com valor ótimo $-a_{iii}+a_{ii}-a_i$, ou $a_{iii}+a_{ii}+a_i$.

Vamos dividir o caso $x_i \in (-1,0)\cup(0,1)$ em dois subproblemas, isto é, x_i sendo negativo ou positivo. Assumindo $x_i \in (-1,0)$, definindo $\alpha := -\frac{\mu}{x_i(1-x_i^2)}$ e substituindo μ por $-x_i(1-x_i^2)\alpha$ em (4.21), obtemos

$$\max_{-\Lambda_i-\alpha I\succ 0} -x_i(1-x_i^2)\alpha - \frac{1}{4}x_i^2(1-x_i^2)(x_ib_i+\tilde{b}_i)^\top \left(x_i(1-x_i^2)(\Lambda_i+\alpha I)\right)^{-1}(x_ib_i+\tilde{b}_i)$$
$$+a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i$$

Portanto, para $x_i \in (-1,0)$, obtemos o limite inferior

$$\min_{\|y\|=1} x_i(1-x_i^2)y^\top\tilde{A}_iy + \sum_{\substack{j=1\\ j\neq i}}^n \left(2a_{iij}x_i^2\sqrt{1-x_i^2} + a_{ij}x_i\sqrt{1-x_i^2}\right)y_j + a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i \geq$$
$$\max_{\alpha<-\lambda_{i_n}} \left(-x_i(1-x_i^2)\alpha - \frac{1}{4}x_i(x_ib_i+\tilde{b}_i)^\top(\Lambda_i+\alpha I)^{-1}(x_ib_i+\tilde{b}_i) + a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i\right). \quad (4.22)$$

Analogamente, quando $x_i \in (0,1)$, definindo $\alpha := \frac{\mu}{x_i(1-x_i^2)}$ e substituindo μ por $-x_i(1-x_i^2)\alpha$ em (4.21), obtemos

$$\min_{\|y\|=1} x_i(1-x_i^2)y^\top\tilde{A}_iy + \sum_{\substack{j=1\\ j\neq i}}^n \left(2a_{iij}x_i^2\sqrt{1-x_i^2} + a_{ij}x_i\sqrt{1-x_i^2}\right)y_j + a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i \geq$$
$$\max_{\alpha<\lambda_{i_1}} \left(x_i(1-x_i^2)\alpha - \frac{1}{4}x_i(x_ib_i+\tilde{b}_i)^\top(\Lambda_i-\alpha I)^{-1}(x_ib_i+\tilde{b}_i) + a_{iii}x_i^3 + a_{ii}x_i^2 + a_ix_i\right). \quad (4.23)$$

Levando em conta que nosso objetivo é encontrar limites inferiores para o problema PC, nossa estratégia é fixar $\alpha = -\lambda_{i_n} - \epsilon$ em (4.22) e $\alpha = \lambda_{i_1} - \epsilon$ em (4.23), com $\epsilon > 0$, para cada $i = 1,\ldots,n$.

De (4.18) e para cada valor de ϵ escolhido, obtemos um limite inferior definido por

$$\min_{\|x\|=1} \mathcal{A}_3x^3 + \mathcal{A}_2x^2 + \mathcal{A}_1x \geq \sum_{i=1}^n \min\{\omega_i, \nu_i, -a_{iii}+a_{ii}-a_i, a_{iii}+a_{ii}+a_i\} \quad (4.24)$$

onde

$$
\begin{aligned}
\omega_i &:= \min_{x_i \in (-1,0)} \left(x_i(1-x_i^2)(\lambda_{i_n}+\epsilon) - \sum_{j=1}^{n} \frac{x_i(x_i b_i + \tilde{b}_i)_j^2}{4(\lambda_{i_j} - \lambda_{i_n} - \epsilon)} + a_{iii}x_i^3 + a_{ii}x_i^2 + a_i x_i \right), \\
\nu_i &:= \min_{x_i \in (0,1)} \left(x_i(1-x_i^2)(\lambda_{i_1}-\epsilon) - \sum_{j=1}^{n} \frac{x_i(x_i b_i + \tilde{b}_i)_j^2}{4(\lambda_{i_j} - \lambda_{i_1} + \epsilon)} + a_{iii}x_i^3 + a_{ii}x_i^2 + a_i x_i \right).
\end{aligned}
\tag{4.25}
$$

O valor $-a_{iii}+a_{ii}-a_i$ e $a_{iii}+a_{ii}+a_i$ em (4.24) correspondem ao caso $x_i \in \{-1,0,1\}$. Note que os problemas de minimização (4.25) são problemas de otimização polinomial unidimensional de grau 3 e portanto podem ser facilmente resolvidos usando uma fórmula fechada.

O esforço computacional para obter este limite inferior é dominado pela diagonalização das n $(n-1)\times(n-1)$-matrizes $\tilde{A}_1,\ldots,\tilde{A}_n$ simétricas.

Observação 4.1 *Note que, analogamente ao caso do Capítulo 3, para obter o limite inferior dado por (4.24), a desigualdade*

$$\min_{x_i} \max_{\alpha} f(x_i,\alpha) \geq \max_{\alpha} \min_{x_i} f(x_i,\alpha) \tag{4.26}$$

é fundamental. De fato, ao fixar $\alpha = -\lambda_{i_n} - \epsilon$ *em* (4.22) *e* $\alpha = \lambda_{i_1} - \epsilon$ *em* (4.23) *estamos discretizando o problema dual com o objetivo de nos aproximar da desigualdade (4.26). Analisaremos esses detalhes na seguinte seção.*

4.2.4 Experimentos numéricos

Dentre os primeiros objetivos alcançados podemos mencionar a introdução de novas abordagens para encontrar limites inferiores para PC.

Experimentos computacionais são apresentados nos quais mostramos a eficiência das relaxações propostas.

Resolvemos ambos problemas em (4.25) pela fórmula fechada, para 100 diferentes valores de ϵ igualmente distribuidos no intervalo $[0.01, 10]$, e reportamos o melhor limite obtido (isto é, o maior de todos).

Geramos tensores simétricos aleatórios $(n\times n\times n)$-dimensional $\mathcal{A}$, com entradas uniformemente distribuidas em $(0,1)$.

Os tensores foram gerados usando o software Tensor Toolbox for MATLAB [39].

Calculamos os limites superiores de cada instância por meio de uma heurística simples, escolhendo a melhor solução primal dentre

(a) $x_i = 1$ e todas as outras componentes de x iguais a zero, para cada $i = 1,\ldots,n$.

(b) $x_i = -1$ e todas as outras componentes de x iguais a zero, para cada $i = 1,\ldots,n$.

(c) $x_i = 1/\sqrt{n}$ para todo $i = 1, \ldots, n$.

(d) $x_i = -1/\sqrt{n}$ para todo $i = 1, \ldots, n$.

A Tabela 4.1 reporta os valores médios para 20 instâncias para cada $n = 3, 5, 10, 30, 50, 100$.

Observamos dos resultados mostrados na Tabela 4.1 que a decomposição feita na Abordagem 3 pode melhorar os limites obtidos pela Abordagem 2 com um esforço computacional mediano.

Notamos também que o tempo computacional da Abordagem 3 ainda pode ser reduzido, porque os tempos computacionais relatados para a Abordagem 3 levam em consideração a solução dos problemas de minimização em (4.25) para 100 valores diferentes de ϵ. Esta estratégia pode tornar-se mais prática, através de uma melhor análise do problema com o objetivo de reduzir o número de soluções duais consideradas, sem comprometer demasiadamente a qualidade dos limites. A melhoria nesses cálculos faz parte de nossa pesquisa futura.

n=3 (limite superior = -1.9281)		
Abordagem	limite inferior	tempo
1	$\mathbf{-2.4723}$	0.0023
3	-2.7252	0.1692
2	-5.8411	0.0054
n=5 (limite superior = -4.4185)		
Abordagem	limite inferior	tempo
1	$\mathbf{-5.0556}$	0.0009
3	-6.9968	0.2781
2	-12.6322	0.0010
n=10 (limite superior = -12.3386)		
Abordagem	limite inferior	tempo
1	$\mathbf{-13.4735}$	0.0017
3	-22.5972	0.5651
2	-32.6065	0.0020
n=20 (limite superior = -36.8650)		
Abordagem	limite inferior	tempo
1	$\mathbf{-38.7062}$	0.0065
3	-84.3078	1.1823
2	-86.9546	0.0066
n=50 (limite superior = -155.2720)		
Abordagem	limite inferior	tempo
1	$\mathbf{-158.4688}$	0.1151
2	-322.5473	0.0520
3	-502.7512	3.6473
n=100 (limite superior = -454.8378)		
Abordagem	limite inferior	tempo
1	$\mathbf{-460.3093}$	0.6890
2	-876.7134	0.2749
3	-1976.1566	20.2721

Tabela 4.1: Comparações entre as Abordagens 1-3.

4.3 Decomposição Lagrangeana

Os Métodos de Decomposição resolvem problemas de grande escala dividindo-os em vários subproblemas menores que são acoplados através de um problema mestre. Decomposição de problemas de otimização começaram com a chamada decomposição Dantzig-Wolfe de programas lineares. Esse método está ligado ao método dual simplex, que ainda é um dos métodos mais eficazes para resolver programas lineares. Existem vários princípios de decomposição, dentre os quais podemos mencionar: decomposição lagrangeana (método dual), método primal de plano de corte , geração de colunas e decomposição de Benders. As abordagens diferem principalmente na definição dos problemas mestre.

Nesta seção, nosso objetivo é usar os fundamentos da técnica de decomposição Lagrangeana juntamente com a abordagem 3 (veja Seção 4.2.3).

A seguir, apresentaremos os princípios do método de Decomposição Lagrangeana.

4.3.1 Fundamentos do método de Decomposição Lagrangeana

Considere um problema de otimização separável por blocos da forma:

$$\begin{array}{ll} \min & f(x) \\ \text{s. a} & g(x) \leq 0 \\ & x \in G, \end{array} \tag{4.27}$$

onde as funções $f : \mathbb{R}^n \to \mathbb{R}$ e $g : \mathbb{R}^n \to \mathbb{R}^m$ e o conjunto $G \subset \mathbb{R}^n$ são separáveis por blocos, ou seja, existe uma partição $\{J_1, J_2, \ldots, J_p\}$ de $\{1, 2, \ldots, n\}$ tal que

$$f(x) = \sum_{k=1}^{p} f_k(x_{J_k}), \quad g(x) = \sum_{k=1}^{p} g_k(x_{J_k})$$

e

$$G = \{x \in \mathbb{R}^n \, : \, x_{J_k} \in G_k, \, k = 1, 2, \ldots, p\}.$$

Sejam

$$L(x; \mu) = f(x) + \mu^\top g(x)$$

a função Lagrangeana de (4.27) e

$$D(\mu) = \inf_{x \in G} L(x; \mu)$$

a função dual associada. Então o problema dual de (4.27) é

$$\sup_{\mu \in \mathbb{R}^m_+} D(\mu). \tag{4.28}$$

Como (4.27) é separável por blocos, a Relaxação Lagrangeana se decompõe em p problemas Lagrangeanos parciais

$$\begin{array}{ll} \inf & L_k(x_{J_k}; \mu) \\ \text{s. a} & x_{J_k} \in G_k \end{array} \tag{4.29}$$

onde

$$L_k(x_{J_k}; \mu) = f_k(x_{J_k}) + \sum_{i=1}^{m} \mu_i g_{i,k}(x_{J_k})$$

é a função Lagrangeana parcial associada ao bloco k.

Seja $D_k(\mu)$ o valor ótimo do (4.29), então

$$D(\mu) = \sum_{k=1}^{p} D_k(\mu).$$

Esta simplificação é denominada Decomposição Lagrangeana.

O seguinte Lema (demonstrado em [44]) descreve duas propriedades da função dual que são exploradas em métodos de solução dual.

Lema 4.1 *(Lema 4.1 em [44])*

(i) *O dominio dom(D) da função dual D é convexo e D é côncavo sobre dom(D).*

(ii) *Seja $\lambda \in$ dom(D) um ponto dual dado. Então para todo x_λ ponto mínimo de $\min_{x \in G} L(x; \lambda)$ o vetor $g(x_\lambda)$ é um supergradiente de D em λ, ou seja,*

$$D(\mu) \leq D(\lambda) + g(x_\lambda)^\top (\mu - \lambda), \quad \forall \mu \in \mathbb{R}^m.$$

A seguir, apresentaremos um método de solução dual baseado em avaliações de subgradientes para a função dual.

4.3.2 Método subgradiente

O método subgradiente é o mais utilizado para a resolução do Lagrangeano dual.

Definamos o algoritmo para resolver o problema dual (4.28).

Seja $\{\alpha^j\}_{j \in \mathbb{N}}$ uma sequência de números reais não negativos. Denote a projeção de um ponto $\mu \in \mathbb{R}^m$ em $\mathbb{R}^m_+$ por $\Pi(\mu)$. O algoritmo subgradiente calcula a sequência de pontos duais $\{\mu^j\}$ de acordo aos seguintes passos:

Algoritmo 1 *Escolher um vetor inicial* $\mu^1 \in \mathbb{R}^m_+$.
Para $j = 1, 2, \ldots, l$
Seja g^j *um subgradiente de* D *em* μ^j.
Seja $\mu^{j+1} = \Pi(\mu^j - \alpha^j g^j / \|g^j\|)$
Fim

Note que o Algoritmo 1 não é necessariamente convergente, pois o subgradiente não é uma direção de subida com respeito a D. O seguinte resultado de convergência de $\{\mu^j\}$ é provado em [45].

Proposição 4.1 *Assumamos que o conjunto solução de (4.28) é não vazio e limitado. Então para qualquer tamanho de passo* $\{\alpha^j\}$ *satisfazendo*

$$\lim_{j \to \infty} \alpha^j = 0, \quad \sum_{j=1}^{\infty} \alpha^j = \infty,$$

a sequência $\{\mu^j\}$ *possui um ponto limite no conjunto solução de (4.28).*

Na prática, essa regra gera uma convergência lenta. A escolha do tamanho de passo $\alpha^j = q_0 q_i^j$, pode render a chamada taxa geométrica de convergência da distância de μ^j para a solução μ^* de (4.28), porém isso requer uma seleção meticulosa dos parámetros q_0, q_1([46]). Uma regra muito bem conhecida na literatura é a regra de Polyak:

$$\alpha^j = \theta^j (D(\mu^j) - D^j_{lev}) / \|g^j\|$$

onde D^j_{lev} é o melhor valor estimado para o ótimo de (4.28), e $0 < \delta < \theta^j$. A convergência das iterações $\{u^j\}$ para μ^* é obtida se D^j_{lev} tende ao valor ótimo de (4.28), veja [47].

A seguir aplicaremos estas metodologias para gerar nossa quarta abordagem.

4.3.3 Limite inferior por Decomposição Lagrangeana: Abordagem 4

Nosso objetivo é aplicar a metodologia de decomposição Lagrangeana junto com a abordagem de decomposição da função objetivo do PC em poucos termos, esperando obter um limite inferior mais apertado. Para isto, usamos

$$\begin{aligned} &\min_{\|x\|=1} \sum_{i,j,k=1}^{n} a_{ijk} x_i x_j x_k + \sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} a_i x_i \\ &= \min_{\|x\|=1} \sum_{i=1}^{n} \Big(x_i \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} x_j x_k + 2x_i^2 \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{iij} x_j + a_{iii} x_i^3 + x_i \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j + a_{ii} x_i^2 + a_i x_i \Big). \end{aligned} \tag{4.30}$$

Na função objetivo de (4.30) renomeamos as variáveis x_q por x_{qp} para todo $q = 1, 2, \dots, n$ e para cada parcela p do somatório ($p = 1, 2, \dots, n$), ou seja, a função objetivo do problema do lado direito de (4.30) torna-se

$$\begin{aligned}
&x_{11} \sum_{\substack{j,k=1 \\ j,k\neq 1}}^{n} a_{1jk} x_{j1} x_{k1} + 2x_{11}^2 \sum_{\substack{j=1 \\ j\neq 1}}^{n} a_{11j} x_{j1} + x_{11} \sum_{\substack{j=1 \\ j\neq 1}}^{n} a_{1j} x_{j1} + a_{111} x_{11}^3 + a_{11} x_{11}^2 + a_1 x_{11} \\
&+ x_{22} \sum_{\substack{j,k=1 \\ j,k\neq 2}}^{n} a_{2jk} x_{j2} x_{k2} + 2x_{22}^2 \sum_{\substack{j=1 \\ j\neq 2}}^{n} a_{22j} x_{j2} + x_{22} \sum_{\substack{j=1 \\ j\neq 2}}^{n} a_{2j} x_{j2} + a_{222} x_{22}^3 + a_{22} x_{22}^2 + a_2 x_{22} \\
&+ \cdots x_{nn} \sum_{\substack{j,k=1 \\ j,k\neq n}}^{n} a_{njk} x_{jn} x_{kn} + 2x_{nn}^2 \sum_{\substack{j=1 \\ j\neq n}}^{n} a_{nnj} x_{jn} + x_{nn} \sum_{\substack{j=1 \\ j\neq n}}^{n} a_{nj} x_{jn} + a_{nnn} x_{nn}^3 + a_{nn} x_{nn}^2 + a_n x_{nn}
\end{aligned} \tag{4.31}$$

Note que para garantir a igualdade com (4.30) devemos considerar as restrições

$$\begin{aligned}
x_{11} &= x_{1p}, \ \text{ para cada } \ p = 2, \dots, n \\
x_{21} &= x_{2p}, \ \text{ para cada } \ p = 2, \dots, n \\
&\vdots \\
x_{n1} &= x_{np}, \ \text{ para cada } \ p = 2, \dots, n,
\end{aligned} \tag{4.32}$$

Como $\|x\| = 1$, tem-se que

$$\begin{aligned}
x_{11}^2 + x_{21}^2 + \cdots + x_{n1}^2 &= 1 \\
x_{12}^2 + x_{22}^2 + \cdots + x_{n2}^2 &= 1 \\
&\vdots \\
x_{1n}^2 + x_{2n}^2 + \cdots + x_{nn}^2 &= 1.
\end{aligned} \tag{4.33}$$

Denotando

$$x^p := (x_{1p}, x_{2p}, \dots, x_{np}) \ \text{ para cada } \ p = 1, 2, \dots, n,$$

(4.32) é equivalente a

$$\begin{aligned}
x^1 &= x^2, \\
x^1 &= x^3, \\
&\vdots \\
x^1 &= x^n.
\end{aligned} \tag{4.34}$$

deste modo, de (4.33) e (4.34), o problema (4.30) torna-se equivalente a

$$\min \quad f(x^1, x^2, \dots, x^n) \tag{4.35}$$

$$\text{s. a} \quad x^1 = x^{k+1}, \ \ k = 1, \dots, n-1 \tag{4.36}$$

$$\|x^p\| = 1, \ \ p = 1, 2, \dots, n \tag{4.37}$$

onde

$$f(x^1, x^2, \dots, x^n) := \sum_{i=1}^{n} \Big(x_{ii} \sum_{\substack{j,k=1 \\ j,k \neq i}}^{n} a_{ijk} x_{ji} x_{ki} + 2x_{ii}^2 \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{iij} x_{ji} + x_{ii} \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_{ji} + a_{iii} x_{ii}^3 + a_{ii} x_{ii}^2 + a_i x_{ii} \Big).$$

Portanto, dualizando a restrição (4.36), obtemos

$$\begin{aligned} \min_{\|x\|=1} \sum_{i,j,k=1}^{n} a_{ijk} x_i x_j x_k + \sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} a_i x_i \geq \\ \min_{\substack{\|x^p\|=1, \\ p=1,2,\dots,n}} f(x^1, x^2, \dots, x^n) + \sum_{k=1}^{n-1} (\mu^k)^\top (x^{k+1} - x^1) \end{aligned} \tag{4.38}$$

onde μ^k são os multiplicadores de Lagrange associados às restrições (4.36).

Note que a função definida em (4.35) é uma função separável, ou seja, tem-se que

$$f(x^1, x^2, \dots, x^n) = f_1(x^1) + f_2(x^2) + \cdots + f_n(x^n)$$

com

$$f_p(x^p) := x_{pp} \sum_{\substack{j,k=1 \\ j,k \neq p}}^{n} a_{pjk} x_{jp} x_{kp} + 2x_{pp}^2 \sum_{\substack{j=1 \\ j \neq p}}^{n} a_{ppj} x_{jp} + x_{pp} \sum_{\substack{j=1 \\ j \neq p}}^{n} a_{pj} x_{jp} + a_{ppp} x_{pp}^3 + a_{pp} x_{pp}^2 + a_p x_{pp} \tag{4.39}$$

para $p = 1, 2, \dots, n$,

assim, de (4.38), obtemos

$$\begin{aligned} \min_{\|x\|=1} \sum_{i,j,k=1}^{n} a_{ijk} x_i x_j x_k &+ \sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} a_i x_i \\ &\geq \min_{\substack{\|x^p\|=1, \\ p=1,2,\dots,n}} f(x^1, x^2, \dots, x^n) + \sum_{k=1}^{n-1} (\mu^k)^\top (x^{k+1} - x^1) \\ &= \sum_{p=1}^{n} \min_{\|x^p\|=1} (f_p(x^p) + (\eta^p)^\top x^p), \end{aligned} \tag{4.40}$$

onde $\eta^p = \begin{cases} -\displaystyle\sum_{k=1}^{n-1} \mu^k, & \text{se } p = 1, \\ \mu^{p-1}, & \text{se } p = 2, 3, \dots, n \end{cases}$

Observação 4.2 *Repare que podemos encontrar limites inferiores dos problemas definidos em (4.40) (para cada $p = 1, 2, \dots, n$), aplicando a metodologia dada na*

Abordagem 3, Seção 4.2.3. Pois, de (4.39), para cada $p = 1, 2, \ldots, n$, *temos*

$$f_p(x^p) := x_{pp} \sum_{\substack{j,k=1\\ j,k\neq p}}^{n} a_{pjk}x_{jp}x_{kp} + 2x_{pp}^2 \sum_{\substack{j=1\\ j\neq p}}^{n} a_{ppj}x_{jp} + x_{pp} \sum_{\substack{j=1\\ j\neq p}}^{n} a_{pj}x_{jp} + a_{ppp}x_{pp}^3 + a_{pp}x_{pp}^2 + a_p x_{pp}$$

a qual é exatamente a mesma função analizada na Abordagem 3 e corresponde a fixarnos $\mu^k = 0$, *para* $k = 1, 2, \ldots, n-1$, *e repetida aqui em (4.31) denotando* $x_{pp} := x_p$. *Porém, agora acrescentamos o termo* $(\eta^p)^\top x^p$ *o que pode nos levar a encontrar limites inferiores mais apertados, a seguir continuaremos com essa análise.*

Agora, passaremos a encontrar limites inferiores para

$$\min_{\|x^p\|=1} (f_p(x^p) + (\eta^p)^\top x^p). \tag{4.41}$$

Note que, para cada índice $p = 1, 2, \ldots, n$, de (3.27) obtemos

$$\begin{aligned}
&\min_{\|x^p\|=1} x_{pp} \sum_{\substack{j,k=1\\ j,k\neq p}}^{n} a_{pjk}x_{jp}x_{kp} + 2x_{pp}^2 \sum_{\substack{j=1\\ j\neq p}}^{n} a_{ppj}x_{jp} + x_{pp} \sum_{\substack{j=1\\ j\neq p}}^{n} a_{pj}x_{jp} \\
&\qquad + a_{ppp}x_{pp}^3 + a_{pp}x_{pp}^2 + a_p x_{pp} + \sum_{j=1}^{n} \eta_{jp}x_{jp} \\
&= \min_{x_{pp}\in[-1,1]} \min_{\|x_{\hat{p}p}\|=\sqrt{1-x_{pp}^2}} x_{pp} \sum_{\substack{j,k=1\\ j,k\neq p}}^{n} a_{pjk}x_{jp}x_{kp} + 2x_{pp}^2 \sum_{\substack{j=1\\ j\neq p}}^{n} a_{ppj}x_{jp} + x_{pp} \sum_{\substack{j=1\\ j\neq p}}^{n} a_{pj}x_{jp} \\
&\qquad + a_{ppp}x_{pp}^3 + a_{pp}x_{pp}^2 + a_p x_{pp} + \sum_{j=1}^{n} \eta_{jp}x_{jp}
\end{aligned} \tag{4.42}$$

assim,

$$
\begin{aligned}
&\min_{\|x^p\|=1} (f_p(x^p) + (\eta^p)^\top x^p) \\
=\ &\min_{\|x^p\|=1} x_{pp} \sum_{\substack{j,k=1\\ j,k\neq p}}^{n} a_{pjk} x_{jp} x_{kp} + 2x_{pp}^2 \sum_{\substack{j=1\\ j\neq p}}^{n} a_{ppj} x_{jp} + x_{pp} \sum_{\substack{j=1\\ j\neq p}}^{n} a_{pj} x_{jp} + a_{ppp} x_{pp}^3 \\
&\quad + a_{pp} x_{pp}^2 + a_p x_{pp} + \sum_{j=1}^{n} (\eta^p)_j x_{jp} \\
=\ &\min_{x_{pp}\in[-1,1]} \min_{\|x_{\hat{p}p}\|=\sqrt{1-x_{pp}^2}} x_{pp} \sum_{\substack{j,k=1\\ j,k\neq p}}^{n} a_{pjk} x_{jp} x_{kp} + 2x_{pp}^2 \sum_{\substack{j=1\\ j\neq p}}^{n} a_{ppj} x_{jp} + x_{pp} \sum_{\substack{j=1\\ j\neq p}}^{n} a_{pj} x_{jp} \\
&\quad + a_{ppp} x_{pp}^3 + a_{pp} x_{pp}^2 + a_p x_{pp} + \sum_{j=1}^{n} (\eta^p)_j x_{jp} \\
=\ &\min_{x_{pp}\in[-1,1]} \min_{\|x_{\hat{p}p}\|=\sqrt{1-x_{pp}^2}} x_{pp} \sum_{\substack{j,k=1\\ j,k\neq p}}^{n} a_{pjk} x_{jp} x_{kp} + 2x_{pp}^2 \sum_{\substack{j=1\\ j\neq p}}^{n} a_{ppj} x_{jp} + x_{pp} \sum_{\substack{j=1\\ j\neq p}}^{n} a_{pj} x_{jp} \\
&\quad + a_{ppp} x_{pp}^3 + a_{pp} x_{pp}^2 + a_p x_{pp} + \sum_{\substack{j=1\\ j\neq p}}^{n} (\eta^p)_j x_{jp} + (\eta^p)_p x_{pp} \\
=\ &\min_{x_{pp}\in[-1,1]} \min_{\|y\|=1} \sum_{\substack{j,k=1\\ j,k\neq p}}^{n} a_{pjk} x_{pp} (1 - x_{pp}^2) y_j y_k + 2 \sum_{\substack{j=1\\ j\neq p}}^{n} a_{ppj} x_{pp}^2 \sqrt{1 - x_{pp}^2}\, y_j + a_{ppp} x_{pp}^3 \\
&\quad + \sum_{\substack{j=1\\ j\neq p}}^{n} a_{pj} x_{pp} \sqrt{1 - x_{pp}^2}\, y_j + a_{pp} x_{pp}^2 + a_p x_{pp} + \sum_{\substack{j=1\\ j\neq p}}^{n} (\eta^p)_j \sqrt{1 - x_{pp}^2} y_j + (\eta^p)_p x_{pp} \\
=\ &\min_{x_{pp}\in[-1,1]} \min_{\|y\|=1} x_{pp}(1 - x_{pp}^2) y^\top \tilde{A}_p y + \sum_{\substack{j=1\\ j\neq p}}^{n} \big(2a_{ppj} x_{pp}^2 \sqrt{1 - x_{pp}^2} + a_{pj} x_{pp} \sqrt{1 - x_{pp}^2} \\
&\quad + (\eta^p)_j \sqrt{1 - x_{pp}^2} \big) y_j + a_{ppp} x_{pp}^3 + a_{pp} x_{pp}^2 + (a_p + (\eta^p)_p) x_{pp}, \qquad (4.43)
\end{aligned}
$$

onde definimos

$$
y := \tfrac{1}{\sqrt{1-x_{pp}^2}} (x_{1p}, x_{2p}, \ldots, x_{(p-1)p}, x_{(p+1)p}, \ldots, x_{np})^\top = \tfrac{1}{\sqrt{1-x_{pp}^2}} x_{\hat{p}p} \in \mathbb{R}^{n-1}
$$

A seguir, tomamos proveito da decomposição espectral de $\tilde{A}_p$. Para isto, sejam $\lambda_{\min}(\tilde{A}_p) = \lambda_{p_1} \leq \ldots \leq \lambda_{p_n}$ os autovalores de $\tilde{A}_p$, e $v_{p_1}, \ldots, v_{p_n}$ seus correspondentes autovetores normalizados. Temos $\tilde{A}_p = V_p \Lambda_p V_p^\top$, onde $V_p := (v_{p_1}, \ldots, v_{p_n}) \in \mathbb{R}^{(n-1)\times(n-1)}$, com $V_p^\top V_p = V_p V_p^\top = I_{n-1}$, e $\Lambda_p := \mathrm{Diag}(\lambda_{p_1}, \ldots, \lambda_{p_n})$. Assim, para cada $p = 1, \ldots, n$ temos

$$
\begin{aligned}
&\min_{\|y\|=1} \; x_{pp}(1 - x_{pp}^2) y^\top \tilde{A}_p y + \sum_{\substack{j=1 \\ j \neq p}}^{n} \big(2a_{ppj} x_{pp}^2 \sqrt{1 - x_{pp}^2} + a_{pj} x_{pp} \sqrt{1 - x_{pp}^2} + (\eta^p)_j \sqrt{1 - x_{pp}^2}\big) y_j \\
&\quad + a_{ppp} x_{pp}^3 + a_{pp} x_{pp}^2 + (a_p + (\eta^p)_p) x_{pp} \\
&\quad = \min_{\|z\|=1} x_{pp}(1 - x_{pp}^2) z^\top \Lambda_p z + (x_{pp}^2 \sqrt{1 - x_{pp}^2}\, b_p + x_{pp} \sqrt{1 - x_{pp}^2} \tilde{b}_p + \sqrt{1 - x_{pp}^2} \bar{b}_p)^\top z \\
&\quad + a_{ppp} x_{pp}^3 + (a_p x_{pp}^2 + (\eta^p)_p) x_{pp},
\end{aligned} \tag{4.44}
$$

onde substituímos $z := V_p^\top y$, $b_p := 2V_p^\top \hat{a}_p$, $\tilde{b}_p := V_p^\top \tilde{a}_p$ e $\bar{b}_p := V_p^\top \bar{\eta}^p$.

$$
\begin{aligned}
\hat{a}_p &:= (a_{pp1}, a_{pp2}, \ldots, a_{pp(p-1)}, a_{pp(p+1)}, \ldots, a_{ppn})^\top \\
\tilde{a}_p &:= (a_{p1}, a_{p2}, \ldots, a_{p(p-1)}, a_{p(p+1)}, \ldots, a_{pn})^\top \\
\bar{\eta}^p &:= ((\eta^p)_1, (\eta^p)_2, \ldots, (\eta^p)_{p-1}, (\eta^p)_{p+1}, \ldots, (\eta^p)_n)^\top.
\end{aligned}
$$

Reparemos que x_{pp} é uma constante neste contexto e portanto o Problema (4.44) minimiza uma função quadrática sobre uma esfera unitária, o qual é uma instância do problema conhecido como subproblema da região de confiança. O problema dual de (4.44), com *gap* de dualidade nulo, é apresentado em [34] (veja Seção 2.3.2). Para sua solução, três casos devem ser discriminados. Em todos os casos, um limite inferior de (4.44) é dado por

$$
\begin{aligned}
\max_{x_{pp}(1-x_{pp}^2)\Lambda_p - \gamma I \succ 0} \gamma - \frac{1}{4}(1 - x_{pp}^2) d_p^\top \left(x_{pp}(1 - x_{pp}^2)\Lambda_p - \gamma I\right)^{-1} d_p \\
+ a_{ppp} x_{pp}^3 + a_{pp} x_{pp}^2 + (a_p + (\eta^p)_p) x_{pp},
\end{aligned} \tag{4.45}
$$

onde $d_p = x_{pp}^2 b_p + x_{pp} \tilde{b}_p + \bar{b}_p$.

Assumindo que d_p têm entradas não nulas, temos que distinguir entre os casos $x_{pp} \in \{-1, 0, 1\}$ e $x_{pp} \in (-1, 0) \cup (0, 1)$.

Note que se $x_{pp} \in \{-1, 0, 1\}$, (4.45) chega ser trivial, com valor ótimo $-a_{ppp} + a_{pp} - (a_p + (\eta^p)_p)$, ou $a_{ppp} + a_{pp} + (a_p + (\eta^p)_p)$.

Vamos dividir o caso $x_{pp} \in (-1, 0) \cup (0, 1)$ em dois subproblemas, isto é, x_{pp} sendo negativo ou positivo. Assumindo $x_{pp} \in (-1, 0)$, definindo $\alpha := -\frac{\gamma}{x_{pp}(1-x_{pp}^2)}$ e substituindo γ por $-x_{pp}(1 - x_{pp}^2)\alpha$ em (4.45), obtemos

$$
\begin{aligned}
\max_{-\Lambda_p - \alpha I \succ 0} -x_{pp}(1 - x_{pp}^2)\alpha - \frac{1}{4}(1 - x_{pp}^2) d_p^\top \left(x_{pp}(1 - x_{pp}^2)(\Lambda_p + \alpha I)\right)^{-1} d_p \\
+ a_{ppp} x_{pp}^3 + a_{pp} x_p^2 + (a_p + (\eta^p)_p) x_{pp}.
\end{aligned}
$$

Portanto, para $x_{pp} \in (-1,0)$, obtemos o limite inferior

$$\begin{aligned}
&\min_{\|y\|=1} \; x_{pp}(1-x_{pp}^2)y^\top \tilde{A}_p y + \sum_{\substack{j=1\\ j\neq p}}^{n} \big(2a_{ppj}x_{pp}^2\sqrt{1-x_{pp}^2} + a_{pj}x_{pp}\sqrt{1-x_{pp}^2} + (\eta^p)_j\sqrt{1-x_{pp}^2}\big)y_j \\
&\qquad +a_{ppp}x_{pp}^3 + a_{pp}x_{pp}^2 + (a_p + (\eta^p)_p)x_{pp} \\
&\geq \max_{\alpha < -\lambda_{p_n}} \; -x_{pp}(1-x_{pp}^2)\alpha - \frac{1}{4}(1-x_{pp}^2)d_p^\top \left(x_{pp}(1-x_{pp}^2)(\Lambda_p + \alpha I)\right)^{-1} d_p \\
&\qquad +a_{ppp}x_{pp}^3 + a_{pp}x_p^2 + (a_p + (\eta^p)_p)x_{pp}
\end{aligned} \tag{4.46}$$

Analogamente, quando $x_{pp} \in (0,1)$, definindo $\alpha := \frac{\gamma}{x_{pp}(1-x_{pp}^2)}$ e substituindo γ por $-x_{pp}(1-x_{pp}^2)\alpha$ em (4.45), obtemos

$$\begin{aligned}
&\min_{\|y\|=1} \; x_{pp}(1-x_{pp}^2)y^\top \tilde{A}_p y + \sum_{\substack{j=1\\ j\neq p}}^{n} \big(2a_{ppj}x_{pp}^2\sqrt{1-x_{pp}^2} + a_{pj}x_{pp}\sqrt{1-x_{pp}^2} + (\eta^p)_j\sqrt{1-x_{pp}^2}\big)y_j \\
&\qquad +a_{ppp}x_{pp}^3 + a_{pp}x_{pp}^2 + (a_p + (\eta^p)_p)x_{pp} \\
&\geq \max_{\alpha < \lambda_{p_1}} \; x_{pp}(1-x_{pp}^2)\alpha - \frac{1}{4}(1-x_{pp}^2)d_p^\top \left(x_{pp}(1-x_{pp}^2)(\Lambda_p - \alpha I)\right)^{-1} d_p \\
&\qquad +a_{ppp}x_{pp}^3 + a_{pp}x_p^2 + (a_p + (\eta^p)_p)x_{pp}
\end{aligned} \tag{4.47}$$

Levando em conta que nosso objetivo é encontrar limites inferiores para o problema PC, nossa estratégia é fixar $\alpha = -\lambda_{p_n} - \epsilon$ em (4.46) e $\alpha = \lambda_{p_1} - \epsilon$ em (4.47), com $\epsilon > 0$, para cada $p = 1, \ldots, n$.

Portanto, de (4.40) e para cada valor de ϵ escolhido, obtemos um limite inferior definido por

$$\begin{aligned}
&\min_{\|x\|=1} \mathcal{A}_3 x^3 + \mathcal{A}_2 x^2 + \mathcal{A}_1 x \geq \\
&\qquad \sum_{p=1}^{n} \min\left\{\omega_p, \nu_p, -a_{ppp} + a_{pp} - (a_p + (\eta^p)_p), a_{ppp} + a_{pp} + (a_p + (\eta^p)_p)\right\}
\end{aligned} \tag{4.48}$$

onde

$$\begin{aligned}
\omega_p &:= \min_{x_{pp}\in(-1,0)} x_{pp}(1-x_{pp}^2)(\lambda_{p_n} + \epsilon) - \sum_{j=1}^{n} \frac{(x_{pp}^2 b_p + x_{pp}\tilde{b}_p + \bar{b}_p)_j^2}{4x_{pp}(\lambda_{p_j} - \lambda_{p_n} - \epsilon)} \\
&\qquad +a_{ppp}x_{pp}^3 + a_{pp}x_{pp}^2 + (a_p + (\eta^p)_p)x_{pp}, \\
\nu_p &:= \min_{x_{pp}\in(0,1)} x_{pp}(1-x_{pp}^2)(\lambda_{p_1} - \epsilon) - \sum_{j=1}^{n} \frac{(x_{pp}^2 b_p + x_{pp}\tilde{b}_p + \bar{b}_p)_j^2}{4x_{pp}(\lambda_{p_j} - \lambda_{p_1} + \epsilon)} \\
&\qquad +a_{ppp}x_{pp}^3 + a_{pp}x_{pp}^2 + (a_p + (\eta^p)_p)x_{pp}\,.
\end{aligned} \tag{4.49}$$

O valor $-a_{ppp}+a_{pp}-(a_p+(\eta^p)_p)$ e $a_{ppp}+a_{pp}+(a_p+(\eta^p)_p)$ em (4.48) correspondem ao caso $x_{pp} \in \{-1, 1\}$. Note que os problemas de minimização (4.49) são problemas de otimização polinomial unidimensional e portanto podem ser resolvidos rápidamente.

Observação 4.3 *De forma geral, de (4.34) podemos decompor o problema PC de s diferentes maneiras ($s = 1, 2, \ldots, n$), na forma*

$$\min \quad f(x^1, x^2, \ldots, x^n) \tag{4.50}$$

$$s.\ a \quad x^s = x^k, \quad k = 1, 2, \ldots, n,\ k \neq s \tag{4.51}$$

$$\|x^p\| = 1, \quad p = 1, 2, \ldots, n \tag{4.52}$$

então obtemos n diferentes limites inferiores usando

$$\begin{aligned} \min_{\|x\|=1} \sum_{i,j,k=1}^{n} a_{ijk}x_i x_j x_k &+ \sum_{i,j=1}^{n} a_{ij}x_i x_j + \sum_{i=1}^{n} a_i x_i \\ &\geq \min_{\substack{\|x^p\|=1,\\ p=1,2,\ldots,n}} f(x^1, x^2, \ldots, x^n) + \sum_{\substack{k=1\\ k\neq s}}^{n} (\mu^k)^\top (x^k - x^s) \\ &= \sum_{p=1}^{n} \min_{\|x^p\|=1} (f_p(x^p) + (\eta^p)^\top x^p), \end{aligned} \tag{4.53}$$

onde

$$\eta^p = \begin{cases} -\sum\limits_{\substack{k=1\\ k\neq s}}^{n} \mu^k, & se\ \ p = s, \\ \mu^p, & se\ \ p \neq s \end{cases}$$

4.3.4 Experimentos numéricos

Resultados computacionais são apresentados nos quais mostramos a eficiência da relaxação proposta.

Implementamos as rotinas para calcular os limites inferiores para o problemas de otimização (PC), baseadas nas abordagens da previas seções, em MATLAB R2017b.

Nossos experimentos foram rodados em um computador com as seguintes caracteristicas: cluster of 64-bit Intel(R) Xeon(R) E5-4620 processador rodando em 2.20GHz com 252.4 GB de memoria.

Resolvemos ambos problemas em (4.49) pela fórmula fechada, para 100 diferentes valores de ϵ igualmente distribuidos no intervalo $[0.01, 10]$, e reportamos o melhor limite obtido (isto é, o maior de todos).

Geramos tensores aleatórios ($n \times n \times n$)-dimensional simétricos $\mathcal{A}_3$, matrizes aleatórios simétricas $\mathcal{A}_2$ e vetores aleatórios $\mathcal{A}_1$ com entradas uniformemente distribuidas em $(0, 1)$. Os tensores foram gerados usando o software Tensor Toolbox for MATLAB [39].

Calculamos os limites superiores de cada instância por meio de uma heurística simples, escolhendo a melhor solução primal entre

(a) $x_i = 1$ e todas as outras componentes de x igual a zero, para cada $i = 1, \ldots, n$.

(b) $x_i = -1$ e todas as outras componentes de x igual a zero, para cada $i = 1, \ldots, n$.

(c) $x_i = 1/\sqrt{n}$ para todo $i = 1, \ldots, n$.

(d) $x_i = -1/\sqrt{n}$ para todo $i = 1, \ldots, n$.

A Tabela 4.3 reporta os valores médios para 20 instâncias para cada $n = 3, 5, 10, 30, 50, 100$.

Vale ressaltar que os valores de μ^k são de muita importancia para obter limites inferiores mais apertados, nestes primeiros experimentos nós consideramos valores aleatórios de μ^k uniformemente distribuidos em $(0, 1)$ e tomamos aquele que outorga o melhor limite inferior.

Observamos dos resultados mostrados na Tabela 4.3 que a decomposição Lagrangeana melhora os limites obtidos pela Abordagem 3 com um esforço computacional não muito maior.

Notamos também que o tempo computacional da Abordagem 4 ainda pode ser reduzido, porque os tempos computacionais relatados para a Abordagem 4 levam em consideração a solução dos problemas de minimização em (4.49) para 100 valores diferentes de ϵ. A Abordagem 4 pode tornar-se mais prática, através de uma atualização dos multiplicadores de Lagrange, e usando o algoritmo dos subgradientes para isso. A melhoria nesses cálculos e o desenvolvimento de uma heurística para atualizar os multiplicadores de Lagrange μ^k faz parte de nossa pesquisa futura.

Nas Tabelas 4.2-4.4 denotamos por $s = \ell$ à dualização das restrições $x^\ell = x^k$, $k = 1, \ldots, n$, $k \neq \ell$. Para 20 instâncias aleatórias consideramos 10 diferentes valores de μ_i^k igualmente distribuidos no intervalo $[0.001, 0.1]$ para cada s decomposição e mostramos na tabela seguinte o melhor limite inferior obtido.

	Exemplo 3.3	Exemplo 3.4	Exemplo 3.5
Abordagem 4 (s=1)	-1.1024	-2.9406	-15.6575
Abordagem 4 (s=2)	-0.8766	-2.9769	-11.8649
Abordagem 4 (s=3)	-1.0617	-2.9771	-15.6573
Abordagem 4 (s=4)			-14.0170
Abordagem 4 (s=5)			-15.6571
Abordagem 1	-1.0967	-2.5984	-16.6820
Melhor solução	$\mathbf{-0.8730}$	$\mathbf{-2.1110}$	$\mathbf{-9.9779}$

Tabela 4.2: Comparações por diferentes dualizações na Abordagem 4.

n=3 (limite superior = $-\mathbf{1.9281}$)		
Abordagem	limite inferior	tempo
4	−2.7206	0.3464
3	−2.7252	0.1692
n=5 (limite superior = $-\mathbf{4.4185}$)		
Abordagem	limite inferior	tempo
4	−6.5648	0.5626
3	−6.9968	0.2781
n=10 (limite superior =$-\mathbf{12.3386}$)		
Abordagem	limite inferior	tempo
4	−22.4734	1.1547
3	−22.5972	0.5651
n=20 (limite superior = $-\mathbf{36.8650}$)		
Abordagem	limite inferior	tempo
4	−82.3516	2.4984
3	−84.3078	1.1823
n=50 (limite superior = $-\mathbf{155.2720}$)		
Abordagem	limite inferior	tempo
4	−496.6177	7.7927
3	−502.7512	3.6473
n=100 (limite superior = $-\mathbf{454.8378}$)		
Abordagem	limite inferior	tempo
4	−1966.2755	36.3679
3	−1976.1566	20.2721

Tabela 4.3: Comparações entre as Abordagens 3 e 4.

n=3 (limite superior = $-\mathbf{1.9281}$)		
Abordagem	limite inferior	tempo
1	-2.4723	0.0023
2	-5.8411	0.0054
3	-2.7252	0.1692
4(s=1)	-2.6084	3.3866
4(s_{best})	-2.5236	10.4828
n=5 (limite superior = $-\mathbf{4.4185}$)		
Abordagem	limite inferior	tempo
1	-5.0556	0.0010
2	-12.6322	0.0010
3	-6.9968	0.2781
4(s=1)	-6.4626	5.8833
4(s_{best})	-6.2582	28.8986
n=10 (limite superior = $-\mathbf{12.3386}$)		
Abordagem	limite inferior	tempo
1	-13.4735	0.0017
2	-32.6065	0.0020
3	-22.5972	0.5651
4 (s=1)	-21.5368	11.8513
4(s_{best})	-21.1035	117.4275
n=20 (limite superior = $-\mathbf{36.8650}$)		
Abordagem	limite inferior	tempo
1	-38.7062	0.0066
2	-86.9546	0.0066
3	-84.3078	1.1823
4 (s=1)	-81.8364	24.2069
4(s_{best})	-81.2475	478.3256

Tabela 4.4: Comparações entre todas as abordagens

Capítulo 5

Conclusões e trabalhos futuros

O principal objetivo da tese foi desenvolver e comparar várias técnicas de relaxação para problemas de otimização esférico cúbico restritos à esfera.

A primeira técnica analizada lineariza tanto a função objetivo como a função que define a restrição (funções cúbica e quadrática respectivamente), esta relaxação do conjunto viável leva a adicionar novas variáveis e restrições à formulação do problema original (problemas PL(PEC) e PL_g(PEC)). Os resultados mostrados indicam que ambas relaxações PL(PEC) e PL_g(PEC) aumentam consideravelmente o tamanho do problema original, porém uma vantagem de ambas relaxações é que os problemas resultantes são lineares e na literatura existem diversos softwares que resolvem problemas lineares de grande porte, de forma eficiente e em poucos segundos. Além disso, o aumento de tamanho do problema original é devido ao fato que as desigualdades que definem as relaxações estão considerando variáveis que podem não estar definidas no problema original, por isso como trabalhos futuros, podemos desenvolver uma análise mais profunda sobre qual grupo de desigualdades válidas devemos considerar para as relaxações PL(PEC) e PL_g(PEC) e desta forma reduzir o tamanho dos problemas relaxados. Os resultados apresentados mostram que utilizando as restrições *bound-grid-factor* na relaxação linear obtemos melhores limitantes inferiores, porém também indicam que a medida que a matriz dos coeficientes da função objetivo é mais densa esses limitantes se afastam cada vez mais do limite superior conhecido. Em recentes investigações têm-se introduzido novas técnicas de relaxações, as quais acrescentam outras desigualdades válidas à formulação do problema relaxado, por exemplo, desigualdades válidas baseadas em programação semidefinida, e/ou cortes disjuntivos, as quais passarão a ser tema de trabalhos de pesquisa futura.

A segunda técnica desenvolvida utiliza quatro diferentes abordagens de decomposição para a função objetivo dos problemas PEC e PC. Diferentemente da primeira técnica desenvolvemos abordagens que decompõe as funções objetivo sem modificar a região viável com a intenção de obter melhores limites inferiores. Foram desenvolvidos experimentos numéricos com instâncias geradas aleatoriamente com a finalidade de comparar os limites inferiores obtidos pelas diferentes abordagens. Os resultados computacionais mostrados reportam os valores médios para 20 instâncias com diferentes tamanhos. Corroboramos que todas as abordagens são tratáveis no sentido de envolverem a solução de subproblemas com complexidade polinomial. Os resultados mostrados indicam que a abordagem via autovalores (Abordagem 1), obtem melhores limites inferiores a um custo computacional muito baixo, provávelmente pela forma da decomposição da função objetivo a qual aproveita resultados de programação quadrática.

Vale ressaltar que as Abordagens 3 e 4 ainda poderão ser melhor desenvolvidas fortalecendo a análise nos seguintes aspectos:

1. Para a Abordagem 3, a desigualdade

$$\min_{x_i} \max_{\alpha} f(x_i, \alpha) \geq \max_{\alpha} \min_{x_i} f(x_i, \alpha)$$

é fundamental. Nossa primeira tentativa para nos aproximar a um bom limite inferior do problema do lado direito da desigualdade, foi fixar os valores de α e minimizar com respeito a x_i, ou seja, discretizar o problema dual com o objetivo de nos aproximarmos da desigualdade acima. Desenvolver uma melhor análise sobre este aspecto introduzindo algoritmos aproximados ou métodos heurísticos para nos aproximarmos da igualdade nesta desigualdade é parte de nossa pesquisa futura.

2. Para a Abordagem 4, os valores dos multiplicadores de Lagrange μ^k são muito importantes para obter um limite inferior mais apertado. A Abordagem 4 pode tornar-se mais prática, através de uma melhor escolha dos multiplicadores de Lagrange, e atualizá-los dentre do algoritmo de subgradiente. A melhoria nesses cálculos e o desenvolvimento de uma heurística para atualizar os multiplicadores de Lagrange μ^k também fazem parte de nossa pesquisa futura.

Referências Bibliográficas

[1] BUCHHEIM, C., SANTIS, M. D., PALAGI, L., et al. "An Exact Algorithm for Nonconvex Quadratic Integer Minimization Using Ellipsoidal Relaxations", *SIAM J. Optim.*, v. 23, n. 3, pp. 1867–1889, jun. 2013.

[2] ZHANG, X., QI, L., YE, Y. "The cubic spherical optimization problems", *Mathematics of Computation*, v. 81, n. 279, pp. 1513–1525, jul. 2012.

[3] NESTEROV, Y. E. "Random walk in a simplex and quadratic optimization over convex polytopes", *CORE Discussion Paper 2003/71*, v. 71, pp. CORE–UCL, 2003.

[4] ZHANG, T., GOLUB, H. "Rank-1 approximation of higher-order tensors", *SIAM Journal on Matrix Analysis and Applications*, v. 23, n. 2, pp. 534–550, nov. 2001.

[5] KOFIDIS, E., REGALIA, P. A. "On the best rank-1 approximation of higher-order supersymmetric tensors", *SIAM. J. Matrix Anal. and Appl.*, v. 23, n. 3, pp. 863–884, mar. 2002.

[6] BASSER, P. J., MATTIELLO, J., LEBIHAN, D. "MR diffusion tensor spectroscopy and imaging", *Biophys. J.*, v. 66, n. 1, pp. 259–267, jan. 1994.

[7] BASSER, P. J., MATTIELLO, J., LEBIHAN, D. "Estimation of the effective seldiffusion tensor from the NMR spin echo", *J. Magn. Reson. B*, v. 103, n. 3, pp. 247–254, mar. 1994.

[8] BASSER, P. J., JONES, D. K. "Diffusion-tensor MRI: theory, experimental design and data analysis-a technical review", *NMR Biomed.*, v. 15, n. 7, pp. 456–467, 2002.

[9] LIU, C. L., BAMMER, R., ACAR, B., et al. "Characterizing non-gaussian diffusion by using generalized diffusion tensors", *Magn. Reson. Med.*, v. 51, n. 5, pp. 924–937, maio 2004.

[10] NIE, J., WANG, L. “Semidefinite Relaxations for Best Rank-1 Tensor Approximations”, *SIAM J. Matrix Anal. Appl.*, v. 1, pp. 1155–1179, maio 2014.

[11] BUCHHEIM, C., FAMPA, M., SARMIENTO, O. “Tractable Relaxations for the Cubic One-Spherical Optimization Problem”, *Advances in Intelligent Systems and Computing. Springer, Cham*, v. 991, pp. 267–276, jun. 2019.

[12] AHMADI, A. A., MAJUMDAR, A. “Some applications of polynomial optimization in operations research and real-time decision making”, *Optim. Lett*, v. 10, n. 4, pp. 709–729, abr. 2016.

[13] BERMAN, O., JAILLET, P., SIMCHI-LEVI, D. “Location-routing problems with uncertaintly”, *Facility location: a survey of applications and methods*, v. 106, pp. 427–452, 1995.

[14] HORST, R. “Deterministic Methods in constrained global optimization: some recent advances and new fields of application”, *Naval Research Logistic Quartely*, v. 37, n. 4, pp. 433–471, ago. 1990.

[15] HORST, R., TUY, H. *Global optimization: deterministic approaches.* 3 ed. Berlin, Springer-Verlag Berlin Heidelberg, 1996.

[16] SHERALI, H. D., ADAMS, W. P. “A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems”, *SIAM J. Discrete Math.*, v. 3, n. 3, pp. 411–430, set. 1990.

[17] SHERALI, H. D., TUNCBILEK, C. H. “A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique”, *J. Global Optim.*, v. 2, n. 1, pp. 101–112, mar. 1992.

[18] SHERALI, H. D., TUNCBILEK, C. H. “New reformulation linearization / convexification relaxations for univariate and multivariate polynomial programming problems”, *Oper. Res.Lett.*, v. 21, n. 1, pp. 1–9, dez. 1992.

[19] ADAMS, W. P., SHERALI, H. D. “A tight linearization and an algorithm for zero-one quadratic programming problems”, *Management Sci.*, v. 32, n. 10, pp. 1274–1290, out. 1986.

[20] DALKIRAN, E., SHERALI, H. D. “Theorical filtering of RLT bound-factor-constraints for solving polynomial programming problems to global optimality”, *J. Glob. Optim.*, v. 57, n. 4, pp. 1147–1172, dez. 2013.

[21] ANSTREICHER, K. M. "Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming", *J. Glob. Optim.*, v. 43, n. 2, pp. 471–484, mar. 2009.

[22] BURER, S., LETCHFORD, A. N. "On nonconvex quadratic programming with box constraints", *SIAM J. Optim.*, v. 20, n. 2, pp. 1073–1089, jul. 2009.

[23] SHERALI, H. D., DALKIRAN, E. "Combined bound-grid-factor constraints for enhancing RLT relaxations for polynomial programs", *J. Glob. Optim.*, v. 51, n. 3, pp. 377–393, nov. 2011.

[24] HOCK, W., SCHITTKOWSKI, K. *Test Examples for Nonlinear Programming Codes*. Berlin, Lectures Notes in Economics and Mathematical Systems, Springer, 1981.

[25] LASERRE, J. B. "Global optimization with polynomials and the problem of moments", *SIAM J. Control Optim.*, v. 12, n. 3, pp. 796–817, jan. 2001.

[26] LASERRE, J. B. "Semidefinite programming vs LP relaxations for polynomial programming", *Math. Oper. Res.*, v. 27, n. 2, pp. 347–360, maio 2002.

[27] SORENSEN, D. C. "Newton's method with a model trust region modification", *SIAM Journal on Scientific and Statistical Computing*, v. 19, n. 2, pp. 409–427, 1982.

[28] MORÉ, J. J., SORENSEN, D. C. "Computing a trust region step", *SIAM Journal on Scientific and Statistical Computing*, v. 4, n. 3, pp. 553–572, 1983.

[29] SORENSEN, D. C. "Minimization of a large scale quadratic function subject to an ellipsoidal constraint", *SIAM Journal on Optimization*, v. 7, n. 1, pp. 141–161, 1997.

[30] FLETCHER, R. *Practical Methods of Optimization*. 2 ed. New York, USA, John Wiley & Sons, 2000.

[31] FORSYTHE, G. E., GOLUB, G. H. "On the stationary values of a second-degree polynomial on the unit sphere", *SIAM J. Appl. Math.*, v. 13, n. 4, pp. 1050–1068, dez. 1965.

[32] GANDER, W., GOLUB, G. H., MATT, U. V. "A constrained eigenvalue problem", *Linear Algebra and its Applications*, v. 114-115, pp. 815–839, 1989.

[33] GAY, D. M. "Computing optimal locally constrained steps", *SIAM Journal on Scientific and Statistical Computing*, v. 2, n. 2, pp. 186–197, 1981.

[34] STERN, R. J., WOLKOWICZ, H. "Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations", *SIAM J. Optimization*, v. 5, n. 2, pp. 286–313, maio 1995.

[35] LOFBERG, J. "Yalmip: A toolbox for modeling and optimization in Matlab", *2004 IEEE International Symposium on Computer Aided Control Systems Design*, pp. 284–289, set. 2004.

[36] SHERALI, H. D., DALKIRAN, E., DESAI, J. "Enhancing RLT - based relaxations for polynomial programming problems via a new class of v-semidefinite cuts", *Comput. Optim. Appl.*, v. 52, n. 2, pp. 483–506, jun. 2012.

[37] SHERALI, H. D., TUNCBILEK, C. H. "Comparison of two reformulation-linearization technique based linear programming relaxations for polynomial programming problems", *J. Global Optim.*, v. 10, n. 4, pp. 381–390, jun. 1997.

[38] LUCIDI, S., PALAGI, L. "Solution of the Trust Region Problem via a Smooth Unconstrained Reformulation", *Fields Institute Communications*, v. 18, pp. 237–250, fev. 1998.

[39] BADER, B. W., KOLDA, T. G., OTHERS. *MATLAB Tensor Toolbox Version 3.0-dev.* Available online https://www.tensortoolbox.org., 2017.

[40] BEAULIEU, C. "The basis of anisotropic water diffusion in the nervous system a technical review", *NMR Biomed.*, v. 15, n. 7-8, pp. 435–455, dez. 2005.

[41] ERIKSSON, S. H. "Diffusion tensor imaging in patients with epilepsy and malformations of cortical development", *Brain.*, v. 124, n. 3, pp. 617–626, mar. 2001.

[42] BASSER, P. J., PAJEVIC, S. "A normal distribution for tensor valued random variables: applications to diffusion tensor MRI", *IEEE Trans. Med. Imaging.*, v. 22, n. 7, pp. 785–794, jul. 2003.

[43] ZHANG, X., LING, C., QI, L., et al. "The measure of diffusion skewness and kurtosis in magnetic resonance imaging", *Pacific Journal of Optimization*, v. 6, pp. 391–404, set. 2010.

[44] NOWAK, I. *Relaxation and Decomposition Methods for Mixed Integer Nonlinear Programming.* 1 ed. Berlin, Birkhaeuser Basel, 2005.

[45] POLYAK, B. T. "A general method for solving extremum problems", *Soviet Mathematics*, v. 8, pp. 593–597, jan. 1993.

[46] GOFFIN, J. L. "On convergence rate of subgradient optimization methods", *Math. Progr.*, v. 13, n. 1, pp. 329–347, dez. 1977.

[47] POLYAK, B. T. *Introduction to optimization.* 1 ed. New York, USA, Optimization Software, 1987.