UNIVERSIDADE DE SÃO PAULO-USP
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

RICARDO BENTO NOGUEIRA MORI PINHEIRO

**UMA ABORDAGEM PRIMAL-DUAL DE
REESCALAMENTO NÃO-LINEAR INTEGRADO
PARA PROBLEMAS DE PROGRAMAÇÃO
MATEMÁTICA DISCRETA-MISTA COM RESTRIÇÕES
DE EQUILÍBRIO E SUAS APLICAÇÕES AO PROBLEMA
DE FLUXO DE POTÊNCIA ÓTIMO REATIVO**

Ricardo Bento Nogueira Mori Pinheiro

**Uma abordagem primal-dual de
reescalamiento não-linear integrado para
problemas de programação matemática
discreta-mista com restrições de equilíbrio
e suas aplicações ao problema de fluxo de
potência ótimo reativo**

Tese de doutorado apresentada ao Programa de Engenharia Elétrica da Escola de Engenharia de São Carlos como parte dos requisitos para a obtenção do título de Doutor em Ciências.

Área de concentração: Sistemas Elétricos de Potência

Orientador: Prof^o Titular Geraldo Roberto Martins da Costa

Coorientador: Prof^o Dr. Guilherme Guimarães Lage

São Carlos

2017

Trata-se da versão corrigida da tese. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

P654a

Pinheiro, Ricardo Bento Nogueira Mori

Uma abordagem primal-dual de reescalamento não-linear integrado para problemas de programação matemática discreta-mista com restrições de equilíbrio e suas aplicações ao problema de fluxo de potência ótimo reativo / Ricardo Bento Nogueira Mori Pinheiro; orientador Geraldo Roberto Martins da Costa; coorientador Guilherme Guimarães Lage. -- São Carlos, 2017.

Tese (Doutorado) - Programa de Pós-Graduação em Engenharia Elétrica e Área de Concentração em Sistemas Elétricos de Potência. -- Escola de Engenharia de São Carlos da Universidade de São Paulo, 2017.

1. Fluxo de potência ótimo reativo. 2. Método de lagrangiano aumentado. 3. Método primal-dual de reescalamento não-linear integrado. 4. Método de penalidade. I. Título.

FOLHA DE JULGAMENTO

Candidato: Licenciado **RICARDO BENTO NOGUEIRA MORI PINHEIRO**.

Título da tese: "Uma abordagem primal-dual de reescalamiento não-linear integrado para problemas de programação matemática discreta-mista com restrições de equilíbrio e suas aplicações ao problema de fluxo de potência ótimo reativo".

Data da defesa: 03/05/2017.

Comissão Julgadora:

Resultado:

Prof. Titular **Geraldo Roberto Martins da Costa**
(Orientador)
(Escola de Engenharia de São Carlos/EESC)

aprovado

Prof. Dr. **Guilherme Guimarães Lage**
(Co-Orientador)
(Universidade Federal de São Carlos/UFSCar)

Aprovado

Profa. Dra. **Edméa Cássia Baptista**
(Universidade Estadual Paulista "Júlio de Mesquita Filho"/UNESP – Bauru)

APROVADO

Prof. Dr. **Leonardo Nepomuceno**
(Universidade Estadual Paulista "Júlio de Mesquita Filho"/UNESP – Bauru)

APROVADO

Prof. Dr. **Aurélio Ribeiro Leite de Oliveira**
(Universidade Estadual de Campinas/UNICAMP)

APROVADO

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Associado **Luís Fernando Costa Alberto**

Presidente da Comissão de Pós-Graduação:
Prof. Associado **Luís Fernando Costa Alberto**

Dedico este trabalho ao caprichoso Criador e querido amigo Jesus Cristo. Este mesmo Jesus a quem Deus, o Pai, refere-se da seguinte maneira: “Ó Deus, o teu trono subsiste pelos séculos dos séculos; Cetro de equidade é o cetro do teu reino. Amaste a justiça e odiaste a iniquidade; por isso Deus, o teu Deus, te ungiu com óleo de alegria mais do que a teus companheiros”. E também do seguinte modo: “Tu, Senhor, no princípio fundaste a terra, e os céus são obra de tuas mãos.”

Agradecimentos

Registro neste trabalho meus agradecimentos a Deus por toda sorte de Graça e capacitação em benefício do desenvolvimento deste trabalho que vem a cooperar com muitas pessoas. Agradeço ao professor Geraldo pela oportunidade que me concebeu em continuar meus estudos em um projeto de tese elegantemente difícil e extremamente motivador e também por compartilharmos conhecimentos e experiências acadêmicas . Ao professor Guilherme que foi o primeiro a me ensinar os primeiros conhecimentos sobre as restrições de complementaridade e por toda sorte de excelentes consequência após aquele dia recebi uma cópia de sua tese em minhas mãos. Agradeço ao pessoal do laboratório da EESC pela agradável companhia. Sou grato aos professores Balbo, Edméa e Leonardo por todo suporte psicológico, conhecimento, encorajamento e pelas diversas e generosas sugestões feitas sobre este trabalho. Agradeço à Marcela, a quem eu escolhi para ser minha esposa e complementar minha vida. Registro neste texto minha gratidão e respeito por sua compreensão e paciência comigo para que eu pudesse mais uma vez honrar o saudoso nome do meu pai e dar alegria a minha mãe por tudo que eles fizeram em mim. Obrigado por estar ao meu lado e por correr junto a mim em todo o tempo.

“Nele [em Cristo] estão escondidos todos os tesouros da sabedoria e do conhecimento.”
Colossenses 2:3

Resumo

Pinheiro, R. B. N. M. **Uma abordagem primal-dual de reescalamento não-linear integrado para problemas de programação matemática discreta-mista com restrições de equilíbrio e suas aplicações ao problema de fluxo de potência ótimo reativo.** 229 p. Tese de doutorado – Escola de Engenharia de São Carlos, Universidade de São Paulo, 2017.

Neste trabalho propomos uma abordagem computacional especificamente talhada para a solução de problemas de programação matemática discreta-mista com restrições de equilíbrio (MPEC). Para isso, inicialmente, transformamos o MPEC discreto-misto em uma sequência de MPECs contínuos. Na formulação dos MPECs contínuos, inserimos restrições de igualdade e de desigualdades artificiais, as quais nos permitem considerar as variáveis discretas como contínuas. Cada MPEC contínuo é transformado em um problema de programação não-linear (PNL) padrão. Isso é feito por meio da reformulação das restrições de complementaridade originais do MPEC contínuo em um conjunto equivalente de restrições usuais de desigualdade. As restrições de igualdade originais do PNL são tratadas por meio da função lagrangiana clássica, as restrições de igualdade artificiais associada às variáveis discretas do PNL são tratadas por meio de uma técnica variante do método de penalidades clássico e as restrições de desigualdade artificiais e originais do problema são tratadas por meio do método de reescalamento não-linear integrado proposto neste trabalho. Cada PNL é resolvido por meio de uma abordagem primal-dual do método de reescalamento não-linear integrado (PDRNLI) com atualização dinâmica dos parâmetros e com a estratégia de convergência global proposta. O método PDRNLI é aplicado ao problema de fluxo de potência ótimo reativo com restrições de atuação de dispositivo de controle de tensão associado aos sistemas elétricos IEEE-14, IEEE-30 e IEEE-118 barras. Os resultados numéricos comprovam a eficiência do método PDRNLI proposto para a solução do problema.

Palavras-chave: Fluxo de Potência Ótimo Reativo. Método de Lagrangiano Aumentado. Método Primal-Dual de Reescalamento Não-Linear Integrado. Método de Penalidade.

Abstract

Pinheiro, R. B. N. M. **A primal-dual integrated nonlinear rescaling approach for mixed-discrete mathematical problems with equilibrium constraints and its application to the reactive optimal power flow problems..** 229 p. Ph.D. Thesis – São Carlos School of Engineering, University of São Paulo, 2017.

In this work we propose a computational approach specifically tailored for solving mixed-discrete mathematical problems with equilibrium constraints (MPEC). For such a purpose, we initially transform the mixed-discrete MPEC problem into a sequence of continuous MPEC problems. In the formulation of the continuous MPECs, we insert artificial equality and inequality constraints, which allow us handling discrete variables as continuous ones. Each continuous MPEC is transformed into a standard nonlinear programming problem (NLP). This is performed by reformulating the original complementarity constraints of the continuous MPEC problems into an equivalent system of standard inequality constraints. The original equality constraints of the NLP problem are handled by means of the classical lagrangian function, while the artificial equality constraints associated with the discrete variables are handled by means of a variant of the classic penalty method. The original and artificial inequality constraints are handled by means of the integrated nonlinear rescaling method proposed in this work. Each NLP is solved by means of a primal-dual version of the integrated nonlinear rescaling approach (PDINLR), with dynamic updating of parameters together with proposed a global convergence strategy. The PDINLR method is applied to the reactive optimal power flow problem with additional constraints associated with the actuation of voltage control devices for the associated with IEEE-14, 30 and 118 bus electrical systems. Numerical results assure the efficiency of the method PDINLR proposed for solving the problem.

Keywords: Reactive Optimal Power Flow. Augmented Lagrangian Methods. Primal-Dual Integrated Nonlinear Rescaling Methods. Penalty methods.

Lista de ilustrações

Figura 2.2.1	<i>Interpretação geométrica da medida de distância de Bregman induzida por ν.</i>	35
Figura 2.7.1	<i>A função logarítmica-quadrática (azul) em perspectiva com a função $\varphi(s) = -\ln(s) + s - 1$.</i>	52
Figura 2.7.2	<i>Conjugado convexo da função logarítmica-quadrática (vermelho) em perspectiva com a função barreira logarítmica modificada.</i>	54
Figura 3.2.1	<i>Aplicação da função interpolante trigonométrica definida por partes aos conjuntos $\mathbb{D}_1 = \{0, 1\}$, $\mathbb{D}_2 = \{0, 2\}$, $\mathbb{D}_3 = \{0, 4\}$, $\mathbb{D}_4 = \{0, 6\}$, $\mathbb{D}_5 = \{0, 8\}$ e $\mathbb{D}_6 = \{0, 10\}$.</i>	60
Figura 3.2.2	<i>Aplicação da função interpolante hiperbólica definida por partes aos conjuntos $\mathbb{D}_1 = \{0, 1\}$, $\mathbb{D}_2 = \{0, 2\}$, $\mathbb{D}_3 = \{0, 4\}$, $\mathbb{D}_4 = \{0, 6\}$, $\mathbb{D}_5 = \{0, 8\}$ e $\mathbb{D}_6 = \{0, 10\}$.</i>	60
Figura 3.2.3	<i>Derivada da função interpolante trigonométrica definida por partes ao longo dos intervalos $\mathbb{I}_1 = [0, 1]$, $\mathbb{I}_2 = [0, 2]$, $\mathbb{I}_3 = [0, 4]$, $\mathbb{I}_4 = [0, 6]$, $\mathbb{I}_5 = [0, 8]$ e $\mathbb{I}_6 = [0, 10]$.</i>	61
Figura 3.2.4	<i>Derivada da função interpolante hiperbólica definida por partes ao longo dos intervalos $\mathbb{I}_1 = [0, 1]$, $\mathbb{I}_2 = [0, 2]$, $\mathbb{I}_3 = [0, 4]$, $\mathbb{I}_4 = [0, 6]$, $\mathbb{I}_5 = [0, 8]$ e $\mathbb{I}_6 = [0, 10]$.</i>	61
Figura 3.3.1	<i>A função $q(y) = \ln(\cosh(y))$ e as suas respectivas funções de regularização $q_{\inf}$ e q_2.</i>	64
Figura 3.3.2	<i>As derivadas de primeira ordem de q e de suas respectivas funções de regularização.</i>	64
Figura 3.3.3	<i>As funções penalidades quadrática e barreira γ-logarítmica aplicada à função $\pi^{-1} \sin(\pi x)$.</i>	66
Figura 4.1.1	<i>Representação geométrica de $0 \leq G_s(\mathbf{x}) \perp H_s(\mathbf{x}) \leq 0$.</i>	71
Figura 4.2.1	<i>Representação geométrica de uma restrição de complementaridade mista.</i>	78
Figura 5.1.1	<i>Exemplo de uma restrição de complementaridade mista definida pela relação entre uma variável contínua e discreta.</i>	86
Figura 5.2.1	<i>Representação geométrica do conjunto de restrições $\mathbb{U}$.</i>	87
Figura 5.12.1	<i>Curvas de nível da função objetivo do problema (5.12.1).</i>	106
Figura 5.12.2	<i>Conjunto de soluções viáveis do problema (5.12.1).</i>	107
Figura 5.12.3	<i>A trajetória dos pontos $\mathbf{x}^k$ gerada pelo método PDRNLI para a classe $\mathbb{P}_1$ de funções penalidades para resolver o problema (5.12.1).</i>	109
Figura 5.12.4	<i>A trajetória dos pontos $\mathbf{x}^k$ gerada pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades para resolver o problema (5.12.1).</i>	110

Figura 5.12.5	Trajatória dos pontos $\mathbf{x}^k$ gerados pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades em função da escolha de α	112
Figura 5.12.6	A trajetória dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ sem a estratégia de convergência global.	114
Figura 5.12.7	A trajetória dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ com a estratégia de convergência global.	114
Figura 5.12.8	As trajetórias dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI em função da escolha de α sem a estratégia de convergência global.	117
Figura 5.12.9	As trajetórias dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI em função da escolha de α com a estratégia de convergência global.	117
Figura 5.12.10	Curvas de nível da função objetivo e região viável do problema (5.12.3)	118
Figura 5.12.11A	trajetória dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades para resolvermos o problema (5.12.3).	120
Figura 5.12.12	Região viável e representação geométrica das restrições de complementaridade mista do problema (5.12.5).	122
Figura 5.12.13A	trajetória dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades para resolvermos o problema (5.12.6).	123
Figura 5.12.14A	trajetória dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades para resolvermos o problema (5.12.6).	126
Figura 6.1.1	Modelo de circuito adotado para o transformador em-fase.	132
Figura 7.2.1	Topologia do sistema IEEE-14 barras [Power 1999].	136
Figura 7.2.2	Topologia do sistema IEEE-30 barras [Power 1999].	137
Figura 7.2.3	Topologia do sistema IEEE-118 barras [Power 1999].	138
Figura 7.3.1	Valores iterativos das magnitudes de tensão obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.	142
Figura 7.3.2	Configuração final das magnitudes de tensão nas barras do sistema elétrico IEEE-14 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.	143
Figura 7.3.3	Os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.6).	145
Figura 7.3.4	Os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.1).	146
Figura 7.3.5	Valores iterativos dos taps dos transformadores em-fase obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.	147
Figura 7.3.6	Os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$, vinculados às restrições $V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{\text{sh}} - b_{k_{\text{spec}}}^{\text{sh}})$, gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.6).	149
Figura 7.3.7	Os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$, vinculados às restrições $V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{\text{sh}} - b_{k_{\text{spec}}}^{\text{sh}})$, gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.1).	150

Figura 7.3.8	Valores iterativos da susceptância capacitiva de barra obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.	151
Figura 7.3.9	Valores iterativos vinculados às componentes do vetor de parâmetros de penalidade ρ^k obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.1).	152
Figura 7.3.10	Valores iterativos das magnitudes de tensão obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.	157
Figura 7.3.11	Configuração final das magnitudes de tensão nas barras do sistema elétrico IEEE-30 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.	158
Figura 7.3.12	Os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.6).	160
Figura 7.3.13	Os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.1).	161
Figura 7.3.14	Valores iterativos dos taps dos transformadores em-fase obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.	162
Figura 7.3.15	Os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$, vinculados às restrições $V_k^{\min} \leq V_k \leq V_k^{\max} \pm (b_k^{\text{sh}} - b_{k_{\text{espec}}}^{\text{sh}})$, gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.6).	164
Figura 7.3.16	Os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$, vinculados às restrições $V_k^{\min} \leq V_k \leq V_k^{\max} \pm (b_k^{\text{sh}} - b_{k_{\text{espec}}}^{\text{sh}})$, gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.1).	165
Figura 7.3.17	Valores iterativos das susceptância shunt de barra obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.	166
Figura 7.3.18	Valores iterativos vinculados às componentes do vetor de parâmetros de penalidade ρ^k obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.1).	167
Figura 7.3.19	Valores iterativos das magnitudes de tensão obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.	174
Figura 7.3.20	Configuração final das magnitudes de tensão nas barras do sistema elétrico IEEE-118 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.	175
Figura 7.3.21	Os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante o modelo (6.1.6).	178
Figura 7.3.22	Os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI para o sistema IEEE-118 por meio do modelo de FPOR (6.1.1).	179
Figura 7.3.23	Valores iterativos dos taps dos transformadores em-fase obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.	181

Figura 7.3.24 Os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$, vinculados às restrições $V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{\text{sh}} - b_{k_{\text{espec}}}^{\text{sh}})$, gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante o modelo (6.1.6).	184
Figura 7.3.25 Os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$, vinculados às restrições $V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{\text{sh}} - b_{k_{\text{espec}}}^{\text{sh}})$, gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante o modelo (6.1.1).	185
Figura 7.3.26 Valores iterativos das susceptância shunt de barra obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.	186
Figura 7.3.27 Valores iterativos vinculados às componentes do vetor de parâmetros de penalidade ρ^k obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante o modelo (6.1.1).	188
Figura A.1.1 O vetor $\mathbf{w}^\perp \in \mathbb{W}^\perp$	206
Figura A.1.2 O vetor $\mathbf{v}$ é resultante da soma de dois vetores: um $\mathbf{w}_k \in \mathbb{W}$ e outro $\mathbf{w}_k^\perp \in \mathbb{W}^\perp$	206
Figura A.1.3 Ilustração da Definição A.1.3.	208
Figura A.1.4 Ilustração para a Proposição A.1.6.	209
Figura A.2.1 A reta r que está o mais próximo possível de passar pelos pontos de $\mathbb{Y}$	210
Figura A.2.2 O efeito das matrizes de projeção ortogonal em subespaços do $\mathbb{R}^m$	213
Figura A.3.1 Projeções oblíquas e ortogonais de $\mathbf{v}_1$ em subespaços do $\mathbb{R}^2$	218
Figura B.3.1 Representação geométrica de um Pseudo-Espaço Euclidiano em $\mathbb{R}^2$	229

Lista de tabelas

Tabela 2.6.1	<i>Funções penalidades do tipo barreira logarítmica modificada integrada (IMBF).</i>	46
Tabela 2.6.2	<i>Funções penalidades do tipo exponencial integrada.</i>	46
Tabela 5.12.1	<i>Dados de entrada para resolvermos o problema (5.12.1).</i>	107
Tabela 5.12.2	<i>Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_1$ de funções penalidades aplicado ao problema (5.12.1).</i>	108
Tabela 5.12.3	<i>As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_1$ de funções penalidades aplicado ao problema (5.12.1).</i>	109
Tabela 5.12.4	<i>Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades aplicado ao problema (5.12.1).</i>	110
Tabela 5.12.5	<i>As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades aplicado ao problema (5.12.1).</i>	111
Tabela 5.12.6	<i>Influência do parâmetro α sobre o desempenho do método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades aplicado ao problema (5.12.1).</i>	111
Tabela 5.12.7	<i>Dados de entrada para resolvermos o problema (5.12.2).</i>	113
Tabela 5.12.8	<i>Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_{int}$ sem aplicação da estratégia de convergência global.</i>	113
Tabela 5.12.9	<i>Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_{int}$ com a estratégia de convergência global.</i>	113
Tabela 5.12.10	<i>As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ sem a estratégia de convergência global.</i>	115
Tabela 5.12.11	<i>As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ com a estratégia de convergência global.</i>	115
Tabela 5.12.12	<i>Influência do parâmetro α sobre o desempenho do método PDRNLI para a classe $\mathbb{P}_{int}$ sem a estratégia de convergência global.</i>	116
Tabela 5.12.13	<i>Influência do parâmetro α sobre o desempenho do método PDRNLI para a classe $\mathbb{P}_{int}$ com a estratégia de convergência global.</i>	116
Tabela 5.12.14	<i>Dados de entrada para resolvermos o problema (5.12.3).</i>	119
Tabela 5.12.15	<i>Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades para resolvermos o problema (5.12.3).</i>	119

Tabela 5.12.16	As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades para resolvermos o problema (5.12.3).	120
Tabela 5.12.17	Influência do parâmetro α sobre o desempenho do método PDRNLI para a classe $\mathbb{P}_{int}$ sem a estratégia de convergência global.	121
Tabela 5.12.18	Dados de entrada para resolvermos o problema (5.12.6).	122
Tabela 5.12.19	Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades para resolvermos o problema (5.12.6).	123
Tabela 5.12.20	As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades para resolvermos o problema (5.12.6).	124
Tabela 5.12.21	As sequência de multiplicadores de Lagrange $\{\pi_{1_k}\}$ e $\{\pi_{2_k}\}$ associadas à restrição de complementaridade mista do problema (5.12.5) determinadas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades.	125
Tabela 5.12.22	Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades para resolvermos o problema (5.12.6).	125
Tabela 5.12.23	As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades para resolvermos o problema (5.12.6).	126
Tabela 5.12.24	As sequência de multiplicadores de Lagrange $\{\pi_{1_k}\}$ e $\{\pi_{2_k}\}$ associadas à restrição de complementaridade mista do problema (5.12.5) determinadas pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades.	127
Tabela 5.12.25	Influência do parâmetro α sobre o desempenho do método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades.	128
Tabela 7.2.1	Características técnicas do sistema IEEE-14 barras.	136
Tabela 7.2.2	Características técnicas do sistema IEEE-30 barras.	137
Tabela 7.2.3	Características técnicas do sistema IEEE-118 barras.	137
Tabela 7.3.1	Dados de entrada para resolvermos problema de FPOR para o sistema IEEE-14.	139
Tabela 7.3.2	Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-14 por meio do modelo (6.1.7).	139
Tabela 7.3.3	Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-14 por meio do modelo (6.1.6).	140
Tabela 7.3.4	Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-14 por meio do modelo (6.1.1).	140
Tabela 7.3.5	Geração de potência reativa, em MVAr, obtida pelo método PDRNLI para o sistema IEEE-14.	140
Tabela 7.3.6	As magnitudes de tensão nas barras de tensão controlada determinadas pelo método PDRNLI para o sistema elétrico IEEE-14.	141
Tabela 7.3.7	Os taps dos transformadores em-fase determinados pelo método PDRNLI para o sistema elétrico IEEE-14.	144
Tabela 7.3.8	Os multiplicadores de Lagrange π_1^{tap} e π_2^{tap} vinculados à V_m e à $-(t_{km} - t_{km\text{espec}})$, respectivamente, determinados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14.	145

Tabela 7.3.9 As susceptâncias shunt de barra determinadas pelo método PDRNLI para o sistema elétrico IEEE-14.	148
Tabela 7.3.10 Os multiplicadores de Lagrange π_1^{sh} e π_2^{sh} e vinculados à V_k e à $(b_k^{sh} - b_{k_{espec}}^{sh})$, respectivamente, determinados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14.	148
Tabela 7.3.11 Valor final do parâmetro de penalidade definido pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.1).	152
Tabela 7.3.12 Monitoramento do condicionamento numérico da matriz hessiana do método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14.	153
Tabela 7.3.13 Comparação entre as proporções das perdas de potência ativa associado ao sistema IEEE-14.	153
Tabela 7.3.14 Dados de entrada para resolvermos problema de FPOR para o sistema IEEE-30.	154
Tabela 7.3.15 Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-30 por meio do modelo (6.1.7).	155
Tabela 7.3.16 Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-30 por meio do modelo (6.1.6).	155
Tabela 7.3.17 Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-30 por meio do modelo (6.1.1).	156
Tabela 7.3.18 Geração de potência reativa, em MVar, obtida pelo método PDRNLI para o sistema IEEE-30.	156
Tabela 7.3.19 As magnitudes de tensão nas barras de tensão controlada determinadas pelo método PDRNLI para o sistema elétrico IEEE-30.	156
Tabela 7.3.20 Os taps dos transformadores em-fase determinados pelo método PDRNLI para o sistema elétrico IEEE-30.	159
Tabela 7.3.21 Os multiplicadores de Lagrange π_1^{tap} e π_2^{tap} vinculados à V_m e à $-(t_{km} - t_{km_{espec}})$, respectivamente, determinados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30.	159
Tabela 7.3.22 As susceptâncias shunt de barra determinadas pelo método PDRNLI para o sistema elétrico IEEE-30.	163
Tabela 7.3.23 Os multiplicadores de Lagrange π_1^{sh} e π_2^{sh} e vinculados à V_k e à $(b_k^{sh} - b_{k_{espec}}^{sh})$, respectivamente, determinados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30.	163
Tabela 7.3.24 Valor final do parâmetro de penalidade definido pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.1).	167
Tabela 7.3.25 Monitoramento do condicionamento numérico da matriz hessiana do método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30.	168
Tabela 7.3.26 Comparação entre as proporções das perdas de potência ativa associado ao sistema IEEE-30.	168
Tabela 7.3.27 Dados de entrada para resolvermos problema de FPOR para o sistema IEEE-118.	169
Tabela 7.3.28 Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-118 por meio do modelo (6.1.7).	170
Tabela 7.3.29 Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-118 por meio do modelo (6.1.6).	170

Tabela 7.3.30	Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-118 por meio do modelo (6.1.1).	171
Tabela 7.3.31	Geração de potência reativa, em MVar, obtida pelo método PDRNLI para o sistema IEEE-118.	172
Tabela 7.3.32	As magnitudes de tensão nas barras de tensão controlada determinadas pelo método PDRNLI para o sistema elétrico IEEE-118.	173
Tabela 7.3.33	Os taps dos transformadores em-fase determinados pelo método PDRNLI para o sistema elétrico IEEE-118.	176
Tabela 7.3.34	Os multiplicadores de Lagrange π_1^{tap} e π_2^{tap} vinculados à V_m e à $-(t_{km} - t_{km_{\text{espec}}})$, respectivamente, determinados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118.	177
Tabela 7.3.35	As susceptâncias shunt de barra determinadas pelo método PDRNLI para o sistema elétrico IEEE-118.	182
Tabela 7.3.36	Os multiplicadores de Lagrange π_1^{sh} e π_2^{sh} e vinculados à V_k e à $(b_k^{\text{sh}} - b_{k_{\text{espec}}}^{\text{sh}})$, respectivamente, determinados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118.	183
Tabela 7.3.37	Valores iterativos vinculados às componentes do vetor de parâmetros de penalidade ρ^k obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante o modelo (6.1.1).	187
Tabela 7.3.38	Monitoramento do condicionamento numérico da matriz hessiana do método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118.	189
Tabela 7.3.39	Comparação entre as proporções das perdas de potência ativa associado ao sistema IEEE-118.	189

Sumário

1	Introdução	27
1.1	Motivações	28
1.1.1	O tratamento das variáveis discretas	29
1.1.2	O tratamento das restrições de complementaridade	29
1.2	Objetivos gerais e específicos	30
1.2.1	Tratamento das variáveis discretas	30
1.2.2	Tratamento das restrições de desigualdade e sistema de complementaridade	31
1.3	Organização deste trabalho	32
2	O método de reescalamento não-linear integrado	33
2.1	Introdução ao método de ponto proximal	34
2.2	Distâncias de Bregman e entrópicas φ -divergente de Csiszar	35
2.3	Medida de distância integrada	38
2.4	Declaração do problema e premissas básicas	40
2.5	Método dual de ponto proximal	41
2.6	Método de Reescalamento não-linear Integrado	43
2.6.1	A função lagrangiana aumentada integrada e a classe $\mathbb{P}_{\text{int}}$ de funções penalidades	44
2.6.2	O método RNLI e suas propriedades	47
2.7	Extensão: O método da função logarítmica-quadrática integrada	51
3	Método de penalidades aplicado à programação discreta	55
3.1	Declaração do problema e premissas básicas	56
3.2	Transformação de um problema discreto em um contínuo equivalente	56
3.3	Método de Penalidades	61
3.3.1	Funções penalidades e extensão	62
3.3.2	Descrição do método de penalidades	66
4	Problemas de programação matemática com restrições de equilíbrio	69
4.1	Programação matemática com restrições de complementaridade	69

4.1.1	Declaração do problema e outros conceitos	70
4.1.2	Condições de estacionaridade: definições e formulações	73
4.2	Restrições de complementaridade mista	76
5	O método primal-dual de reescalamento não-linear integrado para resolver problemas de programação discreta-mista com restrições de equilíbrio	83
5.1	O problema de programação discreta-mista com restrições de equilíbrio	85
5.2	O MPEC discreto-misto como um MPEC formulado em variáveis contínuas	86
5.3	Transformação do MPEC com variáveis contínuas em um problema de programação não-linear	88
5.4	O método PDRNLI: abordagem primal-dual	89
5.5	O método PDRNLI: dedução do vetor de direções de busca primal-dual	90
5.6	O método PDRNLI: uma abordagem predictor-corretor	95
5.7	O método PDRNLI como um método de gradiente projetado	95
5.8	O método PDRNLI: A estratégia de convergência global denominada condição de positividade	98
5.9	O método PDRNLI: Cálculo do tamanho do passo e atualização dos parâmetros	103
5.10	Inicialização do método PDRNLI	104
5.11	Algoritmo computacional do método PDRNLI	104
5.12	Aplicação do método PDRNLI em exemplos matemáticos	106
5.12.1	Programação não-linear convexo	106
5.12.2	Programação não-linear não-convexo	112
5.12.3	Programação discreta	118
5.12.4	Programação matemática com restrições de equilíbrio	121
6	Fluxo de Potência Ótimo Reativo	129
6.1	Modelos de FPOR	129
6.1.1	FPOR com variáveis de controle discretas e com restrições de atuação de dispositivos de controle de tensão	130
6.1.2	FPOR com variáveis de controle contínuas e com restrições de atuação de dispositivos de controle de tensão	133
6.1.3	FPOR clássico	133
7	Resultados numéricos	135
7.1	Especificações do problema de FPOR	135
7.2	Os sistemas elétricos investigados e suas características	136
7.2.1	Características técnicas do sistema IEEE-14 barras	136
7.2.2	Características técnicas do sistema IEEE-30 barras	137
7.2.3	Características técnicas do sistema IEEE-118 barras	137
7.3	Resultados numéricos	138
7.3.1	Resultados para o sistema IEEE-14 barras	138
7.3.2	Resultados para o sistema IEEE-30 barras	154
7.3.3	Resultados para o sistema IEEE-118 barras	169

8	Conclusões	191
	Referências	195
APÊNDICE A	Uma introdução ao estudo das matrizes de projeção em subespaços	205
A.1	Conceitos Iniciais	205
A.2	Aplicação e propriedades da projeção ortogonal em subespaços do $\mathbb{R}^m$	210
A.3	Projeções ortogonais e oblíquas em subespaços do $\mathbb{R}^n$	214
APÊNDICE B	Regularização, lei de inércia de Sylvester e pseudo-espaço euclidiano	221
B.1	Regularização	221
B.2	Lei de inércia de Sylvester	223
B.3	O pseudo-espaço euclidiano	224

Introdução

O Fluxo de Potência Ótimo (FPO) é um problema de otimização não-linear, não-convexo, restrito, estático, de grande porte, com variáveis contínuas e discretas e com restrições de equilíbrio [Rosehart, Roman e Schellenberg 2005, Platbrood et al. 2014]. O propósito do FPO é determinar o melhor ponto de operação do sistema elétrico através da otimização de uma função objetivo que representa um certo desempenho do sistema elétrico ao mesmo tempo que as restrições de operação são satisfeitas.

Dentre as várias formulações possíveis para o problema de FPO, trabalharemos estritamente com a formulação do Fluxo de Potência Ótimo Reativo (FPOR), a qual utiliza um conjunto de controles relacionados à potência reativa. Na formulação deste problema, supomos que as variáveis de controle vinculadas à potência ativa são fixas. Neste trabalho, o FPOR é formulado com a finalidade de minimização das perdas de potência ativa na transmissão por meio do ajuste ótimo das seguintes variáveis de controle: magnitudes de tensão em barras de geração (contínuas); *taps* dos transformadores em-fase (com formulação discreta) e; susceptâncias *shunt* de barra associadas aos bancos de capacitores e reatores (com formulação discreta).

Em [Lage 2013], o autor nos apresenta um novo modelo de FPOR. Este modelo é formulado como um problema de programação matemática com restrições de equilíbrio (MPEC¹) e envolvendo variáveis discretas e contínuas. De acordo com [Luo, Pang e Ralph 1996, p. xxi], um MPEC é um problema de otimização cujas restrições essenciais são definidas por uma desigualdade variacional paramétrica ou por um sistema de complementaridade.

Um MPEC cujas restrições essenciais são definidas por uma desigualdade variacional² possui a seguinte formulação matemática:

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) \\ \text{s.a :} \quad & \mathbf{F}(\mathbf{x})^T (\mathbf{x} - \bar{\mathbf{x}}) \geq 0 \quad , \quad \forall \bar{\mathbf{x}} \in \mathbb{X} \subseteq \mathbb{R}^n, \end{aligned}$$

em que $\mathbf{x} \in \mathbb{R}^n$ é o vetor de variáveis de decisão, $\mathbb{X}$ é um conjunto não-vazio; $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é a função objetivo e $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é uma função vetorial. Em particular, se $\mathbf{F}(\mathbf{x})$ é identificada como $\nabla g(\mathbf{x})$,

¹Esta sigla é universalmente adotada para abreviarmos o termo em inglês *Mathematical Problems with Equilibrium Constraints*.

²O MPEC formulado com restrições de desigualdade variacional paramétrica pode ser encontrado em [Luo, Pang e Ralph 1996, p. 2] e em [Ye 2005].

então a desigualdade variacional $\nabla g(\mathbf{x})^T (\mathbf{x} - \bar{\mathbf{x}}) \geq 0$, para todo $\bar{\mathbf{x}}$ em $\mathbb{X}$, é interpretada como um problema de encontrarmos um minimizador $\mathbf{x}^*$ em $\mathbb{X}$ da função $g(\mathbf{x})$ definida sobre $\mathbb{X}$.

Por outro lado, um MPEC cujas restrições essenciais são definidas por um sistema de complementaridade possui a seguinte formulação:

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) \\ \text{s.a :} \quad & \mathbf{0} \leq \mathbf{G}(\mathbf{x}) \perp \mathbf{H}(\mathbf{x}) \geq \mathbf{0}, \end{aligned}$$

onde $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{G} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ e $\mathbf{H} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ são funções vetoriais e o símbolo “ $\perp$ ” é um operador que exige *ortogonalidade* entre $\mathbf{G}(\mathbf{x})$ e $\mathbf{H}(\mathbf{x})$.

As restrições de equilíbrio inseridas no modelo proposto em [Lage 2013] são denominadas “restrições de dispositivos de controle de tensão”. Tais restrições são formuladas como um sistema de complementaridade. De acordo com o autor, essas restrições permitem a atuação dos *taps* dos transformadores em-fase e das susceptâncias *shunt* de barra somente quando um dos limitantes impostos para a magnitude de tensão na barra controlada por eles for atingido.

Além do modelo proposto, o autor apresentou uma variante da metodologia apresentada em [Solter, Asada e Costa 2013] para o tratamento das variáveis discretas associadas aos *taps* dos transformadores em-fase e às susceptâncias *shunts* de barra. Em resumo, esta variante nos permite trabalhar com variáveis discretas, cujo ordenamento dos elementos do conjunto discreto associado a cada variável não é identificado como uma progressão aritmética; ou seja, os elementos do conjunto discreto não podem ser organizados de forma igualmente espaçada.

No que tange à metodologia de solução, a complexidade do modelo matemático apresentado em [Lage 2013] traz desafios teóricos e práticos à teoria de otimização. Não encontramos uma abordagem teórica que nos assegure das condições de estacionariedade primal-dual para MPECs formulados em variáveis contínuas e discretas. No aspecto prático, os métodos de otimização convencionais podem falhar nos casos mais simples para a solução de MPECs que envolvem somente variáveis contínuas. Em geral, os métodos convencionais são projetados sob as hipóteses de condição de qualificação de independência linear [Ribeiro e Karas 2013, p. 222] ou a de Mangasarian-Fromovitz [Bertsekas 1999, p. 329], as quais não garantem condições de estacionariedade primal-dual de primeira ordem na presença de restrições de complementaridade em nenhum ponto viável.

Desta forma, neste trabalho, propomos uma metodologia de otimização que resolve o modelo de FPOR proposto em [Lage 2013], cujos resultados numéricos indicam a existência de solução primal-dual para uma sequência de problemas de programação não-linear equivalente que tende para uma solução viável do problema de programação discreta-mista com restrição de equilíbrio original, proposto em [Lage 2013].

1.1 Motivações

Segundo citado anteriormente, o modelo matemático apresentado em [Lage 2013] traz desafios teóricos e práticos à teoria de otimização. Como desafios práticos, citamos duas questões fundamentais relacionadas ao modelo de FPOR que motivaram o desenvolvimento deste trabalho:

- O tratamento das variáveis discretas e;

- O tratamento das restrições de complementaridade.

Estas questões práticas fundamentais são discutidas e sintetizadas nas Subseções 1.1.1 e 1.1.2, a seguir.

1.1.1 O tratamento das variáveis discretas

De um modo geral, as variáveis discretas têm sido tratadas por meio de abordagens baseadas em enumeração implícita ou por meio de métodos de penalidade.

As abordagens de enumeração implícita envolvem uma avaliação sistemática de todas as soluções possíveis, sem, entretanto, avaliar de forma explícita todas elas. Dentre as abordagens de enumeração implícita, destacamos os seguintes métodos: *Branch-and-Bound*, *Branch-and-Cut*, *Branch-and-Reduce*, *Branch-and-Price*, *Branch-and-Cut-and-Price*, *Branch-and-Infer*, *Extended cutting-plane*, *Generalized Benders Decomposition*, *Outer Approximation*, *Logic-based approach*, *Generalized Cross Decomposition*, *Successive Semidefinite Relaxation*, *Lagrangian and Domain Cut*, as quais são discutidas em [Nowak 2005, seção 10.2]. Em geral, estes métodos têm sido desenvolvidos no contexto de programação linear discreta. Todavia, destacamos que poucos trabalhos metodológicos têm sido desenvolvidos no campo da programação não-linear discreta [Rao 2009, p. 588]. Em geral, os problemas de otimização não-linear discreta são abordados por meio de programação linear-discreta sequencial. Assim, optamos por não investir na abordagem de enumeração implícita.

As abordagens baseadas em métodos de penalidade consistem em resolver o problema discreto por meio de um sequência de problemas de programação não-linear em variáveis contínuas que convergem para uma solução do problema discreto original [Shin, Gurdal e Jr 1989, Rao 2009, Soler, Asada e Costa 2013]. Cada problema de programação não-linear desta sequência de problemas é definido da seguinte forma: as variáveis discretas são consideradas como variáveis contínuas e são limitadas a um intervalo real compacto; são adicionados termos de penalidade associados a estas variáveis na função objetivo do problema original, com o intuito de penalizar os valores não-discretos para essas variáveis. A principal vantagem das abordagens com base em métodos de penalidade sobre as abordagens de enumeração implícita é a possibilidade de resolvemos os problemas de programação não-linear por métodos do tipo *gradiente*. Por isso, optamos por utilizar esta abordagem para o tratamento das variáveis discretas no FPOR.

1.1.2 O tratamento das restrições de complementaridade

Os MPECs, cujas restrições essenciais são definidas por um sistema de complementaridade, são geralmente reformulados como um problema de otimização não-linear restrito equivalente. As restrições deste problema equivalente são especificamente formuladas com o intuito de representar a ortogonalidade e a não-negatividade do sistema de complementaridade $\mathbf{0} \leq \mathbf{G}(\mathbf{x}) \perp \mathbf{H}(\mathbf{x}) \geq \mathbf{0}$. Observamos na literatura pelo menos duas maneiras de construir estas restrições. Na primeira, o sistema de complementaridade é representado pelo seguinte conjunto de restrições [Flegel e Kanzow

2005]:

$$\begin{aligned} \mathbf{G}(\mathbf{x})^T \mathbf{H}(\mathbf{x}) &= 0 \\ G_i(\mathbf{x}) &\geq 0, \quad \forall i = 1, \dots, p \\ H_i(\mathbf{x}) &\geq 0, \quad \forall i = 1, \dots, p. \end{aligned} \quad (1.1.1)$$

Na segunda, o sistema de complementaridade é representado por um sistema de equações não-diferenciáveis por meio de Funções-NCP³. Um exemplo destas funções é a função de Fischer-Burmeister [Fischer 1992], a seguir:

$$G_i(\mathbf{x}) + H_i(\mathbf{x}) - \sqrt{G_i(\mathbf{x})^2 + H_i(\mathbf{x})^2} = 0, \quad \forall i = 1, \dots, p. \quad (1.1.2)$$

Uma das principais vantagens de trabalharmos com funções-NCP é que as restrições de desigualdade $\mathbf{G}(\mathbf{x}) \geq \mathbf{0}$ e $\mathbf{H}(\mathbf{x}) \geq \mathbf{0}$ são satisfeitas automaticamente, logo, não é necessário incorporá-las ao conjunto de restrições. No entanto, a principal desvantagem da maioria das funções-NCP é a não-diferenciabilidade dessas funções em um ponto $\bar{\mathbf{x}}$ tal que $G_i(\bar{\mathbf{x}}) = H_i(\bar{\mathbf{x}}) = 0$, para algum $i = 1, \dots, p$; isto é, quando ocorre o que é denominado conjunto bi-ativo. Tendo em vista esta desvantagem (embora esta possa ser tratada eficientemente), nesta tese optamos por trabalhar com a ideia da formulação equivalente (1.1.1) para o sistema de complementaridade.

Uma característica inerente à formulação (1.1.1) é que pelo menos uma restrição de desigualdade $G_i(\mathbf{x}) \geq 0$ e $H_i(\mathbf{x}) \geq 0$ sempre será ativa em um ponto estacionário $\mathbf{x}^*$ qualquer. Isso é devido à presença da restrição $\mathbf{G}(\mathbf{x})^T \mathbf{H}(\mathbf{x}) = 0$. Logo, não existe interior relativo às restrições $G_i(\mathbf{x}) \geq 0$ e $H_i(\mathbf{x}) \geq 0$ em $\mathbf{x}^*$, o que faz com que métodos de pontos interiores, por exemplo, tenham maiores dificuldades computacionais na presença de tais restrições [Benson, Shanno e Vanderbei 2002, Liu e Sun 2004]. Assim, para o tratamento das restrições de complementaridade, investigamos a aplicação dos métodos de lagrangiano aumentado desenvolvidos em [Polyak e Teboulle 1997, Mاتيoli e Gonzaga 2008], os quais são bem definidos na fronteira das restrições de desigualdade.

1.2 Objetivos gerais e específicos

O objetivo geral deste trabalho é o desenvolvimento de uma metodologia para solução de problemas de programação não-linear discreta-mista com restrições de equilíbrio, considerando seus aspectos primal e dual, a qual pode ser utilizada para resolvermos o problema de FPOR proposto em [Lage 2013].

Para o desenvolvimento do objetivo geral, estabelecemos os seguintes objetivos específicos que compõem a metodologia proposta neste trabalho e que visam suprir as motivações discutidas na seção anterior.

1.2.1 Tratamento das variáveis discretas

Neste trabalho, o tratamento das variáveis discretas é inspirado no método de penalidades clássico [Bazaraa, Sherali e Shetty 2006, p. 475]. Para a construção desta abordagem, estabelecemos os seguintes objetivos específicos:

³Esta sigla é universalmente usada para abreviarmos o termo em inglês *Nonlinear Complementarity Functions*.

- 1) Construir classe de funções, cujos *zeros* coincidam com os elementos de um conjunto discreto associado a uma variável;
- 2) Construir um problema de programação não-linear em que as funções da classe descrita em 1) são incorporadas como restrições de igualdade artificiais ao passo que as variáveis discretas são tratadas como contínuas e são limitadas a um intervalo contínuo por meio de restrições canalizadas artificiais;
- 3) Tratar as restrições de igualdade artificiais descritas em 2) pelo método de funções de penalidade e as restrições canalizadas pelo método de lagrangiano aumentado proposto neste trabalho;
- 4) Propor novas funções penalidades de natureza não-quadrática que satisfaçam as propriedades específicas para lidarmos com as restrições de igualdade artificiais descritas em 2).

1.2.2 Tratamento das restrições de desigualdade e sistema de complementaridade

Conforme observado na Subseção 1.1.2, é necessário que tenhamos um maior cuidado para lidar com as restrições de desigualdade associadas ao sistema de complementaridade (1.1.1) do MPEC. Assim, é necessário adotar abordagens capazes de tratar restrições de desigualdade sem interior relativo. Um dos objetivos específicos deste trabalho consiste em investigar dois métodos de lagrangiano aumentado que apresentam boas propriedades na fronteira das restrições de desigualdades ativas, a saber: o método de reescalamiento não-linear proposto em [Polyak e Teboulle 1997] e o método de lagrangiano aumentado proposto em [Matioli e Gonzaga 2008], os quais foram especificamente desenvolvidos para problemas de programação não-linear convexa.

Ambos os métodos de lagrangiano aumentado podem ser interpretados como um método dual de ponto proximal. Um método dual de ponto proximal consiste em gerar uma sequência de pontos duais, a qual é obtida pela maximização da função dual (associada ao problema primal original) adicionada a um termo que exprime a *distância* do ponto dual a ser determinado para um ponto dual viável dado. O método de reescalamiento não-linear descrito em [Polyak e Teboulle 1997] pode ser interpretado como um método dual de ponto proximal com a medida φ -divergente de Csiszar [Csiszár 1967], ao passo que o método descrito em [Matioli e Gonzaga 2008] pode ser interpretado como um método dual de ponto proximal com a medida de distância de Bregman [Bregman 1967].

Ambas as distâncias são induzidas por meio de funções não-quadráticas e convexas cujo domínio é o octante positivo, assim, a viabilidade dual é automaticamente tratada. Outro objetivo específico deste trabalho, consiste em propor um método dual de ponto proximal com uma medida de distância que integra as distâncias de Bregman e de Csiszar por meio de uma função de homotopia. Esta composição tem como objetivo o desenvolvimento de um novo método de lagrangiano aumentado que generaliza as abordagens de [Polyak e Teboulle 1997] e de [Matioli e Gonzaga 2008].

Ao contrário dos problemas de otimização não-linear convencionais, os quais possuem apenas uma condição necessária de estacionariedade de primeira ordem, existem vários conceitos de condições necessárias de estacionariedade de primeira ordem associados aos MPECs [Guo, Lin e Ye 2015].

Finalmente, temos como objetivo específico a caracterização dos pontos estacionários primal e dual dos MPECs investigados neste trabalho, isto é, nosso objetivo específico é classificar tais pontos como pelo menos uma dentre as diversas condições de estacionariedade relacionadas aos MPECs.

1.3 Organização deste trabalho

Este trabalho está organizado da seguinte forma. No Capítulo 2, descrevemos os fundamentos e detalhes do método de reescalamiento não-linear integrado proposto. No Capítulo 3, descrevemos uma variante do método de penalidades clássico para o tratamento das variáveis discretas. No Capítulo 4, descrevemos os problemas de programação matemática com restrições de equilíbrio com ênfase em restrições de complementaridade. No Capítulo 5 desenvolvemos o método primal-dual de reescalamiento não-linear integrado (PDRNLI) para resolvermos problemas de programação matemática discreta-mista com restrições de equilíbrio e suas aplicações em exemplos matemáticos. No Capítulo 6 descrevemos as formulações adotadas para o problema de Fluxo de Potência Ótimo Relativo (FPOR) investigado. No Capítulo 7 apresentamos os resultados numéricos da aplicação do método PDRNLI à solução dos problemas de FPOR. As conclusões do trabalho são resumidas no Capítulo 8. Alguns aspectos matemáticos de interesse para a compreensão do método proposto são apresentados nos Apêndices A e B.

O método de reescalamiento não-linear integrado

Neste capítulo, propomos uma nova abordagem para os métodos de lagrangiano aumentado com funções penalidades não-quadráticas para lidarmos com problemas de programação matemática convexos. Para isto, consideramos primeiramente um problema somente com restrições de desigualdades e seu dual associado e, após isso, tratamos o problema dual por meio de uma abordagem variante do método de ponto proximal clássico. Por definição, o método (dual) de ponto proximal consiste em gerar uma sequência de pontos duais, a qual é obtida pela maximização da função dual adicionada a um termo que exprime a *distância* do ponto dual a ser determinado para um ponto dual viável dado. Vários conceitos de distância têm sido utilizados na literatura. Se a medida de distância é a euclidiana, temos o método proximal clássico de Rockafellar [Rockafellar 1973, Rockafellar 1976]; se a medida de distância é a de Bregman, podemos ter os métodos proximais de Teboulle [Teboulle 1992] ou de Matioli e Gonzaga [Matioli e Gonzaga 2008] e; se a medida de distância é a entrópica φ -divergente de Csiszar, temos o método proximal de Polyak e Teboulle [Polyak e Teboulle 1997].

A escolha adequada de uma classe de funções convexas e não-quadráticas nos permite obter uma formulação matemática semelhante associada à distância de Bregman e à de Csiszar [Teboulle 1992, Matioli e Gonzaga 2008, Polyak 2014]. Esta formulação nos propicia propor uma medida de distância que integra as distâncias de Bregman e de Csiszar por meio de uma função de homotopia, e consequentemente, aplicar este novo conceito de medir ao método dual de ponto proximal. A esta função de homotopia, denominamos medida de distância integrada.

A transformação contínua da medida de distância de Bregman para a de Csiszar e vice-e-versa é realizada por meio de um fator aqui denominado de transformação. Este fator define a forma pela qual a medida de distância integrada se transforma de uma para a outra conforme varia o parâmetro de homotopia. Apresentamos três fatores de transformação e, como resultado, apresentamos três diferentes métodos proximais.

Nosso principal resultado é que o método dual de ponto proximal com a medida de distância integrada nos conduz para um equivalente método de lagrangiano aumentado com uma nova família de funções penalidades não-quadráticas associada ao problema primal. Dentre as novas funções penalidades, apresentamos os conceitos de função barreira logarítmica modificada integrada e de

exponencial integrada. Denominamos o método de lagrangiano aumentado proposto por reescalamiento não-linear integrado (RNLI), pois este possui as características dos métodos de lagrangiano aumentado desenvolvidos em [Polyak e Teboulle 1997, Matioli e Gonzaga 2008].

Este capítulo está organizado da seguinte forma:

- Na Seção 2.1, apresentamos uma breve introdução ao método de ponto proximal em programação convexa;
- Na Seção 2.2, apresentamos o conceito de medidas de distância de Bregman e entrópicas φ -divergente e alguns exemplos;
- Na Seção 2.3, propomos o conceito de medida de distância integrada, a qual mescla ponderadamente as medidas de distância de Bregman e entrópicas φ -divergente de Csiszar;
- Na Seção 2.4, apresentamos o problema de programação não-linear convexo, seus pressupostos, seu problema dual equivalente e alguns resultados clássicos;
- Na Seção 2.5, apresentamos alguns métodos duais de ponto proximal já comentados e o método dual de ponto proximal proposto neste trabalho;
- Apresentamos na Seção 2.6 o desenvolvimento do método de reescalamiento não-linear integrado (RNLI). Apresentamos também uma nova família de funções penalidades para o método de lagrangiano aumentado em questão e alguns exemplos de funções que pertencem a esta família. Dentre as possíveis funções penalidades, destacamos o conceito de funções barreira logarítmica modificada integrada e o de funções exponencial integrada. Por fim, introduzimos uma abordagem primal-dual para o método RNLI;
- Apresentamos na Seção 2.7 o método da função logarítmica-quadrática integrada. Para isso, adaptamos a medida de distância de Auslender à medida de distância integrada proposta. Logo, o método de lagrangiano aumentado é substancialmente diferente do descrito em [Auslender, Teboulle e Ben-Tiba 1999].

2.1 Introdução ao método de ponto proximal

O método de ponto proximal em programação convexa foi proposto por Martinet [Martinet 1978] e tratado amplamente por Rockafellar [Rockafellar 1973, Rockafellar 1976]. O método clássico de ponto proximal consiste em minimizar uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ convexa por meio de uma sequência $\{\mathbf{x}^k\}$ gerada de acordo com (2.1.1):

$$\mathbf{x}^{k+1} = \arg \min \left\{ f(\mathbf{x}) + \mu_k (2)^{-1} \|\mathbf{x} - \mathbf{x}^k\|^2 \right\}, \quad (2.1.1)$$

em que $\{\mu_k\}$ é uma sequência em $\mathbb{R}_{++}$ e $\|\cdot\|$ é a norma euclidiana em $\mathbb{R}^n$ [Teboulle 1992, Iusem, Svaiter e Teboulle 1994]. O algoritmo apresentado em (2.1.1) é generalizado de acordo com (2.1.2):

$$\mathbf{x}^{k+1} = \arg \min \left\{ f(\mathbf{x}) + \mu_k \mathcal{D}(\mathbf{x}, \mathbf{x}^k) \right\}, \quad (2.1.2)$$

em que $\mathcal{D} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ é uma função que representa uma medida de distância do vetor $\mathbf{x}$ para o vetor $\mathbf{x}^k$. A função $\mathcal{D}$ deve possuir pelo menos as propriedades apresentadas em (2.1.3):

$$\begin{cases} \mathcal{D}(\mathbf{a}, \mathbf{b}) = 0 & \Leftrightarrow \mathbf{a} = \mathbf{b} \\ \mathcal{D}(\mathbf{a}, \mathbf{b}) > 0 & , \forall \mathbf{a} \neq \mathbf{b}. \end{cases} \quad (2.1.3)$$

As medidas de distância $\mathcal{D}$ mais utilizadas na literatura são as de Bregman [Bregman 1967, Censor e Zenios 1992, Pierro e Iusem 1986, Matioli e Gonzaga 2008] apresentada na Seção 2.2; as entrópicas φ -divergente de Csiszar [Csiszár 1967, Teboulle 1992, Iusem, Svaiter e Teboulle 1994] discutidas na Seção 2.2 e; a de Auslender [Auslender, Teboulle e Ben-Tiba 1999] brevemente discutida na Seção 2.7.

Em geral, as medidas de distâncias $\mathcal{D}$ são induzidas por funções não-quadráticas e convexas cujo domínio pode ser um subconjunto do $\mathbb{R}^n$. A seguir, discutiremos com mais detalhes as medidas de distância de Bregman e entrópicas φ -divergentes.

2.2 Distâncias de Bregman e entrópicas φ -divergente de Csiszar

Seja $\nu : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa de classe C^1 . A medida de distância de Bregman induzida por ν é expressa por [Bregman 1967, Pierro e Iusem 1986, Censor e Zenios 1992]:

$$\mathcal{D}_\nu(\mathbf{x}, \mathbf{y}) = \nu(\mathbf{x}) - \nu(\mathbf{y}) - \nabla \nu(\mathbf{y})^T (\mathbf{x} - \mathbf{y}), \quad \forall (\mathbf{x}, \mathbf{y}) \in (\mathbb{R}^n \times \mathbb{R}^n). \quad (2.2.1)$$

Uma interpretação geométrica da medida de distância de Bregman induzida por uma específica função convexa ν é dada na Figura 2.2.1.

Figura 2.2.1: Interpretação geométrica da medida de distância de Bregman induzida por ν .

É imediato constatararmos que se considerarmos o vetor $\mathbf{s} = (s_1, \dots, s_i, \dots, s_n)^T \in \mathbb{R}^n$ e a função $\nu_1(\mathbf{s}) = \|\mathbf{s}\|^2$, então a distância de Bregman induzida pela função ν_1 é dada por $\mathcal{D}_{\nu_1}(\mathbf{x}, \mathbf{y}) =$

$\|\mathbf{x} - \mathbf{y}\|^2$, a qual é a norma usual em $\mathbb{R}^n$. Logo, a distância euclidiana é um caso particular de distância de Bregman.

Entretanto, uma função ν não necessariamente induz uma métrica no $\mathbb{R}^n$ no sentido clássico, pois, em geral, as distâncias de Bregman não satisfazem as condições de simetria e de desigualdade triangular, mas topologicamente se comportam muito bem como distâncias [Iusem, Svaiter e Teboulle 1994]. Apresentamos na Proposição 2.2.1 um exemplo de medida de distância de Bregman que não induz a norma usual em $\mathbb{R}^n$.

Proposição 2.2.1. *Seja $\nu_2 : \mathbb{R}_{++}^n \rightarrow \mathbb{R}_+$ dada por $\nu_2(\mathbf{s}) = \sum_{i=1}^n (-\ln(s_i) + s_i - 1)$, então a medida de distância de Bregman induzida por ν_2 é a função*

$$\mathcal{D}_{\nu_2}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n (-\ln(y_i^{-1}x_i) + y_i^{-1}x_i - 1), \forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_{++}^{2n}. \quad (2.2.2)$$

Demonstração. Vide [Matioli e Gonzaga 2008]. ■

Não é difícil encontrarmos um ponto $(\bar{\mathbf{x}}, \bar{\mathbf{y}}) \in \mathbb{R}_{++}^{2n}$ tal que $\mathcal{D}_{\nu_2}(\bar{\mathbf{x}}, \bar{\mathbf{y}}) \neq \mathcal{D}_{\nu_2}(\bar{\mathbf{y}}, \bar{\mathbf{x}})$. Logo, $\mathcal{D}_{\nu_2}(\mathbf{x}, \mathbf{y})$ não define uma métrica em $\mathbb{R}^n$. Além disso, mostramos no Teorema 2.2.1 a seguir que a medida de distância $\mathcal{D}_{\nu_2}(\mathbf{x}, \mathbf{y})$ é um caso particular especial dentre uma classe de funções parametrizadas $\mathcal{D}_{\nu_3}(\mathbf{x}, \mathbf{y}, \alpha)$ proposta, onde α está em um intervalo real $[0, 1)$.

Teorema 2.2.1. *Seja $\nu_3 : (\mathbb{R}_{++}^n \times [0, 1)) \rightarrow \mathbb{R}_-$ definida por $\nu_3(\mathbf{s}, \alpha) = -\sum_{i=1}^n (\alpha^{-1}s_i^\alpha)$. Então a medida de distância de Bregman induzida por ν_3 é a função*

$$\mathcal{D}_{\nu_3}(\mathbf{x}, \mathbf{y}, \alpha) = \begin{cases} \sum_{i=1}^n (-\ln(y_i^{-1}x_i) + y_i^{-1}x_i - 1) & , \text{ se } \alpha = 0 \\ \sum_{i=1}^n (-\alpha^{-1}(x_i^\alpha - 1) + \alpha^{-1}(y_i^\alpha - 1) + y_i^{\alpha-1}x_i - y_i^\alpha) & , \text{ se } \alpha \in (0, 1). \end{cases} \quad (2.2.3)$$

Em particular, se $\alpha \rightarrow 0^+$, então $\mathcal{D}_{\nu_3}(\mathbf{x}, \mathbf{y}, \alpha)$ é cônica à função $\mathcal{D}_{\nu_2}(\mathbf{x}, \mathbf{y})$ apresentada em (2.2.2).

Demonstração. Em primeiro lugar, vamos supor que $\alpha \in (0, 1)$. Então, ao aplicarmos (2.2.1) na variável $\mathbf{s}$, temos:

$$\begin{aligned} \mathcal{D}_{\nu_3}(\mathbf{x}, \mathbf{y}, \alpha) &= \sum_{i=1}^n (-\alpha^{-1}x_i^\alpha + \alpha^{-1}y_i^\alpha + y_i^{\alpha-1}(x_i - y_i)) \\ &= \sum_{i=1}^n (-\alpha^{-1}x_i^\alpha + \alpha^{-1}y_i^\alpha + y_i^{\alpha-1}x_i - y_i^\alpha). \end{aligned}$$

Ao adicionarmos e subtraímos α^{-1} , vem:

$$\begin{aligned} \mathcal{D}_{\nu_3}(\mathbf{x}, \mathbf{y}, \alpha) &= \sum_{i=1}^n (-\alpha^{-1}x_i^\alpha + \alpha^{-1}y_i^\alpha + y_i^{\alpha-1}x_i - y_i^\alpha + \alpha^{-1} - \alpha^{-1}) \\ &= \sum_{i=1}^n (-\alpha^{-1}(x_i^\alpha - 1) + \alpha^{-1}(y_i^\alpha - 1) + y_i^{\alpha-1}x_i - y_i^\alpha). \end{aligned}$$

Assim, a medida de distância de Bregman induzida por ν_3 é de acordo com (2.2.3) para todo $\alpha \in (0, 1)$.

De acordo com a definição da função $v_3(\mathbf{s}, \alpha)$ está claro que esta função não existe se $\alpha = 0$. Todavia, sabemos que a função logarítmica $\ln(t)$ é obtida por meio do cálculo do limite fundamental logarítmico dado por $\lim_{\alpha \rightarrow 0^+} \alpha^{-1}(t^\alpha - 1)$; logo, temos que:

$$\begin{aligned} \lim_{\alpha \rightarrow 0^+} \mathcal{D}_{\nu_3}(\mathbf{x}, \mathbf{y}, \alpha) &= \lim_{\alpha \rightarrow 0^+} \left[\sum_{i=1}^n (-\alpha^{-1}(x_i^\alpha - 1) + \alpha^{-1}(y_i^\alpha - 1) + y_i^{\alpha-1}x_i - y_i^\alpha) \right] \\ &= \sum_{i=1}^n \left(-\lim_{\alpha \rightarrow 0^+} \alpha^{-1}(x_i^\alpha - 1) + \lim_{\alpha \rightarrow 0^+} \alpha^{-1}(y_i^\alpha - 1) + \lim_{\alpha \rightarrow 0^+} y_i^{\alpha-1}x_i - \lim_{\alpha \rightarrow 0^+} y_i^\alpha \right) \\ &= \sum_{i=1}^n -\ln(x_i) + \ln(y_i) + y_i^{-1}x_i - 1 \\ &= \sum_{i=1}^n -\ln(y_i^{-1}x_i) + y_i^{-1}x_i - 1; \end{aligned}$$

isto é, $\lim_{\alpha \rightarrow 0^+} \mathcal{D}_{\nu_3}(\mathbf{x}, \mathbf{y}, \alpha)$ existe e $\mathcal{D}_{\nu_3}(\mathbf{x}, \mathbf{y}, \alpha)$ torna-se idêntica a $\mathcal{D}_{\nu_2}(\mathbf{x}, \mathbf{y})$ em (2.2.2). Portanto, a proposição é verificada ao impormos que $\alpha = 0$ implica que $\mathcal{D}_{\nu_3}(\mathbf{x}, \mathbf{y}, 0)$ é cônica à função $\mathcal{D}_{\nu_2}(\mathbf{x}, \mathbf{y})$. ■

Corolário 2.2.1. *Seja $\bar{\mathbf{y}} \in \mathbb{R}_{++}^n$ conhecido, então, para todo $\mathbf{x} \in \mathbb{R}_{++}^n$ e $\alpha \in [0, 1)$, a distância $\mathcal{D}_{\nu_3}(\mathbf{x}, \bar{\mathbf{y}}, \alpha)$ possui as seguintes propriedades:*

- 1) $\mathcal{D}_{\nu_3}(\mathbf{x}, \bar{\mathbf{y}}, \alpha) \geq 0$;
- 2) $\mathcal{D}_{\nu_3}(\mathbf{x}, \bar{\mathbf{y}}, \alpha) = 0 \Leftrightarrow \mathbf{x} = \bar{\mathbf{y}}$;
- 3) $\mathcal{D}_{\nu_3}(\mathbf{x}, \bar{\mathbf{y}}, \alpha)$ é estritamente convexa;
- 4) $\lim_{x_i \rightarrow 0^+} \mathcal{D}_{\nu_3}(\mathbf{x}, \bar{\mathbf{y}}, \alpha) = \sum_{i=1}^n (\alpha^{-1} - 1) \bar{y}_i^\alpha$, se $\alpha \in (0, 1)$; e $\lim_{x_i \rightarrow 0^+} \mathcal{D}_{\nu_3}(\mathbf{x}, \bar{\mathbf{y}}, \alpha) = \infty$, se $\alpha = 0$, para todo $i = 1, \dots, n$;
- 5) $\lim_{x_i \rightarrow 0^+} \frac{\partial \mathcal{D}_{\nu_3}(\mathbf{x}, \bar{\mathbf{y}}, \alpha)}{\partial x_i} = -\infty$, para todo $i = 1, \dots, n$;
- 6) $\lim_{x_i \rightarrow \infty} \frac{\partial \mathcal{D}_{\nu_3}(\mathbf{x}, \bar{\mathbf{y}}, \alpha)}{\partial x_i} = 1$, para todo $i = 1, \dots, n$;
- 7) $\lim_{\alpha \rightarrow 1^-} \mathcal{D}_{\nu_3}(\mathbf{x}, \bar{\mathbf{y}}, \alpha) = 0$, para todo $\mathbf{x} \in \mathbb{R}_{++}^n$ (Degenerescência). —

Demonstração. Todas as propriedades são imediatas devido à convexidade e à diferenciabilidade da função $\nu_3(\mathbf{s}, \alpha) = -\sum_{i=1}^n (\alpha^{-1} s_i^\alpha)$ para todo $\alpha \in (0, 1)$. No caso em que $\alpha = 0$ essas mesmas propriedades são também imediatas devido à convexidade e à diferenciabilidade da função $\nu_2(\mathbf{s}) = \sum_{i=1}^n (-\ln(s_i) + s_i - 1)$. ■

Outra maneira de expressarmos uma medida de distância de um vetor $\mathbf{x}$ para o vetor $\mathbf{y}$ é a medida entrópica φ -divergente de Csiszar [Csiszár 1967] dada em (2.2.4).

$$d_\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n y_i \varphi(y_i^{-1} x_i), \quad \forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_{++}^{2n}, \quad (2.2.4)$$

onde $\varphi: \mathbb{R}_{++} \rightarrow \mathbb{R}_+$ é elemento de um conjunto $\mathbb{F}$ de funções de classe C^∞ que possui as seguintes propriedades [Teboulle 1992, Teboulle 1997, Iusem, Svaiter e Teboulle 1994]:

- 1) $\varphi(1) = \varphi'(1) = 0$;

- 2) $\varphi''(s) > 0, \forall s \in \mathbb{R}_{++}$;
- 3) $\varphi'(s) < 0, \forall s \in (0, 1)$ e $\varphi'(s) > 0, \forall s \in (1, \infty)$;
- 4) $\lim_{s \rightarrow 0^+} \varphi'(s) = -\infty$ e $\lim_{s \rightarrow \infty} \varphi'(s) = b$, com $0 < b \leq \infty$;
- 5) $\lim_{s \rightarrow 0^+} \varphi(s) = a$, com $0 < a \leq \infty$ e;
- 6) $\lim_{s \rightarrow \infty} \varphi(s) = \infty$.

É bem conhecido que a função $\varphi_1(s) = s \ln(s) - s + 1$ induz por meio de (2.2.4) a medida entrópica de Kullback-Liebler definida em (2.2.5), a qual é amplamente usada na teoria de probabilidades e na teoria da informação, [Teboulle 1992].

$$d_{\varphi_1}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n (x_i \ln(y_i^{-1} x_i) - x_i + y_i). \quad (2.2.5)$$

Em particular, é imediato obtermos (2.2.5) ao induzirmos a medida de distância de Bregman com a função $\nu(s) = \sum_{i=1}^n \varphi_1(s_i)$. Assim, temos um exemplo em que $\mathcal{D}_\nu(\mathbf{x}, \mathbf{y}) = d_\varphi(\mathbf{x}, \mathbf{y})$.

Por fim, é possível estabelecer uma relação entre as medidas de distâncias de Bregman e entrópicas φ -divergente sempre que essas são induzidas por qualquer $\varphi \in \mathbb{F}$. De fato, consideremos uma função $\mathcal{B}_\varphi : \mathbb{R}_{++}^2 \rightarrow \mathbb{R}$ definida por $\mathcal{B}_\varphi(u, t) = \varphi(u) - \varphi(t) - \varphi'(t)(u - t)$. Então, através da função $\mathcal{B}_\varphi$, as medidas de distância de Bregman e entrópicas φ -divergente são relacionadas pelas expressões (2.2.6), (2.2.7) e (2.2.8) [Teboulle 1992, Matioli e Gonzaga 2008, Polyak 2014]:

$$\mathcal{D}_\nu(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \mathcal{B}_\varphi(x_i, y_i) = \sum_{i=1}^n [\varphi(x_i) - \varphi(y_i) - \varphi'(y_i)(x_i - y_i)], \quad \forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_{++}^{2n}. \quad (2.2.6)$$

$$\mathcal{Q}_\nu(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \mathcal{B}_\varphi(y_i^{-1} x_i, 1) = \sum_{i=1}^n \varphi(y_i^{-1} x_i), \quad \forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_{++}^{2n}. \quad (2.2.7)$$

$$d_\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n y_i \mathcal{B}_\varphi(y_i^{-1} x_i, 1) = \sum_{i=1}^n y_i \varphi(y_i^{-1} x_i), \quad \forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_{++}^{2n}. \quad (2.2.8)$$

Por exemplo, se a função $\varphi_2(s) = -\ln(s) + s - 1$ for utilizada para induzir uma distância de Bregman através das expressões (2.2.1), (2.2.6) e (2.2.7), então obtermos novamente a expressão (2.2.2), ou seja, as formulações (2.2.1), (2.2.6) e (2.2.7) são equivalentes através da indução por φ_2 .

Entretanto, não é verdade afirmarmos que podemos obter um resultado semelhante para toda $\varphi \in \mathbb{F}$. De fato, é imediato constatarmos que não obtermos (2.2.5) se $\varphi_1(s) = s \ln(s) - s + 1$ for escolhida para induzir uma medida de distância de Bregman por meio da formulação (2.2.7).

2.3 Medida de distância integrada

A classe das funções $\mathbb{F}$ mediante a função $\mathcal{B}_\varphi$ nos permitiu obter uma formulação matemática semelhante associada à distância de Bregman (2.2.7) e entrópica φ -divergente (2.2.8). Nesta seção, o

conceito de homotopia entre duas funções é o fundamento para que as medidas $\mathcal{Q}_\nu(\mathbf{x}, \mathbf{y})$ e $d_\varphi(\mathbf{x}, \mathbf{y})$ sejam integradas de modo a formar um todo coerente e harmônico.

De acordo com [Hatcher 2001, p. 25], duas funções contínuas f e g são homotópicas se existe uma função contínua $\mathcal{H} : (\mathbb{R}^n \times [0, 1]) \rightarrow \mathbb{R}$ tal que $\mathcal{H}(\mathbf{x}, 0) = f(\mathbf{x})$ e $\mathcal{H}(\mathbf{x}, 1) = g(\mathbf{x})$. Mediante a homotopia, propomos o conceito de medida de distância integrada.

Definição 2.3.1 (Medida de Distância Integrada). *Sejam $\varphi \in \mathbb{F}$ e as medidas de distância de Bregman (2.2.7) e entrópica φ -divergente de Csiszar (2.2.8). Denominamos “medida de distância integrada” à coleção de todas as funções $\mathcal{H}_\varphi : (\mathbb{R}_{++}^{2n} \times [0, 1]) \rightarrow \mathbb{R}_+$, para as quais $\mathcal{H}_\varphi(\mathbf{x}, \mathbf{y}, 0) = \sum_{i=1}^n \varphi(y_i^{-1}x_i)$ e $\mathcal{H}_\varphi(\mathbf{x}, \mathbf{y}, 1) = \sum_{i=1}^n y_i \varphi(y_i^{-1}x_i)$. $\square$*

Em [Auslender, Teboulle e Ben-Tiba 1999, Auslender, Teboulle e Ben-Tiba 1999], os autores propõem a medida de distância $d_\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n y_i^2 \varphi(y_i^{-1}x_i)$. Por não ser uma homotopia entre as medidas de distância de Bregman e entrópica φ -divergente, esta medida de distância não será investigada neste trabalho.

Teorema 2.3.1. *Seja $\varphi \in \mathbb{F}$, então as funções $\mathcal{H}_\varphi$ apresentadas em (2.3.1), (2.3.2) e (2.3.3) são medidas de distância integrada.*

$$\mathcal{H}_\varphi(\mathbf{x}, \mathbf{y}, \alpha) = \sum_{i=1}^n (1 + (y_i - 1)\alpha) \varphi(y_i^{-1}x_i), \forall (\mathbf{x}, \mathbf{y}, \alpha) \in (\mathbb{R}_{++}^{2n} \times [0, 1]). \quad (2.3.1)$$

$$\mathcal{H}_\varphi(\mathbf{x}, \mathbf{y}, \alpha) = \sum_{i=1}^n y_i^\alpha \varphi(y_i^{-1}x_i), \forall (\mathbf{x}, \mathbf{y}, \alpha) \in (\mathbb{R}_{++}^{2n} \times [0, 1]). \quad (2.3.2)$$

$$\mathcal{H}_\varphi(\mathbf{x}, \mathbf{y}, \alpha) = \sum_{i=1}^n \frac{y_i}{\alpha + (1 - \alpha)y_i} \varphi(y_i^{-1}x_i), \forall (\mathbf{x}, \mathbf{y}, \alpha) \in (\mathbb{R}_{++}^{2n} \times [0, 1]). \quad (2.3.3)$$

Demonstração. O resultado procede imediatamente ao impormos $\alpha = 0$ e $\alpha = 1$. $\blacksquare$

Por simplicidade, consideremos a função contínua $\gamma : \mathbb{R}_{++} \times [0, 1] \rightarrow \mathbb{R}_+$, a qual possui as seguintes propriedades:

- i) Se $\bar{v} = 0$, então $\gamma(u, \bar{v}) = 1$ para todo $u \in \mathbb{R}_{++}$;
- ii) Se $\bar{v} = 1$, então $\gamma(u, \bar{v}) = u$ para todo $u \in \mathbb{R}_{++}$,

então, uma medida de distância integrada arbitrária será representada doravante por:

$$\mathcal{H}_\varphi(\mathbf{x}, \mathbf{y}, \alpha) = \sum_{i=1}^n \gamma(y_i, \alpha) \varphi(y_i^{-1}x_i), \forall (\mathbf{x}, \mathbf{y}, \alpha) \in (\mathbb{R}_{++}^{2n} \times [0, 1]). \quad (2.3.4)$$

Denominamos $\gamma(y_i, \alpha)$ e $\alpha \in [0, 1]$ respectivamente por fator de transformação e parâmetro de homotopia. O fator de transformação define a forma pela qual a medida de distância integrada se transmuta de distância de Bregman para a distância entrópica φ -divergente e vice-e-versa conforme a variação do parâmetro de homotopia α .

Observação 2.3.1. Por definição, a medida de distância integrada $\mathcal{H}_\varphi(\mathbf{x}, \mathbf{y}, \alpha)$ não é nem uma distância de Bregman nem entrópica para qualquer $\alpha \in (0, 1)$. Mas surpreendentemente esta afirmação não é simultaneamente verdadeira para qualquer combinação entre a função φ em $\mathbb{F}$, fator de transformação $\gamma(y_i, \alpha)$ e parâmetro de homotopia α , conforme o Teorema 2.3.2 a seguir mostra. $\square$

Teorema 2.3.2. *Seja $\mathbb{F} \ni \varphi_3(s) = -2\sqrt{s} + s + 1$ e o fator de transformação $\gamma(y_i, \alpha) = y_i^\alpha$. Se $\alpha = 0,5$; então a medida de distância integrada $\mathcal{H}_{\varphi_3}(\mathbf{x}, \mathbf{y}, 0,5)$ é uma distância de Bregman.* —

Demonstração. De fato, a medida de distância integrada envolvendo os elementos $\varphi_3(s) = -2\sqrt{s} + s + 1$ e $\gamma(y_i, \alpha) = y_i^\alpha$ é dada por:

$$\mathcal{H}_{\varphi_3}(\mathbf{x}, \mathbf{y}, \alpha) = \sum_{i=1}^r y_i^\alpha \left(-2\sqrt{\frac{x_i}{y_i}} + \frac{x_i}{y_i} + 1 \right), \forall (\mathbf{x}, \mathbf{y}, \alpha) \in (\mathbb{R}_{++}^{2n} \times [0, 1]). \quad (2.3.5)$$

Ao impormos $\alpha = 0,5$, obtemos:

$$\mathcal{H}_{\varphi_3}(\mathbf{x}, \mathbf{y}, 0,5) = \sum_{i=1}^r \left(-2\sqrt{x_i} + \sqrt{y_i} + \frac{x_i}{\sqrt{y_i}} \right), \forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_{++}^{2n}. \quad (2.3.6)$$

Agora, consideremos a distância de Bregman $\mathcal{D}_{v_3}(\mathbf{x}, \mathbf{y}, \alpha)$ em (2.2.3). Ao impormos α em 0,5, obtemos:

$$\begin{aligned} \mathcal{D}_{v_3}(\mathbf{x}, \mathbf{y}, 0,5) &= \sum_{i=1}^n \left(-0,5^{-1} \left(x_i^{0,5} - 1 \right) + 0,5^{-1} \left(y_i^{0,5} - 1 \right) + y_i^{0,5-1} x_i - y_i^{0,5} \right) \\ &= \sum_{i=1}^n \left(-2 \left(\sqrt{x_i} - 1 \right) + 2 \left(\sqrt{y_i} - 1 \right) + \left(\sqrt{y_i} \right)^{-1} x_i - \sqrt{y_i} \right) \\ &= \sum_{i=1}^n \left(-2\sqrt{x_i} + \sqrt{y_i} + \left(\sqrt{y_i} \right)^{-1} x_i \right). \end{aligned}$$

Portanto, a medida de distância integrada $\mathcal{H}_{\varphi_3}(\mathbf{x}, \mathbf{y}, 0,5)$ em (2.3.6) é genuinamente uma distância de Bregman. $\blacksquare$

Observação 2.3.2. Em [Matioli e Gonzaga 2008] os autores constataam que a função $\mathcal{D}_{v_2}(\mathbf{x}, \mathbf{y})$ apresentada em (2.2.2) estabeleceu pela primeira vez a conexão entre a função barreira logarítmica modificada desenvolvida em [Polyak 1992, Polyak e Teboulle 1997] com a medida de distância de Bregman. Conjecturamos que o Teorema 2.3.2 estabeleceu pela primeira vez uma conexão envolvendo a função barreira hiperbólica modificada desenvolvida em [Polyak 1992, Polyak e Teboulle 1997] e distâncias de Bregman, pois não encontramos alguma literatura que estabelece esta relação. $\square$

2.4 Declaração do problema e premissas básicas

Seja o problema de programação não-linear dado em (2.4.1):

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) \\ \text{s.a :} \quad & h_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, r, \end{aligned} \quad (2.4.1)$$

em que: $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$, são funções de classe C^1 e convexas. Consideramos, sempre que oportuno, o vetor $\mathbf{h} = (h_1, \dots, h_i, \dots, h_r)^T \in \mathbb{R}^r$. Sobre o problema (2.4.1), assumimos as seguintes premissas:

Premissa 1: O conjunto viável $\mathbb{X} = \{\mathbf{x} \in \mathbb{R}^n; h_i(\mathbf{x}) \leq 0, \forall i = 1, \dots, r\}$ é não-vazio, convexo e compacto.

Premissa 2: O conjunto dos minimizadores $\mathbb{X}^* = \{\mathbf{x}^*; \mathbf{x}^* = \arg \min f(\mathbf{x}), \mathbf{x} \in \mathbb{X}\}$ é não-vazio e limitado.

Premissa 3: A condição de Slater é satisfeita, isto é, existe $\bar{\mathbf{x}} \in \mathbb{R}^n$ tal que $h_i(\bar{\mathbf{x}}) < 0$ para todo i que está no conjunto $\mathbb{I}(\mathbf{x}^*) = \{i; h_i(\mathbf{x}^*) = 0\}$ de restrições ativas [Bertsekas 1999, p. 331].

Em consequência da Premissa 3, temos que as condições necessárias de otimalidade de Karush-Kuhn-Tucker (KKT) estão satisfeitas em $\mathbf{x}^*$. Logo, existe um vetor $\lambda^* = (\lambda_1, \dots, \lambda_i, \dots, \lambda_r)^T \in \mathbb{R}_+^r$ de multiplicadores de Lagrange tal que as equações (2.4.2) e (2.4.3) a seguir são consistentes:

$$\nabla_x L(\mathbf{x}^*, \lambda^*) = \nabla f(\mathbf{x}^*) + \sum_{i=1}^r \lambda_i^* \nabla h_i(\mathbf{x}^*) = \mathbf{0}, \quad (2.4.2)$$

$$\min(\lambda_i^*, -h_i(\mathbf{x}^*)) = 0, \quad \forall i = 1, \dots, r, \quad (2.4.3)$$

em que: $L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \sum_{i=1}^r \lambda_i h_i(\mathbf{x})$ é a função lagrangiana. O problema dual lagrangiano vinculado ao problema primal (2.4.1) é apresentado em (2.4.4).

$$\begin{aligned} \max \quad & d(\lambda) \\ \text{s.a.} \quad & \lambda \geq \mathbf{0}, \end{aligned} \quad (2.4.4)$$

onde $d(\lambda) = \inf \{L(\mathbf{x}, \lambda); \mathbf{x} \in \mathbb{X}\}$ é a função dual. O conjunto viável $\mathbb{L} \subset \mathbb{R}_+^r$ associado ao problema (2.4.1) é não-vazio e limitado [Gauvin 1977]; o conjunto $\text{int}(\mathbb{L}) = \{\lambda \in \mathbb{L}; \lambda_i > 0, \forall i = 1, \dots, r\}$ é não-vazio e o conjunto dos maximizadores $\mathbb{L}^* = \{\lambda^*; \lambda^* = \arg \max d(\lambda), \lambda \in \mathbb{L}\}$ é não-vazio.

Além disso, seja $\mathbb{X}(\bar{\lambda}) = \{\bar{\mathbf{x}}; \bar{\mathbf{x}} = \arg \min \{L(\mathbf{x}, \bar{\lambda}), \mathbf{x} \in \mathbb{X}\}\}$ o conjunto das soluções ótimas do subproblema primal lagrangiano, então $\mathbb{X}(\bar{\lambda})$ é não-vazio para qualquer $\bar{\lambda}$ que está em $\text{int}(\mathbb{L})$ e $\mathbf{h}(\bar{\mathbf{x}}) \in \partial d(\bar{\lambda})$, em que $\partial d(\bar{\lambda}) = \{\xi \in \mathbb{R}^r; d(\lambda) \leq d(\bar{\lambda}) + \xi^T(\lambda - \bar{\lambda}), \forall \lambda \in \mathbb{L}\}$ é o conjunto dos subgradientes de $d(\bar{\lambda})$ em $\bar{\lambda}$ [Bazaraa, Sherali e Shetty 2006].

2.5 Método dual de ponto proximal

O método dual de ponto proximal clássico, amplamente discutido em [Rockafellar 1973, Rockafellar 1976, Rockafellar 1976], é definido em (2.5.1):

$$\lambda^{k+1} = \arg \max \left\{ d(\lambda) - \frac{\mu_k}{2} \|\lambda - \lambda^k\|^2; \lambda \in \mathbb{R}_+^r \right\}. \quad (2.5.1)$$

Dentre as propriedades do método dual de ponto proximal, destacamos o efeito de suavização sobre a função $d(\lambda)$ e a preservação do conjunto de maximizadores associado ao problema dual, [Rockafellar 1976, Seção 4].

Métodos dual de ponto proximal generalizado são propostos em [Teboulle 1992] e em [Matioli e Gonzaga 2008]. A proposta dos autores consiste em substituir a medida euclidiana $\frac{1}{2} \|\lambda - \bar{\lambda}\|^2$ pelas

medidas de Bregman e de Csiszar. Os métodos proximais apresentados em (2.5.2) e (2.5.3) descrevem as abordagens propostas em [Teboulle 1992] com as distâncias entrópicas φ -divergente e de Bregman, respectivamente; enquanto que (2.5.3) descreve o método proximal proposto em [Matioli e Gonzaga 2008] com distâncias de Bregman.

$$\lambda^{k+1} = \arg \max \left\{ d(\lambda) - \mu_k \sum_{i=1}^r \lambda_{i_k} \varphi \left(\lambda_{i_k}^{-1} \lambda_{i_k} \right); \lambda \in \mathbb{R}_{++}^r \right\}. \quad (2.5.2)$$

$$\lambda^{k+1} = \arg \max \left\{ d(\lambda) - \mu_k \sum_{i=1}^r [\varphi(\lambda_i) - \varphi(\lambda_{i_k}) - \varphi'(\lambda_{i_k})(\lambda_i - \lambda_{i_k})], \lambda \in \mathbb{R}_{++}^r \right\}. \quad (2.5.3)$$

$$\lambda^{k+1} = \arg \max \left\{ d(\lambda) - \mu_k \sum_{i=1}^r \varphi \left(\lambda_{i_k}^{-1} \lambda_i \right); \lambda \in \mathbb{R}_{++}^r \right\}. \quad (2.5.4)$$

Assim como o método dual de ponto de proximal clássico, as versões generalizadas preservam o efeito de suavização da função $d(\lambda)$ e o conjunto de maximizadores associado ao problema dual [Teboulle 1992, teor. 4.1]. Além dessas interessantes propriedades, temos que a viabilidade dual é automaticamente tratada por estes métodos generalizados, pois o domínio da função $\varphi \in \mathbb{F}$ é o $\mathbb{R}_{++}$. Desta forma, os métodos duais de ponto proximal desenvolvidos em [Teboulle 1992, Matioli e Gonzaga 2008] podem ser interpretados como um tipo de método pontos interiores não-clássico no espaço dual.

Os métodos proximais (2.5.2) e (2.5.4) nos propiciam integrar as abordagens desenvolvidas em [Teboulle 1992, Iusem, Svaiter e Teboulle 1994, Polyak e Teboulle 1997, Matioli e Gonzaga 2008] ao substituímos as medidas de distância de Bregman e entrópicas φ -divergente pela medida de distância integrada (2.3.4). Desta forma, o método dual de ponto proximal com a medida de distância integrada é definido por:

$$\lambda^{k+1} = \arg \max \left\{ d(\lambda) - \mu_k \sum_{i=1}^r \gamma(\lambda_{i_k}, \alpha) \varphi \left(\lambda_{i_k}^{-1} \lambda_i \right); \lambda \in \mathbb{R}_{++}^r \right\}. \quad (2.5.5)$$

Proposição 2.5.1 (Convergência). *Sejam $\varphi \in \mathbb{F}$, $\mu_0 > 0$, $\alpha \in [0, 1]$, $\gamma(\lambda_i, \alpha)$ o fator de transformação e a medida de distância integrada $\mathcal{H}_\varphi(\lambda, \bar{\lambda}, \alpha) = \sum_{i=1}^r \gamma(\bar{\lambda}_i, \alpha) \varphi(\bar{\lambda}_i^{-1} \lambda_i)$. Suponha que a sequência $\{\lambda^k\}$ é gerada por meio do algoritmo (2.5.5) e que $\mu_{k+1} < \mu_k$. Então a sequência $\{d(\lambda^k)\}$ converge para $d(\lambda^*)$ do problema dual (2.4.4).* —

Observação 2.5.1. A demonstração da Proposição 2.5.1 é um tema de pesquisa para trabalhos futuros. No entanto, acreditamos que com procedimentos de análise similares àqueles apresentados em [Teboulle 1992, Polyak e Teboulle 1997, Matioli e Gonzaga 2008], então a veracidade da proposição deve prevalecer. $\square$

Na próxima seção, vamos deduzir o método de lagrangiano aumentado associado à resolução do problema primal (2.4.1) e que seja equivalente ao método dual de ponto proximal proposto neste trabalho.

2.6 Método de Reescalamamento não-linear Integrado

O método de lagrangiano aumentado clássico associado ao problema (2.4.1) é definido por:

$$\begin{aligned} \mathbf{x}^{k+1} &= \arg \min \{ L(\mathbf{x}, \lambda^k, \mu_k) ; \mathbf{x} \in \mathbb{X} \} \\ \lambda_{i_{k+1}} &= \max(0, \lambda_{i_k} + \mu_k^{-1} h_i(\mathbf{x}^{k+1})), \quad \forall i = 1, \dots, r, \end{aligned} \quad (2.6.1)$$

onde $L(\mathbf{x}, \bar{\lambda}, \mu) = f(\mathbf{x}) + \mu \frac{1}{2} \sum_{i=1}^r \left[\max(0, \bar{\lambda}_i + \mu^{-1} h_i(\mathbf{x}))^2 - \bar{\lambda}_i^2 \right]$ é a função lagrangiana aumentada clássica [Rockafellar 1973].

A estreita relação que existe entre o método de lagrangiano aumentado (2.6.1) e o método dual de ponto proximal (2.5.1) é que eles são equivalentes. Esta equivalência não é imediata, mas nos assegura dizer que um método pode ser interpretado como outro e vice-versa. Por fim, mostra-se que λ^{k+1} em $\mathbb{R}_+^r$ apresentado na segunda expressão em (2.6.1) é exatamente a solução do problema $\max \left\{ d(\lambda) - \mu_k \frac{1}{2} \|\lambda - \lambda^k\|^2 ; \lambda \in \mathbb{R}_+^r \right\}$.

Em problemas de otimização convexa, nos quais a condição de Slater é satisfeita, Matioli e Gonzaga [Matioli e Gonzaga 2008] e Polyak e Teboulle [Polyak e Teboulle 1997] aplicaram, respectivamente, as medidas de distância de Bregman e entrópicas φ -divergente junto ao método de ponto proximal no espaço dual associado ao problema primal (2.4.1). Foi demonstrado pelos autores que resolver o problema dual por meio desta abordagem é equivalente a resolver o problema primal associado por meio de uma específica abordagem do método de lagrangiano aumentado com funções penalidades não-quadráticas. A especificidade do método de lagrangiano aumentado está em acordo com a classe de funções penalidades adotada.

Para distinguir uma classe da outra, Matioli e Gonzaga denominaram por classe $\mathbb{P}_1$ o conjunto de funções penalidades associadas ao método de lagrangiano aumentado, cujo conjugado convexo induz a distância de Bregman para o método dual de ponto proximal; por classe $\mathbb{P}_2$ o conjunto das funções penalidades vinculadas ao método de lagrangiano aumentado, cujo conjugado convexo induz a distância entrópica φ -divergente para o método dual de ponto proximal.

Nesta seção, mostramos que o método dual de ponto proximal com a medida de distância integrada nos conduz para uma nova e equivalente abordagem para os métodos de lagrangiano aumentado com funções penalidades não-quadráticas, a qual denominamos método de reescalamamento não-linear integrado (RNLI). Denominamos por classe $\mathbb{P}_{\text{int}}$ o conjunto de todas as funções penalidades vinculadas ao método RNLI, pois a classe $\mathbb{P}_1$ (a qual define o método de lagrangiano aumentado de Matioli e Gonzaga) e a classe $\mathbb{P}_2$ (a qual define o método de reescalamamento não-linear original de Polyak e Teboulle) são casos particulares da classe $\mathbb{P}_{\text{int}}$.

Dentre as funções da classe $\mathbb{P}_{\text{int}}$, apresentamos os conceitos de funções barreira logarítmica modificada integrada (IMBF) e os de funções exponencial integrada.

Por fim, apresentamos três casos particulares do método RNLI e uma primeira versão geral primal-dual deste. Estes métodos diferem-se mutuamente apenas pela escolha do fator de transformação vinculado à medida de distância integrada. Esta seção está dividida da seguinte forma:

- Na Subseção 2.6.1, deduzimos a proposta da função lagrangiana aumentada integrada e apresentamos a nova classe $\mathbb{P}_{\text{int}}$ de funções penalidades para os métodos de lagrangiano aumentado. Além disso, apresentamos o conceito das funções barreira logarítmica modificada integrada e exponencial integrada, as quais estão em $\mathbb{P}_{\text{int}}$, e de função de reescalamamento integrado.

- Na Subseção 2.6.2, apresentamos o método RNLI e três casos particulares deste.

2.6.1 A função lagrangiana aumentada integrada e a classe $\mathbb{P}_{\text{int}}$ de funções penalidades

Devido à Premissa 3 ser assegurada, o Teorema de Dualidade Forte nos assegura que se o par primal-dual $(\mathbf{x}^*, \lambda^*) \in (\mathbb{X}^* \times \mathbb{L}^*)$ resolve simultaneamente os problemas (2.4.1) e (2.4.4) respectivamente com $f(\mathbf{x}^*) = d(\lambda^*)$ e $\mathbf{h}(\mathbf{x}^*)^T \lambda^* = 0$, então $\sup_{\lambda \in \mathbb{L}} \inf_{\mathbf{x} \in \mathbb{X}} L(\mathbf{x}, \lambda) = \inf_{\mathbf{x} \in \mathbb{X}} \sup_{\lambda \in \mathbb{L}} L(\mathbf{x}, \lambda)$ [Bazaraa, Serali e Shetty 2006, p. 267]. Este resultado nos permite considerar a seguinte abordagem proximal:

$$L_\alpha(\mathbf{x}, \bar{\lambda}, \mu) = \sup \{ L(\mathbf{x}, \lambda) - \mu \mathcal{H}_\varphi(\lambda, \bar{\lambda}, \alpha), \lambda \in \mathbb{R}_{++}^r \}.$$

Com maiores detalhes, a função primal lagrangiana aumentada $L_\alpha(\mathbf{x}, \bar{\lambda}, \mu)$ é apresentada em (2.6.2):

$$\begin{aligned} L_\alpha(\mathbf{x}, \bar{\lambda}, \mu) &= \sup \{ L(\mathbf{x}, \lambda) - \mu \mathcal{H}_\varphi(\lambda, \bar{\lambda}, \alpha), \lambda \in \mathbb{R}_{++}^r \} \\ &= \sup \{ f(\mathbf{x}) + \lambda^T \mathbf{h}(\mathbf{x}) - \mu \mathcal{H}_\varphi(\lambda, \bar{\lambda}, \alpha), \lambda \in \mathbb{R}_{++}^r \} \\ &= f(\mathbf{x}) + \sup \{ \lambda^T \mathbf{h}(\mathbf{x}) - \mu \mathcal{H}_\varphi(\lambda, \bar{\lambda}, \alpha), \lambda \in \mathbb{R}_{++}^r \} \\ &= f(\mathbf{x}) + \sup \left\{ \sum_{i=1}^r [\lambda_i h_i(\mathbf{x}) - \mu \gamma(\bar{\lambda}_i, \alpha) \varphi(\bar{\lambda}_i^{-1} \lambda_i)], \lambda \in \mathbb{R}_{++}^r \right\} \\ &= f(\mathbf{x}) + \mu \sum_{i=1}^r \gamma(\bar{\lambda}_i, \alpha) \sup \left\{ \left(\lambda_i (\mu \gamma(\bar{\lambda}_i, \alpha))^{-1} h_i(\mathbf{x}) - \varphi(\bar{\lambda}_i^{-1} \lambda_i) \right), \lambda \in \mathbb{R}_{++}^r \right\}. \end{aligned} \quad (2.6.2)$$

Nosso objetivo é explicitarmos analiticamente a expressão (2.6.2). Para isso, necessitamos dos conceitos de funções conjugadas e de dualidade de Fenchel [Bertsekas 1999, seção 5.4]. O conjugado convexo de φ é a função $\psi : \mathbb{T} \subseteq \mathbb{R} \rightarrow \mathbb{R}$ de classe C^∞ obtida por meio da transformada de Fenchel apresentada em (2.6.3):

$$\psi(t) = \sup \{ st - \varphi(s), s \in \mathbb{R}_{++} \}. \quad (2.6.3)$$

A função $\varphi \in \mathbb{F}$ é também própria ($\varphi(s) > -\infty$) e inferiormente semi-continua¹. Um resultado importante e de interesse provém da seguinte proposição [Kanniappan 1983]:

Proposição 2.6.1. $\varphi(s) = \sup \{ st - \psi(t), t \in \mathbb{T} \}$ se, e somente se, φ é convexa, própria e inferiormente semi-continua. Isto é, $\varphi(s)$ é idêntica à sua função biconjugada. —

Como $\varphi''(s) > 0$ para todo $s \in \mathbb{R}_{++}$, então $\varphi'(s)$ é estritamente crescente em $\mathbb{R}_{++}$, portanto φ'^{-1} existe. Isso significa que podemos explicitar $\psi(t)$ em função da sua taxa de variação. Ao utilizarmos (2.6.3), obtemos (2.6.4):

$$\psi(t) = t\varphi'^{-1}(t) - \varphi(\varphi'^{-1}(t)). \quad (2.6.4)$$

A partir de (2.6.4), explicitamos a relação fundamental apresentada em (2.6.5):

$$\psi'(t) = \varphi'^{-1}(t). \quad (2.6.5)$$

¹Dizemos que φ é inferiormente semi-continua se para cada $\bar{s} \in \mathbb{R}_{++}$ e sequência $\{s_k\}$ tal que $\lim_{k \rightarrow \infty} s_k = \bar{s}$, temos $\lim_{k \rightarrow \infty} \inf \varphi(s_k) \geq \varphi(\bar{s})$.

Portanto, com esta relação, a função $\psi(t)$ é também segundo (2.6.6):

$$\psi(t) = t\psi'(t) - \varphi(\psi'(t)). \quad (2.6.6)$$

Devido às propriedades da função $\varphi \in \mathbb{F}$, pode-se mostrar que $\psi(t)$ é elemento de um conjunto $\mathbb{H}$ com as seguintes características [Polyak e Teboulle 1997, Mاتيoli e Gonzaga 2008]:

- 1) $\mathbb{T} = (-\infty, b)$, com $0 < b \leq \infty$;
- 2) $\psi(0) = 0$ e $\psi'(0) = 1$;
- 3) $\psi'(t) > 0$ e $\psi''(t) > 0, \forall t \in \mathbb{T}$;
- 4) $\lim_{t \rightarrow -\infty} \psi'(t) = 0$ e $\lim_{t \rightarrow b^-} \psi'(t) = \infty$ (coercividade à direita).

Retornemos à expressão (2.6.2). Seja a função $p(\lambda_i) = (\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} \lambda_i h_i(\mathbf{x}) - \varphi(\bar{\lambda}_i^{-1} \lambda_i)$. Vamos determinar $\hat{\lambda}_i$ que maximiza a função côncava $p(\lambda_i)$. A derivada de $p(\lambda_i)$ é a função $p'(\lambda_i) = (\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} h_i(\mathbf{x}) - (\bar{\lambda}_i)^{-1} \varphi'(\bar{\lambda}_i^{-1} \lambda_i)$. Ao aplicarmos as condições suficientes de primeira ordem, obtemos $\hat{\lambda}_i = \varphi'^{-1}\left((\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} \bar{\lambda}_i h_i(\mathbf{x})\right) \bar{\lambda}_i$. Logo, pela relação fundamental (2.6.5), obtemos (2.6.7):

$$\hat{\lambda}_i = \psi'\left((\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} \bar{\lambda}_i h_i(\mathbf{x})\right) \bar{\lambda}_i, \quad \forall i = 1, \dots, r. \quad (2.6.7)$$

Assim, segue que:

$$\begin{aligned} p(\hat{\lambda}_i) &= \bar{\lambda}_i \psi'\left((\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} \bar{\lambda}_i h_i(\mathbf{x})\right) (\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} h_i(\mathbf{x}) - \varphi\left(\bar{\lambda}_i^{-1} \bar{\lambda}_i \psi'\left((\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} \bar{\lambda}_i h_i(\mathbf{x})\right)\right) \\ &= \psi'\left((\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} \bar{\lambda}_i h_i(\mathbf{x})\right) (\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} \bar{\lambda}_i h_i(\mathbf{x}) - \varphi\left(\psi'\left((\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} \bar{\lambda}_i h_i(\mathbf{x})\right)\right). \end{aligned}$$

Se identificarmos $t = (\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} \bar{\lambda}_i h_i(\mathbf{x})$, então, pela expressão (2.6.6), obtemos:

$$p(\hat{\lambda}_i) = \psi\left(\frac{\bar{\lambda}_i h_i(\mathbf{x})}{\mu\gamma(\bar{\lambda}_i, \alpha)}\right).$$

Logo, segue que:

$$\begin{aligned} L_\alpha(\mathbf{x}, \bar{\lambda}, \mu) &= f(\mathbf{x}) + \mu \sum_{i=1}^r \gamma(\bar{\lambda}_i, \alpha) \sup \left\{ \left((\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} \lambda_i h_i(\mathbf{x}) - \varphi(\lambda_i^{-1} \lambda_i) \right), \lambda \in \mathbb{R}_{++}^r \right\} \\ &= f(\mathbf{x}) + \mu \sum_{i=1}^r \gamma(\bar{\lambda}_i, \alpha) \psi\left((\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} \bar{\lambda}_i h_i(\mathbf{x})\right). \end{aligned}$$

Portanto, a função lagrangiana aumentada $L_\alpha(\mathbf{x}, \bar{\lambda}, \mu)$ proposta neste trabalho é de acordo com a expressão (2.6.8) a seguir:

$$L_\alpha(\mathbf{x}, \bar{\lambda}, \mu) = f(\mathbf{x}) + \mu \sum_{i=1}^r \gamma(\bar{\lambda}_i, \alpha) \psi\left((\mu\gamma(\bar{\lambda}_i, \alpha))^{-1} \bar{\lambda}_i h_i(\mathbf{x})\right). \quad (2.6.8)$$

Definição 2.6.1. *Sejam $\gamma(\bar{\lambda}_i, \alpha)$ o fator de transformação e $\alpha \in [0, 1]$. A função $L_\alpha(\mathbf{x}, \bar{\lambda}, \mu)$ apresentada em (2.6.8) é denominada “função lagrangiana aumentada integrada”.* $\square$

Com referência aos autores em [Mاتيoli e Gonzaga 2008], apresentamos na Definição 2.6.2 uma nova classe de funções penalidades para os métodos de lagrangiano aumentado.

Definição 2.6.2. Sejam $\psi \in \mathbb{H}$, $\alpha \in [0, 1]$ e o fator de transformação $\gamma(\lambda, \alpha)$. Denominamos por classe $\mathbb{P}_{\text{int}}$ o conjunto das funções penalidades integrada $\eta_\psi : (\mathbb{T} \times \mathbb{R}_{++} \times [0, 1]) \rightarrow \mathbb{R}$ dadas por $\eta_\psi(t, \lambda, \alpha) = \gamma(\lambda, \alpha) \psi(\gamma(\lambda, \alpha)^{-1} \lambda t)$. $\square$

Consideremos as funções barreira logarítmica modificada $\psi(t) = -\ln(-t + 1)$ [Polyak 1992] e exponencial $\psi(t) = \exp(t) - 1$ [Kort e Bertsekas 1973]. Então, embasados na Definição 2.6.2, mostramos na Tabela 2.6.1 e na Tabela 2.6.2 alguns exemplos de funções penalidades da classe $\mathbb{P}_{\text{int}}$ geradas por meio destas funções.

Tabela 2.6.1: Funções penalidades do tipo barreira logarítmica modificada integrada (IMBF).

Fator de transformação $\gamma(\lambda, \alpha)$	$\psi(t) = -\ln(-t + 1)$
$\lambda(\alpha + (1 - \alpha)\lambda)^{-1}$	$\eta_\psi(t, \lambda, \alpha) = -\frac{\lambda}{(\alpha + (1 - \alpha)\lambda)} \ln(-(\alpha + (1 - \alpha)\lambda)t + 1)$
λ^α	$\eta_\psi(t, \lambda, \alpha) = -\lambda^\alpha \ln(-\lambda^{1-\alpha}t + 1)$
$1 + (\lambda - 1)\alpha$	$\eta_\psi(t, \lambda, \alpha) = -(1 + (\lambda - 1)\alpha) \ln\left(\frac{-\lambda t}{1 + (\lambda - 1)\alpha} + 1\right)$

Tabela 2.6.2: Funções penalidades do tipo exponencial integrada.

Fator de transformação $\gamma(\lambda, \alpha)$	$\psi(t) = \exp(t) + 1$
$\lambda(\alpha + (1 - \alpha)\lambda)^{-1}$	$\eta_\psi(t, \lambda, \alpha) = \frac{\lambda}{\alpha + (1 - \alpha)\lambda} (\exp((\alpha + (1 - \alpha)\lambda)t) + 1)$
λ^α	$\eta_\psi(t, \lambda, \alpha) = \lambda^\alpha (\exp(\lambda^{1-\alpha}t) + 1)$
$1 + (\lambda - 1)\alpha$	$\eta_\psi(t, \lambda, \alpha) = (1 + (\lambda - 1)\alpha) \left(\exp\left(\frac{\lambda t}{1 + (\lambda - 1)\alpha}\right) + 1\right)$

Com a finalidade de esclarecermos ao leitor o conceito de função e parâmetro de mudança de escala, consideremos a seguinte abordagem. Sabemos que o gráfico de uma função $q : \mathbb{R} \rightarrow \mathbb{R}$ real é o conjunto $\mathbb{U} = \{\mathbf{u} \in \mathbb{R}^2; \mathbf{u} = (x, q(x))\}$. Seja o escalar $v \in \mathbb{R}_{++}$. Então, a transformação de reescalamento dos elementos de $\mathbb{U}$ é conforme (2.6.9):

$$\begin{pmatrix} v & 0 \\ 0 & v \end{pmatrix} \begin{pmatrix} x \\ q(x) \end{pmatrix} = \begin{pmatrix} vx \\ vq(x) \end{pmatrix}. \quad (2.6.9)$$

Ao definirmos $t = vx$ e $q_v = vq(x)$, então, temos $x = v^{-1}t$ e $q_v = vq(v^{-1}t)$. Portanto, se $0 < v < 1$, então q_v é interpretada como uma contração de q em v unidades; se $v = 1$, então q_v é idêntica à q e; se $v > 1$, então q_v significa uma dilatação de q em v unidades. Nesse sentido, o escalar $v \in \mathbb{R}_{++}$ é denominado “parâmetro de reescala” e a função q_v é denominada “função de reescalamento”.

Em [Matioli e Gonzaga 2008], os autores denominaram por classe de funções penalidades de tipo 1 o conjunto $\mathbb{P}_1$ de funções $\eta_\psi(t, \lambda) = \psi(\lambda t)$ e por $\mathbb{P}_2$ a classe de funções penalidades de tipo 2 dadas por $\eta_\psi(t, \lambda) = \lambda\psi(t)$. Em nossa abordagem, a classe $\mathbb{P}_{\text{int}}$ é o conjunto de todas as funções $\eta_\psi \in \mathbb{P}_1$

em que tanto o domínio quanto a imagem de η_ψ são reescaladas em $\gamma(\lambda, \alpha)$ unidades para qualquer $\alpha \in [0, 1]$. Em particular, se $\alpha = 1$, temos que $\mathbb{P}_{\text{int}}$ é idêntica à classe $\mathbb{P}_2$.

Por fim, segue a Definição 2.6.3, a qual apresentamos os conceitos de função de reescalamento integrado e de parâmetro de reescala.

Definição 2.6.3. Sejam $\eta_\psi \in \mathbb{P}_{\text{int}}$ o escalar $\mu \in \mathbb{R}_{++}$. A função $\eta_{\psi, \mu} : \mathbb{T} \times \mathbb{R}_{++}^2 \times [0, 1] \rightarrow \mathbb{R}$ dada por:

$$\eta_{\psi, \mu}(t, \lambda, \mu, \alpha) = \mu \gamma(\lambda, \alpha) \psi \left((\mu \gamma(\lambda, \alpha))^{-1} \lambda t \right) \quad (2.6.10)$$

é denominada função de reescalamento integrado e μ é denominado parâmetro de penalidade ou parâmetro de reescala. $\square$

2.6.2 O método RNLI e suas propriedades

O método RNLI é definido por:

$$\mathbf{x}^{\mathbf{k}+1} = \arg \min \left\{ f(\mathbf{x}) + \mu_k \sum_{i=1}^r \gamma(\lambda_{i_k}, \alpha) \psi \left((\mu_k \gamma(\lambda_{i_k}, \alpha))^{-1} \lambda_{i_k} h_i(\mathbf{x}) \right), \mathbf{x} \in \mathbb{X} \right\} \quad (2.6.11)$$

$$\lambda^{\mathbf{k}+1} = \arg \max \left\{ f(\mathbf{x}^{\mathbf{k}+1}) + \lambda^T \mathbf{h}(\mathbf{x}^{\mathbf{k}+1}) - \mu_k \sum_{i=1}^r \gamma(\lambda_{i_k}, \alpha) \varphi \left(\lambda_{i_k}^{-1} \lambda_i \right), \lambda \in \mathbb{R}_{++}^r \right\}.$$

A solução analítica para a segunda expressão do algoritmo (2.6.11) é apresentada em (2.6.7). Assim, o algoritmo (2.6.11) é reescrito de acordo com (2.6.12) a seguir:

$$\mathbf{x}^{\mathbf{k}+1} = \arg \min \left\{ f(\mathbf{x}) + \mu_k \sum_{i=1}^r \gamma(\lambda_{i_k}, \alpha) \psi \left((\mu_k \gamma(\lambda_{i_k}, \alpha))^{-1} \lambda_{i_k} h_i(\mathbf{x}) \right), \mathbf{x} \in \mathbb{X} \right\} \quad (2.6.12)$$

$$\lambda_{i_{k+1}} = \psi' \left((\mu_k \gamma(\lambda_{i_k}, \alpha))^{-1} \lambda_{i_k} h_i(\mathbf{x}^{\mathbf{k}+1}) \right) \lambda_{i_k}, \quad \forall i = 1, \dots, r.$$

Estabelecemos no Teorema 2.6.1 a estreita relação que existe entre o método RNLI e o método dual de ponto proximal (2.5.5).

Teorema 2.6.1. O método de reescalamento não-linear integrado em (2.6.12) é equivalente ao método dual de ponto proximal (2.5.5). —

Demonstração. Pela segunda expressão em (2.6.11), temos:

$$\frac{\lambda_{i_{k+1}}}{\lambda_{i_k}} = \psi' \left(\frac{\lambda_{i_k} h_i(\mathbf{x}^{\mathbf{k}+1})}{\mu_k \gamma(\lambda_{i_k}, \alpha)} \right), \quad \forall i = 1, \dots, r.$$

Como $\psi \in \mathbb{H}$, temos que ψ'^{-1} é bem definida. Logo, temos:

$$h_i(\mathbf{x}^{\mathbf{k}+1}) - \frac{\mu_k \gamma(\lambda_{i_k}, \alpha)}{\lambda_{i_k}} \psi'^{-1} \left(\frac{\lambda_{i_{k+1}}}{\lambda_{i_k}} \right) = 0, \quad \forall i = 1, \dots, r.$$

A Proposição 2.6.1 nos assegura que o conjugado convexo de $\psi \in \mathbb{H}$ é a função $\varphi \in \mathbb{F}$. Logo, ao aplicarmos o análogo da relação fundamental (2.6.5), isto é, $\psi'^{-1} = \varphi'$, temos:

$$h_i(\mathbf{x}^{\mathbf{k}+1}) - \frac{\mu_k \gamma(\lambda_{i_k}, \alpha)}{\lambda_{i_k}} \varphi' \left(\frac{\lambda_{i_{k+1}}}{\lambda_{i_k}} \right) = 0, \quad \forall i = 1, \dots, r. \quad (2.6.13)$$

O fato de $\mathbf{x}^{k+1}$ ser um minimizador da função lagrangiana aumentada integrada $L_\alpha(\mathbf{x}, \lambda^k, \mu_k)$, temos que a condição suficiente² de primeira ordem dada por:

$$\nabla_x L_\alpha(\mathbf{x}^{k+1}, \lambda^k, \mu_k) = \nabla f(\mathbf{x}^{k+1}) + \sum_{i=1}^r \lambda_{i_k} \psi' \left(\frac{\lambda_{i_k} h_i(\mathbf{x}^{k+1})}{\mu_k \gamma(\lambda_{i_k}, \alpha)} \right) \nabla h_i(\mathbf{x}^{k+1}) = \mathbf{0} \quad (2.6.14)$$

é consistente. Novamente pela segunda expressão de (2.6.12), temos que (2.6.14) é equivalente a (2.6.15) a seguir:

$$\nabla_x L(\mathbf{x}^{k+1}, \lambda^{k+1}) = \nabla f(\mathbf{x}^{k+1}) + \sum_{i=1}^r \lambda_{i_{k+1}} \nabla h_i(\mathbf{x}^{k+1}) = \mathbf{0}. \quad (2.6.15)$$

Este resultado significa que $\mathbf{x}^{k+1} = \arg \min \left\{ f(\mathbf{x}) + \sum_{i=1}^r \lambda_{i_{k+1}} h_i(\mathbf{x}), \mathbf{x} \in \mathbb{X} \right\}$, isto é, $\mathbf{x}^{k+1} \in \mathbb{X}(\lambda^{k+1})$, e portanto, temos que $\mathbf{h}(\mathbf{x}^{k+1}) \in \partial d(\lambda^{k+1})$.

Assim, concluímos que a expressão (2.6.13) representa as condições suficientes de primeira ordem do problema $\max \left\{ d(\lambda) - \mu_k \sum_{i=1}^r \gamma(\lambda_{i_k}, \alpha) \varphi(\lambda_{i_k}^{-1} \lambda_i), \lambda \in \mathbb{R}_{++}^r \right\}$ para o qual λ^{k+1} é o argumento máximo procurado. ■

O algoritmo do método RNLI é sistematizado a seguir:

Algoritmo 2.6.1 Método de reescalamamento não-linear integrado.

Procedimento Inicial:

Escolha um $\alpha \in [0, 1]$ e uma função $\psi \in \mathbb{H}$. Além disso, escolha $\varepsilon > 0$, $\mu_0 > 0$, $\alpha_\mu \in (0, 1)$, $\lambda^0 \in \mathbb{R}_{++}^r$ e um $\gamma(\lambda_{i_0}, \alpha)$. Defina $k = 0$ e vá para o procedimento principal.

Procedimento Principal:

Passo 1: Determine $\mathbf{x}^{k+1} = \arg \min \left\{ f(\mathbf{x}) + \mu_k \sum_{i=1}^r \gamma(\lambda_{i_k}, \alpha) \psi((\mu_k \gamma(\lambda_{i_k}, \alpha))^{-1} \lambda_{i_k} h_i(\mathbf{x})); \mathbf{x} \in \mathbb{X} \right\}$;

Passo 2: Faça $\lambda_{i_{k+1}} = \psi'((\mu_k \gamma(\lambda_{i_k}, \alpha))^{-1} \lambda_{i_k} h_i(\mathbf{x}^{k+1})) \lambda_{i_k}$, para todo $i = 1, \dots, r$;

Passo 3: Se $|d(\lambda^{k+1}) - f(\mathbf{x}^{k+1})| < \varepsilon$, então **pare**. Caso contrário, vá para o *Passo 4*.

Passo 4: Faça $\lambda^k \leftarrow \lambda^{k+1}$, $\mu_{k+1} \leftarrow \alpha_\mu \mu_k$, $k \leftarrow k + 1$ e volte ao *Passo 1*.

Com procedimentos de análise análogos aos apresentados no Teorema 4.1 em [Matioli e Gonzaga 2008], o próximo resultado associado à satisfação da condição de folgas complementares pode ser estabelecido.

Proposição 2.6.2 (Convergência). *Sejam $\{\mu_k\}$ uma sequência monotonicamente decrescente e $\{\lambda^k\}$ e $\{\lambda^k\}$ respectivamente a sequência primal e dual gerada pelo Algoritmo 2.6.1 aplicado ao problema primal (2.4.1) com a função penalidade $\eta_\psi(t, \lambda, \alpha) = \gamma(\lambda, \alpha) \psi(\gamma(\lambda, \alpha)^{-1} \lambda t)$ da classe $\mathbb{P}_{\text{int}}$. Então, para qualquer parâmetro de homotopia $\alpha \in [0, 1]$ e fator de transformação $\gamma(\lambda_{i_k}, \alpha)$, $\lim_{k \rightarrow \infty} (\mu_k^{-1} \gamma(\lambda_{i_k}, \alpha))^{-1} \lambda_{i_k} h_i(\mathbf{x}^{k+1}) = 0$, para todo $i = 1, \dots, r$.*

²Por hipótese, o problema primal (2.4.1) é convexo.

Agora, vamos apresentar alguns métodos particulares do método RNLI. Cada um destes três métodos diferem-se mutuamente apenas pela escolha do fator de transformação vinculado à medida de distância integrada $\mathcal{H}_\varphi$.

- Se impusermos que o fator de transformação seja $\bar{\lambda}_i(\alpha + (1 - \alpha) \bar{\lambda}_i)^{-1}$, temos que a função lagrangiana aumentada integrada particulariza-se em:

$$L_\alpha(\mathbf{x}, \bar{\lambda}, \mu) = f(\mathbf{x}) + \mu \sum_{i=1}^r \frac{\bar{\lambda}_i}{(\alpha + (1 - \alpha) \bar{\lambda}_i)} \psi \left(\frac{(\alpha + (1 - \alpha) \bar{\lambda}_i) h_i(\mathbf{x})}{\mu} \right). \quad (2.6.16)$$

Assim, o método RNLI é definido por:

$$\begin{aligned} \mathbf{x}^{\mathbf{k}+1} &= \arg \min \left\{ f(\mathbf{x}) + \mu_k \sum_{i=1}^r \frac{\lambda_{i_k}}{(\alpha + (1 - \alpha) \lambda_{i_k})} \psi \left(\frac{(\alpha + (1 - \alpha) \lambda_{i_k}) h_i(\mathbf{x})}{\mu_k} \right), \mathbf{x} \in \mathbb{X} \right\} \\ \lambda_{i_{k+1}} &= \psi' \left(\frac{(\alpha + (1 - \alpha) \lambda_{i_k}) h_i(\mathbf{x}^{\mathbf{k}+1})}{\mu_k} \right) \lambda_{i_k}, \quad \forall i = 1, \dots, r. \end{aligned} \quad (2.6.17)$$

- Se impusermos que o fator de transformação seja $\bar{\lambda}_i^\alpha$, temos que a função lagrangiana aumentada integrada particulariza-se em:

$$L_\alpha(\mathbf{x}, \bar{\lambda}, \mu) = f(\mathbf{x}) + \mu \sum_{i=1}^r \bar{\lambda}_i^\alpha \psi \left(\frac{\bar{\lambda}_i^{1-\alpha} h_i(\mathbf{x})}{\mu} \right). \quad (2.6.18)$$

Assim, o método RNLI é definido por:

$$\begin{aligned} \mathbf{x}^{\mathbf{k}+1} &= \arg \min \left\{ f(\mathbf{x}) + \mu_k \sum_{i=1}^r \lambda_{i_k}^\alpha \psi \left(\frac{\lambda_{i_k}^{1-\alpha} h_i(\mathbf{x})}{\mu_k} \right), \mathbf{x} \in \mathbb{X} \right\} \\ \lambda_{i_{k+1}} &= \psi' \left(\frac{\lambda_{i_k}^{1-\alpha} h_i(\mathbf{x}^{\mathbf{k}+1})}{\mu_k} \right) \lambda_{i_k}, \quad \forall i = 1, \dots, r. \end{aligned} \quad (2.6.19)$$

- Se impusermos que o fator de transformação seja $1 + (\bar{\lambda}_i - 1) \alpha$, temos que a função lagrangiana aumentada integrada particulariza-se em:

$$L_\alpha(\mathbf{x}, \bar{\lambda}, \mu) = f(\mathbf{x}) + \mu \sum_{i=1}^r (1 + (\bar{\lambda}_i - 1) \alpha) \psi \left(\frac{\bar{\lambda}_i h_i(\mathbf{x})}{(\mu + \mu (\bar{\lambda}_i - 1) \alpha)} \right). \quad (2.6.20)$$

Assim, o método RNLI é definido por:

$$\begin{aligned} \mathbf{x}^{\mathbf{k}+1} &= \arg \min \left\{ f(\mathbf{x}) + \mu_k \sum_{i=1}^r (1 + (\lambda_{i_k} - 1) \alpha) \psi \left(\frac{\lambda_{i_k} h_i(\mathbf{x})}{(\mu_k + \mu_k (\lambda_{i_k} - 1) \alpha)} \right), \mathbf{x} \in \mathbb{X} \right\} \\ \lambda_{i_{k+1}} &= \psi' \left(\frac{\lambda_{i_k} h_i(\mathbf{x}^{\mathbf{k}+1})}{(\mu_k + \mu_k (\lambda_{i_k} - 1) \alpha)} \right) \lambda_{i_k}, \quad \forall i = 1, \dots, r. \end{aligned} \quad (2.6.21)$$

O método RNLI torna-se equivalente ao método de reescalamamento não-linear desenvolvido em [Polyak e Teboulle 1997] se o parâmetro de homotopia α é fixado em 1. De fato, se $\alpha = 1$, então o método RNLI particulariza-se em:

$$\begin{aligned} \mathbf{x}^{\mathbf{k}+1} &= \arg \min \left\{ f(\mathbf{x}) + \mu_k \sum_{i=1}^r \lambda_{i_k} \psi \left(\frac{h_i(\mathbf{x})}{\mu_k} \right), \mathbf{x} \in \mathbb{X} \right\} \\ \lambda_{i_{k+1}} &= \psi' \left(\frac{h_i(\mathbf{x}^{\mathbf{k}+1})}{\mu_k} \right) \lambda_{i_k}, \quad \forall i = 1, \dots, r, \end{aligned} \quad (2.6.22)$$

o qual é idêntico ao apresentado em [Polyak e Teboulle 1997, Expressões 2.2 e 2.3]. Por fim, o método RNLI particulariza-se no método de lagrangiano aumentado desenvolvido em [Matioli e Gonzaga 2008] se o parâmetro de homotopia α é fixado 0. De fato, se $\alpha = 0$, então o método RNLI particulariza-se em:

$$\begin{aligned} \mathbf{x}^{\mathbf{k}+1} &= \arg \min \left\{ f(\mathbf{x}) + \mu_k \sum_{i=1}^r \psi \left(\frac{\lambda_{i_k} h_i(\mathbf{x})}{\mu_k} \right), \mathbf{x} \in \mathbb{X} \right\} \\ \lambda_{i_{k+1}} &= \psi' \left(\frac{\lambda_{i_k} h_i(\mathbf{x}^{\mathbf{k}+1})}{\mu_k} \right) \lambda_{i_k}, \quad \forall i = 1, \dots, r, \end{aligned} \quad (2.6.23)$$

o qual é idêntico ao apresentado em [Polyak 2014, Expressões 8 e 9] e em [Matioli e Gonzaga 2008, Alg. 2.1 onde p é dado pela expressão 9].

Uma abordagem primal-dual do método RNLI é concebida a partir das expressões (2.6.11), (2.6.12) e (2.6.15). De fato, por meio delas temos que um ciclo completo do método RNLI é equivalente a resolver o seguinte sistema primal-dual na variável primal $\mathbf{x} = \mathbf{x}(\lambda^{\mathbf{k}}, \mu_k)$ e dual $\lambda = \lambda(\lambda^{\mathbf{k}}, \mu_k)$:

$$\begin{aligned} \nabla f(\mathbf{x}) + \sum_{i=1}^r \lambda_i \nabla h_i(\mathbf{x}) &= \mathbf{0} \\ \lambda - \text{diag} \left(\psi' \left(\frac{\lambda_{i_k} h_i(\mathbf{x})}{\mu_k \gamma(\lambda_{i_k}, \alpha)} \right) \right)_{i=1}^r \lambda^{\mathbf{k}} &= \mathbf{0}. \end{aligned} \quad (2.6.24)$$

Uma abordagem para lidarmos com o sistema não-linear (2.6.24) é apresentada em [Griva e Polyak 2006] para $\alpha = 1$.

Consideremos a expressão (2.6.13). Ao impormos $-\frac{\mu_k \gamma(\lambda_{i_k}, \alpha)}{\lambda_{i_k}} \varphi' \left(\frac{\lambda_{i_{k+1}}}{\lambda_{i_k}} \right) \geq 0$, para todo $i = 1, \dots, r$, então temos que esta desigualdade pode ser interpretada como um tipo de folga vinculada à restrição $h_i(\mathbf{x}) \leq 0$.

De fato, seja $z_i = -\frac{\mu \gamma(\bar{\lambda}_i, \alpha)}{\bar{\lambda}_i} \varphi' \left(\frac{\lambda_i}{\bar{\lambda}_i} \right)$ a componente do vetor $\mathbf{z} = (z_1, \dots, z_i, \dots, z_r)^T \in \mathbb{R}_+^r$, então, pela a relação fundamental $\varphi'^{-1} = \psi'$, obtemos $\lambda_i - \psi' \left(\frac{-\bar{\lambda}_i z_i}{\mu \gamma(\bar{\lambda}_i, \alpha)} \right) \bar{\lambda}_i = 0$. Logo, com referência às expressões (2.6.13) e (2.6.24) e ao impormos $\mu = \mu_k$ e $\bar{\lambda} = \lambda^{\mathbf{k}}$, temos que resolver um ciclo completo do método RNLI equivale resolvermos o sistema primal-dual dado por:

$$\begin{aligned} \nabla f(\mathbf{x}) + \sum_{i=1}^r \lambda_i \nabla h_i(\mathbf{x}) &= \mathbf{0} \\ \lambda - \text{diag} \left(\psi' \left(\frac{-\lambda_{i_k} z_i}{\mu_k \gamma(\lambda_{i_k}, \alpha)} \right) \right)_{i=1}^r \lambda^{\mathbf{k}} &= \mathbf{0} \\ \mathbf{h}(\mathbf{x}) + \mathbf{z} &= \mathbf{0} \end{aligned} \quad (2.6.25)$$

nas variáveis primais $\mathbf{x} = \mathbf{x}(\lambda^k, \mu_k)$ e $\mathbf{z} = \mathbf{z}(\lambda^k, \mu_k)$ e dual $\lambda = \lambda(\lambda^k, \mu_k)$. É imediato constatar-mos que o sistema não-linear (2.6.25) representa as condições de primeira ordem vinculadas à função lagrangiana aumentada integrada expressa por:

$$L_\alpha(\mathbf{x}, \mathbf{z}, \lambda, \bar{\lambda}, \mu) = f(\mathbf{x}) + \lambda^T (\mathbf{h}(\mathbf{x}) + \mathbf{z}) + \mu \sum_{i=1}^r \gamma(\bar{\lambda}_i, \alpha) \psi\left(\frac{-\bar{\lambda}_i z_i}{\mu \gamma(\bar{\lambda}_i, \alpha)}\right) \quad (2.6.26)$$

associada à solução do problema de programação não-linear convexo dado por:

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) \\ \text{s.a :} \quad & \mathbf{h}(\mathbf{x}) + \mathbf{z} = \mathbf{0} \\ & -\mathbf{z} \leq \mathbf{0}. \end{aligned} \quad (2.6.27)$$

Uma abordagem para lidarmos com o sistema não-linear (2.6.25) é apresentada em [Pineiro et al. 2015] para o caso em que $\alpha = 1$ e $\psi(t) = -\ln(-t + 1)$. No Capítulo 5, desenvolveremos uma abordagem primal-dual do método RNLI para lidarmos com o problema de programação não-linear discreta-mista com restrições de equilíbrio envolvendo qualquer função $\psi \in \mathbb{H}$ e parâmetro de homotopia $\alpha \in [0, 1]$.

2.7 Extensão: O método da função logarítmica-quadrática integrada

Nesta seção, apresentamos uma versão do método da função logarítmica-quadrática no sentido de método RNLI. A função logarítmica-quadrática foi desenvolvida em [Auslender, Teboulle e Ben-Tiba 1999] para um tipo de método de ponto proximal em problemas de desigualdade variacional. A aplicabilidade desta função foi estendida para os métodos de lagrangiano aumentado em [Auslender, Teboulle e Ben-Tiba 1999, Polyak 2006] e para problemas multi-objetivo em [Gregório e Oliveira 2010, Rocha et al. 2016].

Como veremos, a função $\psi(t)$ associada ao conjugado convexo da função logarítmica quadrática é bem diferente da função barreira logarítmica modificada $\psi(t) = -\ln(-t + 1)$, mas, ao contrário desta última, a nova função não necessita de técnicas de extrapolação quadrática ou exponencial [Polyak e Teboulle 1997, seção 6], pois não há pontos que extrapolam seu domínio.

A função logarítmica-quadrática é a função $\varphi \in \mathbb{F}$ definida por:

$$\varphi(s) = \frac{\beta_1(s-1)^2}{2} + \beta_2(-\ln(s) + s - 1), \quad (2.7.1)$$

onde $(\beta_1, \beta_2) \in (\mathbb{R}_+ \times \mathbb{R}_{++})$, e a medida de distância de Auslender é definida por:

$$\mathcal{D}_\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \left(\frac{\beta_1}{2} (x_i - y_i)^2 + \beta_2 \left(-y_i^2 \ln\left(\frac{x_i}{y_i}\right) + x_i y_i - y_i^2 \right) \right), \quad (2.7.2)$$

para algum par (β_1, β_2) [Auslender, Teboulle e Ben-Tiba 1999].

De acordo com autores, a parcela quadrática presente na expressão (2.7.1) é o termo clássico para designar regularização ou suavização no sentido de método de ponto proximal. Portanto, a função

logarítmica-quadrática é interpretada como uma regularização da função $\varphi(s) = -\ln(s) + s - 1$. Na Figura 2.7.1, a função $\varphi(s) = -\ln(s) + s - 1$ é posta em perspectiva com a sua forma regularizada.

Figura 2.7.1: A função logarítmica-quadrática (azul) em perspectiva com a função $\varphi(s) = -\ln(s) + s - 1$.

Neste trabalho, consideramos apenas o caso particular $\beta_1 = \beta_2 = \frac{1}{2}$. Assim sendo, temos que (2.7.1) é particularizada em:

$$\varphi(s) = \frac{1}{2} \left(-\ln(s) + \frac{s^2 - 1}{2} \right). \quad (2.7.3)$$

Além disso, distância de Auslender (2.7.2) é reduzida a:

$$\mathcal{D}_\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \frac{1}{2} \left(-y_i^2 \ln \left(\frac{x_i}{y_i} \right) + \frac{x_i^2 - y_i^2}{2} \right), \quad \forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_{++}^{2n}. \quad (2.7.4)$$

É imediato verificarmos que a função $\mathcal{D}_\varphi$ é equivalente à formulação (2.7.5), a qual não é nem uma distância de Bregman nem uma medida entrópica φ -divergente de Csiszar.

$$\mathcal{D}_\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n y_i^2 \varphi(y_i^{-1} x_i), \quad \forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_{++}^{2n}. \quad (2.7.5)$$

A proposta da função logarítmica-quadrática integrada é muito simples. Nós substituímos o fator y_i^2 da expressão (2.7.5) por y_i^α , com $\alpha \in [0, 1]$, para todo $i = 1, \dots, n$. Esta substituição modifica a medida de Auslender apresentada em (2.7.5) para um caso particular da medida de distância integrada proposta, Definição 2.3.1. Logo, para um caso geral, a medida de distância de Auslender é substituída pela medida de distância integrada apresentada em (2.7.6):

$$\mathcal{H}_\varphi(\mathbf{x}, \mathbf{y}, \alpha) = \sum_{i=1}^n \frac{\gamma(y_i, \alpha)}{2} \left(-\ln(y_i^{-1} x_i) + \frac{y_i^{-2} x_i^2 - 1}{2} \right), \quad (2.7.6)$$

onde $\gamma(y_i, \alpha)$ é um fator de transformação apresentado em (2.3.4).

Por fim, apresentamos no Teorema 2.7.1 o conjugado convexo da função logarítmica-quadrática (2.7.3).

Teorema 2.7.1. *Seja $\mathbb{F} \ni \varphi(s) = \frac{1}{2} \left(-\ln(s) + \frac{s^2-1}{2} \right)$ a função logarítmica-quadrática. Então o conjugado convexo de φ é a função:*

$$\mathbb{H} \ni \psi(t) = \frac{1}{2} \left(\arg \sinh(t) + t^2 + t\sqrt{t^2+1} \right). \quad (2.7.7)$$

Demonstração. Pela definição de conjugado convexo, $\psi(t)$ é o $\sup \{q(s), s \in \mathbb{R}_{++}\}$, em que $q(s) = st - \frac{1}{2} \left(-\ln(s) + \frac{s^2-1}{2} \right)$. A condição de primeira ordem é dada por:

$$q'(s) = t - \frac{1}{2} \left(-\frac{1}{s} + s \right) = 0 \Leftrightarrow -\frac{s^2}{2} + st + \frac{1}{2} = 0,$$

cujas solução ótima condizente com o domínio de φ é $s^* = t + \sqrt{t^2+1}$. Logo, temos:

$$\begin{aligned} \psi(t) = q(s^*) &= t \left(t + \sqrt{t^2+1} \right) - \frac{1}{2} \left(-\ln \left(t + \sqrt{t^2+1} \right) + \frac{(t+\sqrt{t^2+1})^2-1}{2} \right) \\ &= t^2 + t\sqrt{t^2+1} - \frac{1}{2} \left(-\ln \left(t + \sqrt{t^2+1} \right) + \frac{2t^2+2t\sqrt{t^2+1}}{2} \right) \\ &= t^2 + t\sqrt{t^2+1} + \frac{1}{2} \ln \left(t + \sqrt{t^2+1} \right) - \frac{t^2+t\sqrt{t^2+1}}{2} \\ &= \frac{1}{2} \left(\ln \left(t + \sqrt{t^2+1} \right) + t^2 + t\sqrt{t^2+1} \right). \end{aligned}$$

Todavia, temos que $\arg \sinh(t) = \ln \left(t + \sqrt{t^2+1} \right)$. Portanto, a expressão (2.7.7) procede. ■

As propriedades de $\psi \in \mathbb{H}$ são: o domínio $\mathbb{T} = \mathbb{R}$; trata-se de uma função convexa imprópria, isto é, $\psi(t) \geq -\infty$, para todo $t \in \mathbb{T}$; $\psi'(t) = t + \sqrt{t^2+1}$; $\psi''(t) = 1 + \frac{t}{\sqrt{t^2+1}}$; $\lim_{t \rightarrow \infty} t^{-1}\psi(t) = \infty$ (coercividade à direita). O gráfico da função $\psi(t)$ é apresentado na Figura 2.7.2 e é posto em perspectiva com o gráfico da função $\psi(t) = \exp(t) - 1$, a qual é uma função coerciva à direita e é própria, isto é, para esta função temos que $\psi(t) > -1$, para todo $t \in \mathbb{T} = \mathbb{R}$.

Com a função $\psi(t)$ apresentada em (2.7.7) aplicada ao Algoritmo 2.6.1 do método de reescalamiento não-linear integrado, obtemos um método de lagrangiano aumentado substancialmente diferente ao de Auslender. Além disso, apesar de $\psi(t)$ ser bem diferente da função barreira logarítmica modificada, a forma regularizada da função logarítmica modificada não necessita de técnicas de extrapolação quadrática ou exponencial [Polyak e Teboulle 1997, seção 6], pois não há pontos que extrapolam seu domínio.

Figura 2.7.2: Conjugado convexo da função logarítmica-quadrática (vermelho) em perspectiva com a função barreira logarítmica modificada.

Método de penalidades aplicado à programação discreta

O propósito deste capítulo é desenvolver uma variante do método de penalidades clássico para resolvermos problemas de programação discreta¹.

As abordagens baseadas em métodos de penalidade consistem em resolver o problema discreto por meio de uma sequência de problemas de programação não-linear em variáveis contínuas que convergem para uma solução do problema discreto original. Cada problema de programação não-linear desta sequência de problemas é definido da seguinte forma: as variáveis discretas são tratadas como variáveis contínuas e restritas em intervalo real compacto e; são adicionados termos de penalidade associados a estas variáveis na função objetivo do problema original, com o intuito de penalizar os valores não-discretos para essas variáveis. O que torna o método de penalidades atraente para os propósitos deste trabalho é a possibilidade de aplicarmos ao problema de programação não-linear associado à sequência de problemas alguma abordagem do tipo *gradiente* para resolvê-lo.

Embasados neste princípio, os autores em [Rinaldi 2009, Lucidi e Rinaldi 2010, Santis e Rinaldi 2011] propõem métodos de penalidades para lidarmos com problemas inteiros 0-1. Em [Rao 2009, p. 619], o autor descreve um método de penalidades em que as variáveis discretas são penalizadas por meio de uma função polinomial definida por partes; em [Soler, Asada e Costa 2013] e [Lage 2013, seção 2.1.1], os autores apresentam métodos de penalidades em que as variáveis discretas são penalizadas por meio de uma função polinomial aplicada a uma função trigonométrica.

Algumas funções penalidades não-quadráticas são apresentadas em [Bertsekas 1996, Birgin, Castillo e Martínez 2005]. Todavia, desconhecemos aplicações de tais funções no contexto da programação discreta.

Assim, em síntese, a variante do método de penalidade clássico consiste em utilizar uma classe de funções não-quadráticas adequada com o intuito de penalizar um funcional não-linear cujos zeros são valores discretos de interesse que uma variável contínua deve assumir. Além disso, o método de reescalamento não-linear integrado é a abordagem adotada para assegurar a viabilidade dessas variáveis contínuas.

Este capítulo está organizado da seguinte forma:

¹Programação discreta também é conhecida por programação combinatorial ou por programação discreta-mista.

- Na Seção 3.1, apresentamos a formulação do problema de programação discreta e seus pressupostos.
- Na Seção 3.2, apresentamos uma estratégia geral que nos permite transformar o problema discreto descrito na Seção 3.1 em um problema de programação não-linear contínuo e equivalente. Para isso, a minimização da função objetivo é feita em variáveis contínuas e esta está sujeita a um conjunto artificial de restrições de igualdade e de desigualdade adequado. Essas restrições são satisfeitas se, e somente, a variável contínua assume algum valor discreto vinculado ao problema original. Propomos o conceito de função interpolante hiperbólica definida por partes. Esta função possui propriedades analíticas que nos permitem controlar o fenômeno de Runge, o qual é comum em processos de interpolação, e utilizá-la como restrição artificial de igualdade.
- Na Seção 3.3, apresentamos o método de penalidades. Em um primeiro estudo, definimos a classe de funções penalidades que nos propicia tratar as restrições de igualdade artificiais apresentadas na Seção 3.2. Apresentamos exemplos de funções clássicas, não-quadráticas e os conceitos de funções- $q_{\inf}$ e de função barreira y-logarítmica. Em um segundo estudo, discutimos brevemente as dificuldades computacionais associadas ao método e apresentamos melhorias.

3.1 Declaração do problema e premissas básicas

O problema de programação discreta discutido neste Capítulo é apresentado em (3.1.1).

$$\begin{aligned} \text{Min } f(\mathbf{x}) \\ x_i \in \mathbb{D}_i \subset \mathbb{R}, \forall i = 1, \dots, n, \end{aligned} \quad (3.1.1)$$

onde $\mathbb{D}_i = \{d_{i,1}, \dots, d_{i,q}, \dots, d_{i,p_i}\}$ é o conjunto viável associado à variável x_i e $\mathbb{D} = \times_{i=1}^n \mathbb{D}_i$ é o conjunto viável. Sobre o problema (3.1.1), declaramos as seguintes premissas básicas:

Premissa 1: Os elementos de $\mathbb{D}_i$ são mutuamente distintos e;

Premissa 2: O cardinal de $\mathbb{D}_i$ é o número $2 \leq p_i < \infty$, para todo $i = 1, \dots, n$.

3.2 Transformação de um problema discreto em um contínuo equivalente

Consideramos a seguinte classe de funções:

Definição 3.2.1. Seja $\mathbb{M}$ o conjunto das funções $m : \mathbb{R} \rightarrow \mathbb{R}$ que possuem as seguintes propriedades:

- 1) $m(z)$ é pelo menos de classe C^1 ;
- 2) $m(x_i) = 0$ se, e somente se, $x_i \in \mathbb{D}_i$, para todo $i = 1, \dots, n$ e;
- 3) se $x_i \in \mathbb{D}_i$, então o valor absoluto de $\frac{\partial m(x_i)}{\partial x_i}$ é 1.

□

Uma maneira simples de transformarmos o problema (3.1.1) em um problema de programação não-linear equivalente com restrições artificiais de igualdade e de desigualdade é dado por:

$$\begin{aligned} \text{Min } & f(\mathbf{x}) \\ & m_i(x_i) = 0, \forall i = 1, \dots, n. \\ & d_{i,1} \leq x_i \leq d_{i,p_i} \end{aligned} \quad (3.2.1)$$

O conjunto viável $\mathbb{X} = \{\mathbf{x} \in \mathbb{R}^n; m_i(x_i) = 0 \text{ e } d_{i,1} \leq x_i \leq d_{i,p_i}, \forall i = 1, \dots, n\}$ associado ao problema (3.2.1) é idêntico ao conjunto $\mathbb{D}$ devido à definição da função m e este é evidentemente compacto. A escolha adequada de uma função m_i é de acordo com o arranjo dos elementos do conjunto $\mathbb{D}_i$. Ao assumirmos o arranjo $d_{i,1} < \dots < d_{i,q} < \dots < d_{i,p_i}$ segue, por exemplo, que se a distribuição é identificada como uma progressão aritmética de razão $r \neq 0$, então a função:

$$m_i(x_i) = r\pi^{-1} \text{sen}(r^{-1}\pi(x_i - d_{i,q})) \quad (3.2.2)$$

é uma possibilidade. A função apresentada em (3.2.2) é de classe C^∞ e limitada. Tais propriedades não são simultaneamente verificadas por qualquer interpolação polinomial². Um caso particular interessante é se $x_i \in \mathbb{Z}$. Se utilizarmos a expressão (3.2.2) com $r = 1$, segue imediatamente que $m_i(x_i) = \pi^{-1} \text{sen}(\pi(x_i - d_{i,q}))$ é anulada para todo x_i inteiro.

Uma das dificuldades para lidarmos com as funções interpolantes do tipo polinomial é a possibilidade de ocorrer um severo fenômeno de Runge quando um conjunto discreto $\mathbb{D}_i$ possui muitos pontos [Ruggiero e Lopes 1996]. Em [Lage 2013], o autor nos apresenta uma alternativa interessante às técnicas de interpolação polinomial ao utilizar, indiretamente, uma função interpolante trigonométrica definida por partes (Função-ITDP). Esta estratégia é a função $\bar{m}_i(x_i)$ segundo (3.2.3), a seguir:

$$\bar{m}_i(x_i) = \sum_{q=1}^{p_i-1} (-1)^q \text{sen}\left(\frac{\pi}{d_{i,q+1} - d_{i,q}}(x_i - d_{i,q})\right) (u_{d_{i,q}}(x_i) - u_{d_{i,q+1}}(x_i)), \quad (3.2.3)$$

onde:

$$u_a(x) = \begin{cases} 0 & , \text{ se } x < a \\ 1 & , \text{ se } x \geq a \end{cases}$$

é a função degrau de Heaviside.

A menos que a distribuição do arranjo dos elementos do conjunto $\mathbb{D}_i$ seja uma progressão aritmética, então a derivada de primeira ordem da $\bar{m}_i(x_i)$ apresentada em (3.2.3) é descontínua por partes; logo, em geral, a função $\bar{m}_i(x_i)$ não é de classe C^1 sobre o intervalo $d_{i,1} < x_i < d_{i,p_i}$. De fato, consideremos o conjunto discreto $\mathbb{D}_i = \{-2, 0, 1, 3\}$ [Lage 2013, p. 97]. Então, $\bar{m}_i(x_i)$ é dada por:

$$\bar{m}_i(x_i) = \begin{cases} 0 & , \text{ se } x_i < -2 \\ \text{sen}\left(\frac{\pi}{2}(x_i + 2)\right) & , \text{ se } -2 \leq x_i < 0 \\ -\text{sen}(\pi x_i) & , \text{ se } 0 \leq x_i < 1 \\ \text{sen}\left(\frac{\pi}{2}(x_i - 1)\right) & , \text{ se } 1 \leq x_i < 3 \\ 0 & , \text{ se } x_i \geq 3. \end{cases}$$

²Vide diversos métodos de interpolação polinomial em [Franco 2007, cap. 10].

É evidente que $\bar{m}_i(x_i)$ é contínua para todo $x_i \in \mathbb{R}$. Todavia, $\lim_{x_i \rightarrow 0^+} \frac{\pi}{2} \cos\left(\frac{\pi}{2}(x_i + 2)\right) = -\frac{\pi}{2}$, ao passo que $\lim_{x_i \rightarrow 0^-} -\pi \cos(\pi x_i) = -\pi$. Logo, a derivada da função $\bar{m}_i(x_i)$ apresentada em (3.2.3) é descontínua por partes.

Reparamos o problema da descontinuidade da derivada de $\bar{m}_i(x_i)$ com a Função-ITDP $m_i(x_i)$ dada por:

$$m_i(x_i) = \sum_{q=1}^{p_i-1} (-1)^q \frac{(d_{i,q+1} - d_{i,q})}{\pi} \operatorname{sen}\left(\frac{\pi}{d_{i,q+1} - d_{i,q}}(x_i - d_{i,q})\right) (u_{d_{i,q}}(x_i) - u_{d_{i,q+1}}(x_i)). \quad (3.2.4)$$

Consideremos novamente o conjunto discreto $\mathbb{D}_i = \{-2, 0, 1, 3\}$. Então, a função $m_i(x_i)$ associada a este conjunto é dada por:

$$m_i(x_i) = \begin{cases} 0 & , \text{ se } x_i < -2 \\ \frac{2}{\pi} \operatorname{sen}\left(\frac{\pi}{2}(x_i + 2)\right) & , \text{ se } -2 \leq x_i < 0 \\ -\pi^{-1} \operatorname{sen}(\pi x_i) & , \text{ se } 0 \leq x_i < 1 \\ \frac{2}{\pi} \operatorname{sen}\left(\frac{\pi}{2}(x_i - 1)\right) & , \text{ se } 1 \leq x_i < 3 \\ 0 & , \text{ se } x_i \geq 3. \end{cases}$$

Analisemos a continuidade da função $m_i'(x_i)$ na vizinhança do ponto $x_i = 0$. Imediatamente constatamos que $m_i'(0) = -\cos(\pi 0) = -1$ e que $\lim_{x_i \rightarrow 0^+} \cos\left(\frac{\pi}{2}(x_i + 2)\right) = -1 = \lim_{x_i \rightarrow 0^-} -\cos(\pi x_i)$. Logo, $m_i(x_i)$ é de classe C^1 .

A Função-ITDP, embora limitada, não elimina a possibilidade de ocorrer o fenômeno de Runge em relação a cada eixo dos x_i . Conforme o Teorema 3.2.1 a seguir mostra, quanto maior é a medida do intervalo $(d_{i,q}; d_{i,q+1})$, $q = 1, \dots, p_i$ maior será a magnitude da altura máxima da função senoidal em relação ao eixo dos x_i neste intervalo, isto é, maior é o fenômeno de Runge.

Teorema 3.2.1. *Seja $x_i \in \mathbb{D}_i = \{d_{i,1}, \dots, d_{i,q}, \dots, d_{i,p_i}\}$, com $d_{i,1} < \dots < d_{i,q} < \dots < d_{i,p_i}$, e a função interpolante trigonométrica definida por partes apresentada em (3.2.4). Então a magnitude máxima da função $|m_i(x_i)|$ no intervalo $[d_{i,q}; d_{i,q+1}] \subset \mathbb{R}$, dado qualquer $q = 1, \dots, p_i - 1$, é $\frac{d_{i,q+1} - d_{i,q}}{\pi}$. —*

Demonstração. Escolhamos qualquer $q = 1, \dots, p_i - 1$, então

$$|m_i(x_i)| = \frac{(d_{i,q+1} - d_{i,q})}{\pi} \operatorname{sen}\left(\frac{\pi}{d_{i,q+1} - d_{i,q}}(x_i - d_{i,q})\right) (u_{d_{i,q}}(x_i) - u_{d_{i,q+1}}(x_i))$$

para todo $x_i \in [d_{i,q}; d_{i,q+1}] \subset \mathbb{R}$. Como $|m_i(x_i)|$ é contínua, derivável e estritamente côncava em $(d_{i,q}; d_{i,q+1})$; temos que o maximizador global x_i^* é bem definido. Tal maximizador é o ponto $x_i^* = \frac{d_{i,q+1} + d_{i,q}}{2}$. Então temos:

$$\begin{aligned} |m_i(x_i^*)| &= \frac{(d_{i,q+1} - d_{i,q})}{\pi} \operatorname{sen}\left(\frac{\pi}{d_{i,q+1} - d_{i,q}}\left(\frac{d_{i,q+1} + d_{i,q}}{2} - d_{i,q}\right)\right) \\ &= \frac{(d_{i,q+1} - d_{i,q})}{\pi} \operatorname{sen}\left(\frac{\pi}{d_{i,q+1} - d_{i,q}}\left(\frac{d_{i,q+1} - d_{i,q}}{2}\right)\right) \\ &= \frac{(d_{i,q+1} - d_{i,q})}{\pi} \operatorname{sen}\left(\frac{\pi}{2}\right) \\ &= \frac{(d_{i,q+1} - d_{i,q})}{\pi}. \end{aligned}$$

Como os valores $d_{i,q}$ e $d_{i,q+1}$ são números reais distintos e arbitrários, então a proposição é verificada. ■

Com base nas expressões (3.2.3) e (3.2.4) e no Teorema 3.2.1, julgamos interessante propor uma função interpolante definida por partes $m_i(x_i)$ com a seguinte propriedade adicional:

- 4) Dado qualquer $q = 1, \dots, p_i - 1$ e para todo $\bar{x}_i \in [d_{i,q}, d_{i,q+1}]$; então, para algum $u \in \mathbb{R}_{++}$ e fixo, sempre é verdade que $|m_i(\bar{x}_i)| < u$.

Para esta finalidade, propomos o conceito da função interpolante hiperbólica definida por partes (Função-IHDP) dada por:

$$m_i(x_i) = \sum_{q=1}^{p_i-1} (-1)^q \left(\frac{\cosh(2x_i - d_{i,q} - d_{i,q+1})}{2 \sinh(d_{i,q} - d_{i,q+1})} - \frac{1}{2 \tanh(d_{i,q} - d_{i,q+1})} \right) (u_{d_{i,q}}(x_i) - u_{d_{i,q+1}}(x_i)). \quad (3.2.5)$$

Teorema 3.2.2. *Seja $x_i \in \mathbb{D}_i = \{d_{i,1}, \dots, d_{i,q}, \dots, d_{i,p_i}\}$, com $d_{i,1} < \dots < d_{i,q} < \dots < d_{i,p_i}$; e a função interpolante hiperbólica definida por partes apresentada em (3.2.5). Então a magnitude máxima da função $|m_i(x_i)|$ no intervalo $[d_{i,q}; d_{i,q+1}] \subset \mathbb{R}$, dado qualquer $q = 1, \dots, p_i - 1$, é $u = \frac{1}{2}$.* —

Demonstração. Escolhamos qualquer $q = 1, \dots, p_i - 1$, então

$$|m_i(x_i)| = \left(\frac{\cosh(2x_i - d_{i,q} - d_{i,q+1})}{2 \sinh(d_{i,q} - d_{i,q+1})} - \frac{1}{2 \tanh(d_{i,q} - d_{i,q+1})} \right) (u_{d_{i,q}}(x_i) - u_{d_{i,q+1}}(x_i))$$

para todo $x_i \in [d_{i,q}; d_{i,q+1}] \subset \mathbb{R}$. Como $|m_i(x_i)|$ é contínua, derivável e estritamente côncava em $(d_{i,q}; d_{i,q+1})$; temos que o maximizador global x_i^* é bem definido. Este maximizador é o ponto $x_i^* = \frac{d_{i,q+1} + d_{i,q}}{2}$. Então temos:

$$\begin{aligned} |m_i(x_i^*)| &= \frac{\cosh\left(2 \frac{d_{i,q+1} + d_{i,q}}{2} - d_{i,q} - d_{i,q+1}\right)}{2 \sinh(d_{i,q} - d_{i,q+1})} - \frac{1}{2 \tanh(d_{i,q} - d_{i,q+1})} \\ &= \frac{\cosh(0)}{2 \sinh(d_{i,q} - d_{i,q+1})} - \frac{1}{2 \tanh(d_{i,q} - d_{i,q+1})} \\ &= \frac{-\cosh(d_{i,q} - d_{i,q+1}) + 1}{2 \sinh(d_{i,q} - d_{i,q+1})}. \end{aligned}$$

Desde que $d_{i,q+1} > d_{i,q}$, então segue que $\sinh(d_{i,q} - d_{i,q+1}) = -\sinh(d_{i,q+1} - d_{i,q})$ e que $\cosh(d_{i,q} - d_{i,q+1}) = \cosh(d_{i,q+1} - d_{i,q})$. Assim, temos:

$$|m_i(x_i^*)| = \frac{\cosh(d_{i,q+1} - d_{i,q}) - 1}{2 \sinh(d_{i,q+1} - d_{i,q})} > 0.$$

Seja $t = d_{i,q+1} - d_{i,q}$ a medida do intervalo $[d_{i,q}, d_{i,q+1}]$ e a função $n(t) = \frac{\cosh(t) - 1}{2 \sinh(t)}$. Assim, temos:

$$\lim_{t \rightarrow 0^+} \frac{\cosh(t) - 1}{2 \sinh(t)} \stackrel{L'Hospital}{=} \lim_{t \rightarrow 0^+} \frac{\sinh(t)}{2 \cosh(t)} = \lim_{t \rightarrow 0^+} \frac{1}{2} \tanh(t) = 0$$

e

$$\lim_{t \rightarrow \infty} \frac{\cosh(t) - 1}{2 \sinh(t)} \stackrel{L'Hospital}{=} \lim_{t \rightarrow \infty} \frac{\sinh(t)}{2 \cosh(t)} = \lim_{t \rightarrow \infty} \frac{1}{2} \tanh(t) = \frac{1}{2}.$$

Como $d_{i,q}$ e $d_{i,q+1}$ são valores distintos e quaisquer e $m_i(x_i)$ é contínua nesses pontos, temos que $|m_i(x_i)|$ é limitada no intervalo $[0; \frac{1}{2})$ para qualquer que seja o intervalo se $[d_{i,q}, d_{i,q+1}]$. ■

Consideremos os seguintes conjuntos: $\mathbb{D}_1 = \{0, 1\}$, $\mathbb{D}_2 = \{0, 2\}$, $\mathbb{D}_3 = \{0, 4\}$, $\mathbb{D}_4 = \{0, 6\}$, $\mathbb{D}_5 = \{0, 8\}$ e $\mathbb{D}_6 = \{0, 10\}$. Ilustramos respectivamente nas Figuras 3.2.1 e 3.2.2 a aplicação das funções ITDP e IHDP sobre esses conjuntos para ilustrar os resultados apresentados obtidos nos Teoremas 3.2.1 e 3.2.2.

Figura 3.2.1: Aplicação da função interpolante trigonométrica definida por partes aos conjuntos $\mathbb{D}_1 = \{0, 1\}$, $\mathbb{D}_2 = \{0, 2\}$, $\mathbb{D}_3 = \{0, 4\}$, $\mathbb{D}_4 = \{0, 6\}$, $\mathbb{D}_5 = \{0, 8\}$ e $\mathbb{D}_6 = \{0, 10\}$.

Figura 3.2.2: Aplicação da função interpolante hiperbólica definida por partes aos conjuntos $\mathbb{D}_1 = \{0, 1\}$, $\mathbb{D}_2 = \{0, 2\}$, $\mathbb{D}_3 = \{0, 4\}$, $\mathbb{D}_4 = \{0, 6\}$, $\mathbb{D}_5 = \{0, 8\}$ e $\mathbb{D}_6 = \{0, 10\}$.

Em ambos os casos temos que os zeros das funções ITDP e IHDP são os valores discretos definidos pelos conjuntos $\mathbb{D}_i$, $i = 1, \dots, 6$. Entretanto, temos que a magnitude máxima da função IHDP não ultrapassa o valor $u = \frac{1}{2}$, conforme demonstrado no Teorema 3.2.2. Agora, vamos ilustrar respectivamente nas Figuras 3.2.1 e 3.2.2 o gráfico das derivadas destas funções.

Figura 3.2.3: Derivada da função interpolante trigonométrica definida por partes ao longo dos intervalos $\mathbb{I}_1 = [0, 1]$, $\mathbb{I}_2 = [0, 2]$, $\mathbb{I}_3 = [0, 4]$, $\mathbb{I}_4 = [0, 6]$, $\mathbb{I}_5 = [0, 8]$ e $\mathbb{I}_6 = [0, 10]$.

Figura 3.2.4: Derivada da função interpolante hiperbólica definida por partes ao longo dos intervalos $\mathbb{I}_1 = [0, 1]$, $\mathbb{I}_2 = [0, 2]$, $\mathbb{I}_3 = [0, 4]$, $\mathbb{I}_4 = [0, 6]$, $\mathbb{I}_5 = [0, 8]$ e $\mathbb{I}_6 = [0, 10]$.

Apesar dos gráficos das derivadas das funções ITDP e IHDP serem significativamente diferentes ao longo dos intervalos reais $\mathbb{I}_i$, para todo $i = 1, \dots, 6$, temos que a magnitude de $m'(x)$ nos extremos dos respectivos conjuntos discretos é 1. Portanto, ilustramos que a propriedade 3 da Definição 3.2.1 é satisfeita.

3.3 Método de Penalidades

Nesta seção, apresentamos uma variante do método de penalidades clássico como uma alternativa para lidarmos somente com as restrições de igualdade artificiais do problema (3.2.1). As restrições de desigualdade artificiais são tratadas com o método de reescalamento não-linear integrado. Para

leituras complementares sobre o método de penalidades clássico, vide [Bazaraa, Sherali e Shetty 2006, seção 9.2], [Martínez e Santos 1998, p. 172], [Martínez 2009, caps. 2 e 3] e [Griva, Nash e Sofer 2008, seção 16.2]. Esta seção está organizada da seguinte forma:

- na subseção 3.3.1 apresentamos as características da função penalidade, a qual é induzida por uma função $v \in \mathbb{G}$ a ser definida, e, sob certos critérios, uma proposta de gerarmos novas funções v , as quais são obtidas através de estratégias de regularização. Dentre as novas funções, destacamos a proposta da função barreira y -logarítmica.
- na subseção 3.3.2 descrevemos o método de penalidades. Discutimos brevemente as dificuldades computacionais e uma alternativa para melhorias.

3.3.1 Funções penalidades e extensão

Considere o problema de programação não-linear:

$$\begin{aligned} \text{Min } & f(\mathbf{x}) \\ \text{s.a : } & \mathbf{x} \in \mathbb{X}, \end{aligned} \quad (3.3.1)$$

onde $\mathbb{X} = \{\mathbf{x} \in \mathbb{R}^n; g_t(\mathbf{x}) = 0, \forall t = 1, \dots, m\}$ é o conjunto viável e também a função indicatriz dada por:

$$\eta(\mathbf{x}) = \begin{cases} 0 & , \text{ se } \mathbf{x} \in \mathbb{X} \\ +\infty & , \text{ se } \mathbf{x} \notin \mathbb{X} \end{cases} \quad (3.3.2)$$

a qual é interpretada como uma função penalidade ideal para lidarmos com pontos inviáveis. Pela definição da função indicatriz, o problema (3.3.1) é equivalente ao problema irrestrito dado por:

$$\text{Min } f(\mathbf{x}) + \eta(\mathbf{x}). \quad (3.3.3)$$

Poderíamos resolver o problema irrestrito (3.3.3) se a função objetivo $f(\mathbf{x}) + \eta(\mathbf{x})$ fosse definida fora da região viável. No entanto, mesmo que se substituíssemos o “ ∞ ” por um número muito grande, o problema irrestrito continuaria sendo difícil de resolver devido à descontinuidade. Por esses motivos, a função indicatriz é substituída por uma função penalidade contínua e parametrizada tal que esta aproxima-se da função indicatriz conforme o parâmetro cresce.

Neste trabalho, consideramos o conjunto $\mathbb{G}$ das funções $v : \mathbb{Y} \subseteq \mathbb{R} \rightarrow \mathbb{R}_+$ que possuem as propriedades:

- 1) $\mathbb{Y} = (a, b) \cup \{0\}$, com $-\infty \leq a < 0$ e $0 < b \leq \infty$;
- 2) v é pelo menos de classe C^2 ;
- 3) $v(y) = v'(y) = 0$ se, e somente se, $y = 0$;
- 4) $v'(y) < 0$ para todo $y < 0$, $v'(y) > 0$ para todo $y > 0$, $v''(y) > 0$ para todo $y \in \mathbb{Y}$ e;
- 5) $\lim_{y \rightarrow a^+} y^{-1}v(y) = \lim_{y \rightarrow b^-} y^{-1}v(y) = \infty$ (Coercividade à esquerda e à direita, respectivamente).

Ao impormos, por exemplo, $y_t = g_t(\mathbf{x})$; então a função indicatriz é aproximada pela função penalidade $\eta_\rho : \mathbb{R}^n \times \mathbb{R}_{++}^m \rightarrow \mathbb{R}_+$ dada por:

$$\eta_\rho(\mathbf{x}, \rho) = \sum_{t=1}^m \rho_t v(g_t(\mathbf{x})), \quad (3.3.4)$$

em que a componente ρ_i do vetor $\rho = (\rho_1, \dots, \rho_t, \dots, \rho_m)^T \subseteq \mathbb{R}_{++}^m$ é denominada parâmetro de penalidade. Por fim, resolver o problema irrestrito (3.3.3) torna-se equivalente a resolvermos uma sequência de problemas dada por:

$$\mathbf{x}^{k+1} = \arg \min \left\{ f(\mathbf{x}) + \sum_{t=1}^m \rho_{t_k} v(g_t(\mathbf{x})), \mathbf{x} \in \mathbb{R}^n \right\}, \quad (3.3.5)$$

com $\rho_{t_k} \rightarrow \infty$. A função $\phi(\mathbf{x}, \rho) = f(\mathbf{x}) + \sum_{t=1}^m \rho_t v(g_t(\mathbf{x}))$ é geralmente denominada função auxiliar.

Alguns exemplos de funções $v \in \mathbb{G}$ são as funções $v_1(y) = (2\beta)^{-1}|y|^{2\beta}$, com $\beta \in \mathbb{N}_{++}$, e $v_2(y) = \cosh(y) - 1$, as quais são apresentadas respectivamente em [Bazaraa, Sherali e Shetty 2006, p. 470], [Bertsekas 1996, p. 304].

A menos da coercividade à esquerda e à direita, outras funções de classe C^∞ poderiam estar em $\mathbb{G}$. Alguns exemplos de tais funções são: $q_1(y) = \ln(\cosh(y))$ e $q_2(y) = \frac{1}{2}y \arctan(y)$. Neste trabalho, lidamos com tais funções para gerarmos novas funções para o conjunto $\mathbb{G}$. A ideia é aplicarmos uma variante da estratégia de regularização de Auslender, vide Seção 2.7, com objetivo de gerarmos novas funções da classe $\mathbb{G}$.

Definição 3.3.1 (Função- $q_{\inf}$). *Sejam $v \in \mathbb{G}$ e $q : \mathbb{Y} \rightarrow \mathbb{R}_+$ uma função que possui as propriedades da função v exceto, possivelmente, a propriedade 5. A função de regularização*

$$\mathbb{G} \ni q_n(y) = q(y) + v(y) \quad (3.3.6)$$

é denominada “ q n -regularizada” ou “função- q_n ” se $v \in \mathbb{G}$ é de classe C^n , com $2 \leq n < \infty$ e de “ q infinitamente regularizada” ou “função- $q_{\inf}$ ” se $n = \infty$. □

Um exemplo de função $q_2(y)$ é dada por $q_2(y) = q(y) + \frac{1}{2}y^2$ e um exemplo de função $q_{\inf}(y)$ é dada por $q_{\inf}(y) = q(y) + \cosh(y) - 1$. As Figuras 3.3.1 e 3.3.2 ilustram respectivamente o efeito das funções $q_{\inf}$ e q_2 sobre a função $q(y) = \ln(\cosh(y))$.

Figura 3.3.1: A função $q(y) = \ln(\cosh(y))$ e as suas respectivas funções de regularização $q_{\inf}$ e q_2 .

Figura 3.3.2: As derivadas de primeira ordem de q e de suas respectivas funções de regularização.

Por fim, suponha o conjunto $\mathbb{Y} = (a, \infty)$, com a finito³. Então podemos adaptar as funções $\varphi \in \mathbb{F}$, vide Seção 2.2, para gerarmos novas funções para o conjunto $\mathbb{G}$. De fato, sejam $\varphi \in \mathbb{F}$ e um escalar $\kappa > -a$ (parâmetro de mudança de escala), então temos a seguinte função- $q_{\inf}$.

$$q_{\inf}(y, \kappa) = \kappa (\varphi(\kappa^{-1}y + 1) + v(\kappa^{-1}y)). \quad (3.3.7)$$

Um primeiro exemplo de aplicação desta nova função $q_{\inf}$ consiste em utilizarmos a função $\varphi_1(s) = -\ln(s) + s - 1$, a qual está em $\mathbb{F}$. Seja $s = y + 1$, então a função $\varphi_1(y) = q(y) = -\ln(y + 1) + y$ pode ser elemento do conjunto $\mathbb{G}$. Todavia, é fácil verificarmos que $\lim_{y \rightarrow (-1)^+} y^{-1}q(y) = \infty$ e que $\lim_{y \rightarrow \infty} y^{-1}q(y) < \infty$, logo, uma regularização é necessária. Agora, aplicamos a expressão

³Um raciocínio análogo pode ser se supusermos $\mathbb{Y} = (-\infty, b)$, com b finito.

(3.3.7) para gerarmos uma função- q_{inf} vinculada à $q(y)$, de acordo com (3.3.8):

$$q_{\text{inf}}(y, \kappa) = \kappa \left(-\ln(\kappa^{-1}y + 1) + \kappa^{-1}y + v(\kappa^{-1}y) \right). \quad (3.3.8)$$

Antes de prosseguirmos, consideremos a função entrópica de Kullback-Leibler $\varphi_2(s) = s \ln(s) - s + 1$. Para esta função, temos que $\lim_{s \rightarrow 0^+} s^{-1}\varphi_2(s) = \infty = \lim_{s \rightarrow \infty} s^{-1}\varphi_2(s)$. Assim, esta função entrópica é coerciva à esquerda e à direita. Com o intuito de torná-la uma função $v \in \mathbb{G}$, fazamos $s = y + 1$. Com esta aplicação, obtemos:

$$v(y) = (y + 1) \ln(y + 1) - y, \quad (3.3.9)$$

a qual não necessita de uma estratégia de regularização. Agora, consideremos a $q_{\text{inf}}(y, \kappa)$ apresentada (3.3.8) e a função entrópica $v(y, \kappa) = \kappa \left((\kappa^{-1}y + 1) \ln(\kappa^{-1}y + 1) - \kappa^{-1}y \right)$. Assim, temos que a expressão (3.3.8) particulariza-se em:

$$\begin{aligned} \mathbb{G} \ni q_{\text{inf}}(y, \kappa) &= \kappa \left(-\ln(\kappa^{-1}y + 1) + \kappa^{-1}y + (\kappa^{-1}y + 1) \ln(\kappa^{-1}y + 1) - \kappa^{-1}y \right) \\ &= y \ln(\kappa^{-1}y + 1). \end{aligned} \quad (3.3.10)$$

Definição 3.3.2. A função $v : \mathbb{Y} = (-1, \infty) \rightarrow \mathbb{R}_+$ dada por $v(y) = y \ln(y + 1)$ é denominada “função barreira y -logarítmica”. $\square$

Para avaliarmos a aplicabilidade da função barreira y -logarítmica como uma função penalidade, suponhamos que desejamos penalizar a restrição de igualdade $m(x) = 0$, onde $m(x) = \frac{1}{\pi} \sin(\pi x)$ (impomos x irrestrito por simplicidade). Como $|m(x)| < \frac{1}{\pi}$ para todo x , então a função barreira y -logarítmica é bem-definida sobre $m(x)$ ao atribuírmos qualquer valor $\kappa > \frac{1}{\pi}$. Ao escolhermos $\kappa = 1$, obtemos a função penalidade dada por:

$$\eta_\rho(x, \rho) = \rho \frac{1}{\pi} \sin(\pi x) \ln \left(\frac{1}{\pi} \sin(\pi x) + 1 \right).$$

Um outro exemplo segue se escolhermos a função $v(y) = \frac{1}{2}y^2$ para penalizarmos a mesma restrição de igualdade $m(x) = \frac{1}{\pi} \sin(\pi x) = 0$. Então, para este caso, obtemos a função penalidade dada por:

$$\eta_\rho(x, \rho) = \rho \frac{1}{2\pi^2} \sin(\pi x)^2,$$

a qual é exatamente a função penalidade senoidal apresentada em [Soler, Asada e Costa 2013]. A Figura 3.3.3 mostra o gráfico das funções penalidades induzidas pelas função quadrática e barreira y -logarítmica ao impormos $\rho = 1$.

Figura 3.3.3: As funções penalidades quadrática e barreira γ -logarítmica aplicada à função $\pi^{-1} \sin(\pi x)$.

3.3.2 Descrição do método de penalidades

Retornemos ao problema (3.2.1), o qual é reapresentado a seguir:

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) \\ & m_i(\mathbf{x}) = 0, \quad \forall i = 1, \dots, n. \\ & d_{i,1} \leq x_i \leq d_{i,p_i} \end{aligned}$$

Para aplicarmos o método de penalidades proposto, substituímos (3.2.1) pelo problema (3.3.11) a seguir:

$$\begin{aligned} \text{Min} \quad & \phi(\mathbf{x}, \rho) = f(\mathbf{x}) + \sum_{i=1}^n \rho_i v(m_i(x_i)) \\ \text{s.a.} \quad & d_{i,1} \leq x_i \leq d_{i,p_i}, \quad \forall i = 1, \dots, r. \end{aligned} \quad (3.3.11)$$

O problema (3.3.11) é evidentemente equivalente ao problema (3.3.12) a seguir:

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) + \sum_{i=1}^n \rho_i v(m_i(x_i)) \\ \text{s.a.} \quad & -x_i + d_{i,1} \leq 0, \quad \forall i = 1, \dots, n. \\ & x_i - d_{i,p_i} \leq 0 \end{aligned} \quad (3.3.12)$$

Seja o funcional $\mathbf{h}(\mathbf{x}) = (-x_1 + d_{1,1}, x_1 - d_{1,p_1}, \dots, -x_i + d_{i,1}, x_i - d_{i,p_i})^T \in \mathbb{R}^{2n}$, então, definimos $\bar{\mathbb{X}} = \{\mathbf{x} \in \mathbb{R}^n; \mathbf{h}(\mathbf{x}) \leq \mathbf{0}\}$ o conjunto viável associado ao problema (3.3.12) para evitarmos ambiguidades com o conjunto viável $\mathbb{X}$ vinculado ao problema (3.2.1).

Cada restrição de desigualdade $h_i(\mathbf{x}) \leq 0$ do problema (3.3.12) é tratada por meio do método de reescalamamento não-linear integrado. A função lagrangiana aumentada integrada associada ao problema (3.3.12) é definida por:

$$\begin{aligned} L_\alpha(\mathbf{x}, \bar{\lambda}, \rho, \mu) = & f(\mathbf{x}) + \sum_{i=1}^n \rho_i v(m_i(x_i)) + \\ & + \mu \sum_{i=1}^{2n} \gamma(\bar{\lambda}_i, \alpha) \psi\left((\mu \gamma(\bar{\lambda}_i, \alpha))^{-1} \bar{\lambda}_i h_i(\mathbf{x})\right), \end{aligned} \quad (3.3.13)$$

Descrevemos no Algoritmo 3.3.1 os passos que definem o método de penalidades proposto via técnica sequencial de minimizações irrestritas seguidas de atualizações dos multiplicadores de Lagrange e dos parâmetros de penalidades.

Algoritmo 3.3.1 : Método de penalidades para programação discreta.

Procedimento Inicial:

Escolha uma função $v \in \mathbb{G}$ e $\psi \in \mathbb{H}$, um $\alpha \in [0, 1]$ e ponto inicial $\mathbf{x}^0$ tal que $d_{i,1} \leq x_{i_0} \leq d_{i,p_i}$. Além disso, escolha $\varepsilon > 0$; os parâmetros $\rho_{0_i} \in \mathbb{R}_{++}$, para todo $i = 1, \dots, n$; $\alpha_{\rho_i} > 1$, $\mu_0 \in \mathbb{R}_{++}$, $\alpha_\mu \in (0, 1)$; um multiplicador de Lagrange λ_{i_0} e o fator de transformação $\gamma(\lambda_{i_0}, \alpha)$, para todo $i = 1, \dots, 2n$. Por fim, defina um escalar $\tau \in [0, 1)$ e $k = 0$. Vá para o procedimento principal.

Procedimento Principal:

Passo 1: Determine $\mathbf{x}^{k+1} = \arg \min \{L_\alpha(\mathbf{x}, \lambda^k, \rho^k, \mu_k), \mathbf{x} \in \bar{\mathbb{X}}\}$;

Passo 2: Se $|m_i(x_{i_{k+1}})| \leq \tau |m_i(x_{i_k})|$, então $\rho_{i_{k+1}} = \rho_{i_k}$. Caso contrário, faça $\rho_{i_{k+1}} = \alpha_{\rho_i} \rho_{i_k}$;

Passo 3: Faça $\mu_{k+1} = \alpha_\mu \mu_k$;

Passo 4: Faça $\lambda_{i_{k+1}} = \psi' \left(\frac{\lambda_{i_k} h_i(\mathbf{x}^{k+1})}{\mu_{k+1} \gamma(\lambda_{i_k}, \alpha)} \right) \lambda_{i_k}$, para todo $i = 1, \dots, 2n$;

Passo 4: Se $\left| \sum_{i=1}^n \rho_{i_{k+1}} m_i(x_{i_{k+1}}) + \mu_{k+1} \sum_{i=1}^n \gamma(\lambda_{i_k}, \alpha) \psi \left(\frac{\lambda_{i_{k+1}} h_i(\mathbf{x}^{k+1})}{\mu_{k+1} \gamma(\lambda_{i_k}, \alpha)} \right) \right| < \varepsilon$, então **pare**. Caso contrário, faça $k \leftarrow k + 1$ e volte ao Passo 1;

Observação 3.3.1. O Passo 2 do Algoritmo 3.3.1 é denominado estratégia de controle de viabilidade e é apresentada em [Martínez 2009, seção 2.2]. $\square$

Em geral, determinamos o ponto $\mathbf{x}^{k+1}$ através do método de Newton. Para apontarmos uma das principais dificuldades computacionais com o método de penalidades, vamos omitir as restrições de desigualdade do problema (3.3.11) sem ignorarmos que $\mathbf{x}$ está em $\bar{\mathbb{X}}$.

Através do método de Newton, temos $\mathbf{x}^{k+1} = \mathbf{x}^k - \beta [\nabla^2 \phi(\mathbf{x}^k, \rho^k)]^{-1} \nabla \phi(\mathbf{x}^k, \rho^k)$, onde $\beta \in (0, \bar{\beta})$ é o tamanho do passo;

$$\nabla \phi(\mathbf{x}^k, \rho^k) = \nabla f(\mathbf{x}^k) + \sum_{i=1}^n \rho_{i_k} v'(m_i(x_{i_k})) \frac{\partial m_i(x_{i_k})}{\partial x_i} \mathbf{e}_i \quad (3.3.14)$$

é o vetor gradiente da função auxiliar no ponto $(\mathbf{x}^k, \rho^k)$, $\mathbf{e}_i$ é a i -ésima coluna da matriz identidade $\mathbf{e}$;

$$\nabla^2 \phi(\mathbf{x}^k, \rho^k) = \nabla^2 f(\mathbf{x}^k) + \text{diag} \left(\rho_{i_k} v''(m_i(x_{i_k})) \left(\frac{\partial m_i(x_{i_k})}{\partial x_i} \right)^2 + \rho_{i_k} v'(m_i(x_{i_k})) \frac{\partial^2 m_i(x_{i_k})}{\partial x_i^2} \right)_{i=1}^n \quad (3.3.15)$$

é a matriz hessiana da função auxiliar no ponto $(\mathbf{x}^k, \rho^k)$. Se a matriz $\nabla^2 \phi(\mathbf{x}^k, \rho^k)$ é singular ou apresenta mau condicionamento numérico em alguma iteração k , então resolvemos este problema

com a estratégia de regularização apresentada na Proposição B.1.1, p. 221. Assumamos que $\bar{x}$ seja um ponto crítico e $\lim_{k \rightarrow \infty} x^k = \bar{x}$. Não é necessariamente verdade que $\|\nabla \phi(\bar{x}, \rho^k)\|_\infty = 0$ indica que $\bar{x}$ seja um minimizador local x^* com $\rho^k = \infty$, pois este pode ser um *maximizador*. Assim, mesmo que a matriz hessiana seja regularizada em um maximizador $\bar{x}$, tal matriz não é *suficientemente* definida positiva na vizinhança deste ponto. Isso significa que pelo menos uma componente do vetor $\bar{x}$ não assume valor discreto, ou seja, claramente $\bar{x}$ é um ponto inviável. Por consequência da inviabilidade deste ponto, temos que $\sum_{i=1}^n \rho_{i_k} v(m_i(\bar{x}_i)) = \infty$, logo, os procedimentos do algoritmo devem ser interrompidos.

Uma possível alternativa para tratarmos esta dificuldade computacional é aplicarmos o método de Levenberg-Marquardt⁴, o qual é uma proposta de melhoria do método de Newton. Se aplicado junto ao algoritmo do método de penalidades, o método de Levenberg-Marquardt consiste em transformar a matriz hessiana $\nabla^2 \phi(x^k, \rho^k)$ em uma outra matriz estritamente definida positiva em todas as iterações. Portanto, esta abordagem nos permite gerar direções de busca primal sempre de descida.

A principal consequência de sempre gerarmos uma sequência de direções de descida é obtermos $\lim_{k \rightarrow \infty} [\sum_{i=1}^n \rho_{i_k} v(m_i(x_{i_k}))] = 0$ com $\lim_{k \rightarrow \infty} \rho_{i_k} = \infty$; pois, para qualquer $\rho \in \mathbb{R}_{++}^n$ e pelo fato de existir um $\bar{x} \in \mathbb{X}$ tal que $m_i(\bar{x}_i) = 0$, então é verdade que $\inf \{\sum_{i=1}^n \rho_i v(m_i(x_i)), x \in \mathbb{X}\} = 0$. Portanto, como x^k está em $\mathbb{X}$, o qual é compacto, temos que o corolário do Teorema 9.2.2 em [Bazaraa, Sherali e Shetty 2006, p. 477] nos assegura que $\lim_{k \rightarrow \infty} x^k = \bar{x} = x^*$, isto é, nos assegura determinarmos um ponto discreto viável e que minimiza localmente a função objetivo associada ao problema original (3.1.1).

⁴Vide originais em [Levenberg 1944, Marquardt 1963] e uma descrição deste método em [Bazaraa, Sherali e Shetty 2006, p. 398].

Problemas de programação matemática com restrições de equilíbrio

Neste capítulo, apresentamos uma breve dissertação sobre o problema de programação matemática com restrições de complementaridade, o qual é um sub-ramo de uma classe geral de problemas de programação matemática com restrições de equilíbrio [Luo, Pang e Ralph 1996, Ferris e Pang 1997, Ye 2005].

Este capítulo está organizado da seguinte forma:

- Na Seção 4.1 é apresentada a formulação de uma classe especial de problemas de programação matemática com restrições de equilíbrio (MPEC) denominada problema de programação matemática com restrições de complementaridade (MPCC). Nós definimos o que são as restrições de complementaridade, algumas formulações que as representam por meio de restrições de igualdade e de desigualdade e uma interpretação geométrica. Apresentamos os conjuntos de restrições ativas e as diversas condições de estacionaridade de primeira ordem para o problema. Ao final desta seção, apresentamos sistemas de equações não-lineares, cujo significado é a representação matemática para cada uma das condições de estacionaridade.
- Na Seção 4.2 é apresentada uma extensão para o MPCC, cuja formulação contém o que é denominado restrições de complementaridade mista. Definimos o que são essas restrições, suas formulações por meio de restrições de igualdade e de desigualdade e apresentamos uma interpretação geométrica. Mostramos que se um minimizador resolve o problema otimização não-linear equivalente ao MPCC com restrições de complementaridade mista, então o sistema de KKT de primeira ordem associado ao problema de programação não-linear padrão é equivalente a uma condição de estacionaridade específica para o MPCC.

4.1 Programação matemática com restrições de complementaridade

Nesta seção, apresentamos o problema de programação matemática com restrições de complementaridade (MPCC) com suas principais definições. Esta seção está organizada da seguinte forma:

- Na Subseção 4.1.1, nós apresentamos a formulação do MPCC e as premissas básicas sobre o mesmo. Além disso, apresentamos algumas formulações matemáticas que representam as restrições de complementaridade como um conjunto de restrições de igualdade e de desigualdade. Por fim, apresentamos o conjunto das restrições ativas vinculado ao problema e algumas referências bibliográficas sobre investigação em teoria e métodos de solução.
- Na Subseção 4.1.2, nós apresentamos as diversas condições de estacionaridade e culmina com a apresentação de formulações matemáticas como um sistema de equações não-lineares que as representam.

4.1.1 Declaração do problema e outros conceitos

A formulação geral de um MPCC é apresentado em (4.1.1) [Flegel e Kanzow 2005].

$$\begin{aligned}
 & \text{Min } f(\mathbf{x}) \\
 & \text{s.a : } g_t(\mathbf{x}) = 0, \quad \forall t = 1, \dots, m \\
 & \quad h_i(\mathbf{x}) \leq 0, \quad \forall i = 1, \dots, r \\
 & \quad 0 \leq G_s(\mathbf{x}) \perp H_s(\mathbf{x}) \geq 0, \quad \forall s = 1, \dots, p
 \end{aligned} \tag{4.1.1}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g_t : \mathbb{R}^n \rightarrow \mathbb{R}$, $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $G_s : \mathbb{R}^n \rightarrow \mathbb{R}$ e $H_s : \mathbb{R}^n \rightarrow \mathbb{R}$. Assumimos que essas funções são pelo menos de classe C^2 . Os vetores $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_t(\mathbf{x}), \dots, g_m(\mathbf{x}))^T \in \mathbb{R}^m$, $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \dots, h_i(\mathbf{x}), \dots, h_r(\mathbf{x}))^T \in \mathbb{R}^r$, $\mathbf{G}(\mathbf{x}) = (G_1(\mathbf{x}), \dots, G_s(\mathbf{x}), \dots, G_p(\mathbf{x}))^T \in \mathbb{R}^p$ e $\mathbf{H}(\mathbf{x}) = (H_1(\mathbf{x}), \dots, H_s(\mathbf{x}), \dots, H_p(\mathbf{x}))^T \in \mathbb{R}^p$ serão utilizados sempre que oportuno.

O problema (4.1.1) é também denominado programação matemática com restrições de equilíbrio (MPEC). Geralmente, um MPEC é um problema de otimização cujas restrições essenciais são definidas por uma desigualdade variacional. Em um espaço com dimensão finita, a desigualdade variacional é definida como um problema de determinarmos $\bar{\mathbf{x}} \in \mathbb{X} \subseteq \mathbb{R}^n$ tal que para todo $\mathbf{x} \in \mathbb{X}$, temos: $\mathbf{F}(\bar{\mathbf{x}})^T (\mathbf{x} - \bar{\mathbf{x}}) \geq 0$; onde $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é uma aplicação linear (ou não) e contínua. Vide o livro [Luo, Pang e Ralph 1996] para uma extensa discussão sobre esse tema.

Sob certas circunstâncias, no entanto, os MPECs podem ser particularizados para o MPCC (4.1.1). Em geral, os MPCCs baseiam-se na ideia de que a condição de complementaridade pode ser usada para modelar uma mudança no comportamento de um sistema [Rosehart, Roman e Schellenberg 2005]. No contexto deste trabalho, as siglas MPEC e MPCC não nos trarão sentido ambíguo, portanto, elas serão utilizadas de maneira indistinta doravante para designar todo problema de programação matemática que contém restrições de complementaridade em sua formulação.

A notação $\mathbf{0} \leq \mathbf{G}(\mathbf{x}) \perp \mathbf{H}(\mathbf{x}) \geq \mathbf{0}$ é uma abreviação para o conjunto de restrições de igualdade e de desigualdade apresentado em (4.1.2):

$$\begin{aligned}
 & G_s(\mathbf{x}) H_s(\mathbf{x}) = 0 \\
 & G_s(\mathbf{x}) \geq 0, \quad \forall s = 1, \dots, p. \\
 & H_s(\mathbf{x}) \geq 0
 \end{aligned} \tag{4.1.2}$$

Este sistema de equações e inequações são denominados restrições de complementaridade devido à exigência de *ortogonalidade* entre os funcionais $\mathbf{G}$ e $\mathbf{H}$. A presença das restrições $G_s(\mathbf{x}) \geq 0$ e

$H_s(\mathbf{x}) \geq 0$ implica que (4.1.2) é equivalente às expressões definidas por:

$$\begin{aligned} \mathbf{G}(\mathbf{x})^T \mathbf{H}(\mathbf{x}) &= 0 \\ \mathbf{G}(\mathbf{x}) &\geq \mathbf{0} \\ \mathbf{H}(\mathbf{x}) &\geq \mathbf{0} \end{aligned} \quad (4.1.3)$$

e

$$\begin{aligned} G_s(\mathbf{x}) H_s(\mathbf{x}) &\leq 0 \\ G_s(\mathbf{x}) &\geq 0, \forall s = 1, \dots, p. \\ H_s(\mathbf{x}) &\geq 0 \end{aligned} \quad (4.1.4)$$

Cabe ressaltar que se $\mathbf{G} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ com $G_s(\mathbf{x}) = x_s$ e $\mathbf{H} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é um funcional não-linear adequado, então as restrições de complementaridade apresentadas em (4.1.5) são denominadas restrições de complementaridade não-linear [Chen e Mangasarian 1996, Ferris e Pang 1997, Chen e Harker 1997].

$$\begin{aligned} x_s H_s(\mathbf{x}) &= 0 \\ x_s &\geq 0, \forall s = 1, \dots, p. \\ H_s &\geq 0 \end{aligned} \quad (4.1.5)$$

Em particular, se $\mathbf{H}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$, com $M \in \mathbb{R}^{n \times n}$ e $\mathbf{b} \in \mathbb{R}^n$, então (4.1.5) é denominado restrição de complementaridade linear [Kojima et al. 1991, Chen e Harker 1993].

Se uma restrição de complementaridade $0 \leq G_s(\mathbf{x}) \perp H_s(\mathbf{x}) \leq 0$ é tal que os conjuntos imagem de $G_s(\mathbf{x})$ e de $H_s(\mathbf{x})$ são o $\mathbb{R}_+$, então a representação geométrica dessas restrições é conforme a Figura 4.1.1.

Figura 4.1.1: Representação geométrica de $0 \leq G_s(\mathbf{x}) \perp H_s(\mathbf{x}) \leq 0$.

Da própria definição de restrições de complementaridade, constatamos imediatamente que a região viável é não-convexa e o conjunto dos pontos interiores definido por estas restrições é vazio. Além disso, em termos de algoritmo e de desempenho computacional, as expressões (4.1.2) – (4.1.4) podem apresentar desempenho diferente para resolver o problema (4.1.1) [Benson, Shanno e Vanderbei 2002].

Definição 4.1.1 ([Ye 2005]). Seja $\mathbf{x}^* \in \mathbb{X}$ um minimizador para o MPEC (4.1.1). Definimos os seguintes conjuntos de restrições ativas:

$$\begin{cases} \mathbb{I}(\mathbf{x}^*) = \{i; h_i(\mathbf{x}^*) = 0\} \\ \mathbb{A}_1(\mathbf{x}^*) = \{s; G_s(\mathbf{x}^*) = 0, H_s(\mathbf{x}^*) > 0\} \\ \mathbb{A}_2(\mathbf{x}^*) = \{s; G_s(\mathbf{x}^*) = 0, H_s(\mathbf{x}^*) = 0\} \\ \mathbb{A}_3(\mathbf{x}^*) = \{s; G_s(\mathbf{x}^*) > 0, H_s(\mathbf{x}^*) = 0\}. \end{cases} \quad (4.1.6)$$

o conjunto $\mathbb{A}_2(\mathbf{x}^*)$ é denominado “conjunto degenerado” ou “conjunto das restrições bi-ativas”. Se $\mathbb{A}_2(\mathbf{x}^*) = \emptyset$, então o vetor $\mathbf{x}^*$ satisfaz as condições de complementaridade estrita. $\square$

Está claro que $\{\mathbb{A}_1(\mathbf{x}^*), \mathbb{A}_2(\mathbf{x}^*), \mathbb{A}_3(\mathbf{x}^*)\}$ é uma partição de $\{1, 2, \dots, p\}$. Apesar da simplicidade estrutural das restrições de complementaridade, a inserção dessas restrições nos problemas de programação matemática implicam que as condições de qualificação elementares de independência linear e de Mangasarian-Fromovitz não são satisfeitas em nenhum ponto viável [Chen e Florian 1995, Ye, Zhu e Zhu 1997]. Isso significa que as condições de KKT de primeira ordem não estão definidas através dessas condições de qualificação. Portanto, há um grande desafio tanto no aspecto teórico quanto no tratamento numérico do MPEC (4.1.1).

No aspecto teórico, temos o desenvolvimento de condições de qualificação, condições necessárias e suficientes de estacionaridades de primeira e de segunda ordem, análise de sensibilidade e análise de estabilidade [Ye 2005, Scheel e Scholtes 2000, Guo, Lin e Ye 2012, Guo et al. 2014]. Quanto ao aspecto do tratamento numérico, temos as seguintes metodologias principais: métodos de pontos interiores, programação quadrática sequencial, método de lagrangiano aumentado e os de região de confiança [Ferris e Pang 1997, Liu e Sun 2004, Jiang et al. 2014, Jiang e Ralph 2000, Yang e Qi 2005, Guo, Lin e Ye 2015].

Em [Flegel e Kanzow 2005], os autores estabeleceram a contextualização da condição de Guignard [Guignard 1969, Gould e Tolle 1971] no campo dos MPECs frente a outras condições de qualificação. Dentre essas outras condições, citamos a condição de qualificação MPEC Independência Linear (MPEC-LICQ) e a condição de MPEC-Abadie, as quais são generalizações das condições de qualificação de independência linear e de Abadie vinculadas ao problema de programação não-linear padrão.

A condição MPEC-LICQ é a premissa básica para o projeto de muitos algoritmos no campo dos MPECs, veja exemplos em [Hu e Ralph 2004, Liu e Sun 2004, Fletcher et al. 2006]. Tal condição, se satisfeita, implica que condição de Guignard também é, mas a recíproca não é verdadeira [Flegel e Kanzow 2005, teor. 4.5].

Apesar de que a condição de qualificação de Abadie ser muito fraca e suficiente para uma grande variedade de problemas de programação não-linear, ela não é suficientemente fraca no contexto de MPECs. Dentre as consequências deste fato, mesmo que uma certa implementação computacional esteja embasada na condição de qualificação de restrição de Abadie, ela não é suficientemente boa para resolver muitos MPECs.

Por exemplo, consideremos o sistema de complementaridade mais simples dado por:

$$\begin{aligned} xy &= 0 \\ x &\geq 0 \\ y &\geq 0. \end{aligned} \quad (4.1.7)$$

Em [Ribeiro e Karas 2013, p. 208], os autores mostram que a condição de qualificação de Abadie não é satisfeita no ponto de degenerescência $(x, y) = (0, 0)$. Em [Flegel e Kanzow 2005], os autores frisam que a condição de Guignard parece ser amplamente negligenciada na literatura no contexto de MPECs e, pelo fato de que a condição de Abadie falhar em muitos MPECs, houve, então, a necessidade da investigação da condição de Guignard no contexto de MPEC.

A saber, a condição de qualificação de Guignard é satisfeita para o sistema de complementaridade (4.1.7) para todo par (x, y) viável [Flegel 2005, p. 22].

4.1.2 Condições de estacionaridade: definições e formulações

Devido à variedade de MPECs existentes na literatura, surgem diversas condições de estacionaridade vinculadas às variáveis primal e dual do problema. A diversidade está diretamente relacionada aos conjuntos das restrições ativas $\mathbb{A}_1(\mathbf{x}^*)$, $\mathbb{A}_2(\mathbf{x}^*)$ e $\mathbb{A}_3(\mathbf{x}^*)$, os quais implicam diretamente no arranjo dos multiplicadores de Lagrange associados às restrições de complementaridade. Em outras palavras, o arranjo das variáveis duais do MPEC (4.1.1) implica em uma determinada condição de estacionaridade.

Frente à programação não-linear padrão, a qual possui apenas um conceito de estacionaridade para o problema dual, KKT, as condições de estacionaridade para o MPEC (4.1.1) são uma generalização das condições de KKT. Os conceitos de estacionaridade para MPECs são: Weak (W-), Clarke (C-), Alternativo (A-), Mordukhovich (M-) e Strong (S-) estacionaridade. O termo “A” não só significa “Alternativo”, pois descreve as propriedades dos multiplicadores de Lagrange, mas também pode significar “Abadie”, pois ocorreu primeiro quando a condição de qualificação de restrição MPEC-Abadie foi aplicada ao MPEC (4.1.1) [Flegel e Kanzow 2005, Flegel e Kanzow 2005].

Em particular, a S-estacionaridade é a mais forte destas condições, no sentido de que suas exigências são muito restritivas. Sua relevância teórica é ser equivalente às condições de KKT quando formulamos o MPEC (4.1.1) como um problema de programação não-linear padrão com o conjunto de equações e inequações (4.1.2) ou (4.1.3), [Flegel e Kanzow 2005, Guo e Lin 2012].

Sejam $\nabla \mathbf{g}(\mathbf{x}) \in \mathbb{R}^{m \times n}$, $\nabla \mathbf{h}(\mathbf{x}) \in \mathbb{R}^{r \times n}$, $\nabla \mathbf{G}(\mathbf{x}) \in \mathbb{R}^{p \times n}$ e $\nabla \mathbf{H}(\mathbf{x}) \in \mathbb{R}^{p \times n}$ respectivamente as matrizes jacobianas associadas aos funcionais $\mathbf{g}(\mathbf{x})$, $\mathbf{h}(\mathbf{x})$, $\mathbf{G}(\mathbf{x})$ e $\mathbf{H}(\mathbf{x})$ do MPEC (4.1.1) e os vetores $\lambda_0 = (\lambda_{0_1}, \dots, \lambda_{0_t}, \dots, \lambda_{0_m})^T \in \mathbb{R}^m$, $\lambda_1 = (\lambda_{1_1}, \dots, \lambda_{1_i}, \dots, \lambda_{1_r})^T \in \mathbb{R}_+^r$, $\pi_1 = (\pi_{1_1}, \dots, \pi_{1_s}, \dots, \pi_{1_p})^T \in \mathbb{R}^p$ e $\pi_2 = (\pi_{2_1}, \dots, \pi_{2_s}, \dots, \pi_{2_p})^T \in \mathbb{R}^p$. Então segue as Definições 4.1.2 - 4.1.6 [Ye 2005].

Definição 4.1.2 (W-estacionaridade). *Seja o MPEC (4.1.1). Um ponto $\mathbf{x}^* \in \mathbb{X}$ satisfaz a W-estacionaridade se existem multiplicadores de Lagrange $\lambda_0 \in \mathbb{R}^m$, $\lambda_1 \in \mathbb{R}_+^r$, $\pi_1 \in \mathbb{R}^p$ e $\pi_2 \in \mathbb{R}^p$ tais*

que:

$$\left\{ \begin{array}{l} \nabla f(\mathbf{x}^*) + \nabla \mathbf{g}(\mathbf{x}^*)^T \lambda_0 + \nabla \mathbf{h}(\mathbf{x}^*)^T \lambda_1 - \nabla \mathbf{G}(\mathbf{x}^*)^T \pi_1 - \nabla \mathbf{H}(\mathbf{x}^*)^T \pi_2 = \mathbf{0} \\ \text{diag}(\lambda_{1r})_{r=1}^r \mathbf{h}(\mathbf{x}^*) = \mathbf{0} \\ \lambda_1 \geq \mathbf{0} \\ \mathbf{h}(\mathbf{x}^*) \leq \mathbf{0} \\ \mathbf{g}(\mathbf{x}^*) = \mathbf{0} \\ \text{diag}(G_s(\mathbf{x}^*))_{s=1}^p \mathbf{H}(\mathbf{x}^*) = \mathbf{0} \\ \mathbf{G}(\mathbf{x}^*) \geq \mathbf{0} \\ \mathbf{H}(\mathbf{x}^*) \geq \mathbf{0} \end{array} \right. \quad (4.1.8)$$

e

$$\left\{ \begin{array}{l} \pi_{1s} = 0 \quad , \quad \forall s \in \mathbb{A}_3(\mathbf{x}^*) \\ \pi_{2s} = 0 \quad , \quad \forall s \in \mathbb{A}_1(\mathbf{x}^*) . \end{array} \right. \quad (4.1.9)$$

□

Definição 4.1.3 (C-estacionaridade). Um ponto $\mathbf{x}^* \in \mathbb{X}$ satisfaz a C-estacionaridade se existem multiplicadores de Lagrange $\lambda_0 \in \mathbb{R}^m$, $\lambda_1 \in \mathbb{R}_+^r$, $\pi_1 \in \mathbb{R}^p$ e $\pi_2 \in \mathbb{R}^p$ tais que (4.1.8) e (4.1.9) são satisfeitas e:

$$\pi_{1s} \pi_{2s} \geq 0 \quad , \quad \forall s \in \mathbb{A}_2(\mathbf{x}^*) . \quad (4.1.10)$$

□

Definição 4.1.4 (A-estacionaridade). Um ponto $\mathbf{x}^* \in \mathbb{X}$ satisfaz a A-estacionaridade se existem multiplicadores de Lagrange $\lambda_0 \in \mathbb{R}^m$, $\lambda_1 \in \mathbb{R}_+^r$, $\pi_1 \in \mathbb{R}^p$ e $\pi_2 \in \mathbb{R}^p$ tais que (4.1.8) e (4.1.9) são satisfeitas e:

$$\pi_{1s} \geq 0 \quad \text{ou} \quad \pi_{2s} \geq 0 \quad , \quad \forall s \in \mathbb{A}_2(\mathbf{x}^*) . \quad (4.1.11)$$

□

Definição 4.1.5 (M-estacionaridade). Um ponto $\mathbf{x}^* \in \mathbb{X}$ satisfaz a M-estacionaridade se existem multiplicadores de Lagrange $\lambda_0 \in \mathbb{R}^m$, $\lambda_1 \in \mathbb{R}_+^r$, $\pi_1 \in \mathbb{R}^p$ e $\pi_2 \in \mathbb{R}^p$ tais que (4.1.8) e (4.1.9) são satisfeitas e:

$$\text{ou} \quad \pi_{1s} \pi_{2s} = 0 \quad \text{ou} \quad \pi_{1s} > 0, \pi_{2s} > 0 \quad , \quad \forall s \in \mathbb{A}_2(\mathbf{x}^*) . \quad (4.1.12)$$

□

Definição 4.1.6 (S-estacionaridade). Um ponto $\mathbf{x}^* \in \mathbb{X}$ satisfaz a S-estacionaridade se existem multiplicadores de Lagrange $\lambda_0 \in \mathbb{R}^m$, $\lambda_1 \in \mathbb{R}_+^r$, $\pi_1 \in \mathbb{R}^p$ e $\pi_2 \in \mathbb{R}^p$ tais que (4.1.8) e (4.1.9) são satisfeitas e:

$$\pi_{1s} \geq 0, \quad \pi_{2s} \geq 0 \quad , \quad \forall s \in \mathbb{A}_2(\mathbf{x}^*) . \quad (4.1.13)$$

□

A relação condicional entre as condições de estacionaridade são dadas por:

Observação 4.1.1. O arranjo dos multiplicadores de Lagrange das restrições de complementaridade no conjunto bi-ativo é decisivo para sabermos a qual condição de estacionaridade um determinado MPEC está relacionado. $\square$

No sentido de álgebra abstrata, os conjuntos ativos são indispensáveis para uma análise geral nos MPECs. Os principais objetos matemáticos para a álgebra são as operações envolvendo cones por meio de relações entre conjuntos. As principais consequências estão relacionadas à determinação de condição de qualificação, como a de Abadie e MPEC-Abadie apresentadas em [Flegel e Kanzow 2005], que, por sua vez, implicam diretamente para as condições de estacionaridade primal-dual e para o planejamento de algoritmos computacionais. Vide, por exemplo, o método de pontos interiores para MPECs desenvolvido em [Liu e Sun 2004].

É evidente que as operações entre os cones podem nos dizer as condições para o qual um ponto primal seja estacionário. No caso dos MPECs e da programação não-linear padrão, caso um ponto primal seja estacionário, então tal ponto deve satisfazer a condição do cone das direções críticas definida em [Guo, Lin e Ye 2012].

Quanto ao aspecto prático, como poderemos resolver o problema (4.1.1) sem conhecermos antecipadamente que o minimizador $\mathbf{x}^* \in \mathbb{X}$ satisfaz a S-, M-, A- ou a C-estacionaridade? Além disso, não há um método que resolva as condições de estacionaridade diretamente. Um modo eficaz consiste em propor um conjunto de restrições simples e que sejam equivalentes a uma condição de estacionaridade ao incorporá-las no sistema (4.1.8). Está claro que a finalidade é eliminarmos o conjunto de índices associados ao conjunto das restrições ativas (4.1.6). Uma solução para essas questões é apresentada no Teorema 4.1.1 demonstrado em [Guo, Lin e Ye 2015].

Teorema 4.1.1. Para qualquer $\mathbf{x}^* \in \mathbb{X}$, temos:

1. As condições (4.1.9) e (4.1.10), C-estacionaridade, são equivalentes às restrições (4.1.14):

$$\begin{aligned} \text{diag}(\pi_{1_s})_{s=1}^p \mathbf{G}(\mathbf{x}^*) &= \mathbf{0} \\ \text{diag}(\pi_{2_s})_{s=1}^p \mathbf{H}(\mathbf{x}^*) &= \mathbf{0} \\ \text{diag}(\pi_{1_s})_{s=1}^p \pi_2 &\geq \mathbf{0}. \end{aligned} \quad (4.1.14)$$

2. As condições (4.1.9) e (4.1.12), M-estacionaridade, são equivalentes às restrições (4.1.15):

$$\begin{aligned} \text{diag}(\pi_{1_s})_{s=1}^p \mathbf{G}(\mathbf{x}^*) &= \mathbf{0} \\ \text{diag}(\pi_{2_s})_{s=1}^p \mathbf{H}(\mathbf{x}^*) &= \mathbf{0} \\ \text{diag}(\pi_{1_s})_{s=1}^p \pi_2 &\geq \mathbf{0} \\ \max\{\pi_{1_s}, \pi_{2_s}\} &\geq 0, \quad \forall s = 1, \dots, p. \end{aligned} \quad (4.1.15)$$

3. As condições (4.1.9) e (4.1.13), S-estacionaridade, são equivalentes às restrições (4.1.16):

$$\begin{aligned} \text{diag}(\bar{\lambda}_{1_s})_{s=1}^p \mathbf{G}(\mathbf{x}^*) &= \mathbf{0} \\ \text{diag}(\bar{\lambda}_{2_s})_{s=1}^p \mathbf{H}(\mathbf{x}^*) &= \mathbf{0} \\ \bar{\lambda}_1 &\geq \mathbf{0} \\ \bar{\lambda}_2 &\geq \mathbf{0} \\ \pi_1 &= \bar{\lambda}_1 - \text{diag}(\bar{\lambda}_{0_s})_{s=1}^p \mathbf{H}(\mathbf{x}^*) \\ \pi_2 &= \bar{\lambda}_2 - \text{diag}(\bar{\lambda}_{0_s})_{s=1}^p \mathbf{G}(\mathbf{x}^*), \end{aligned} \quad (4.1.16)$$

em que $\bar{\lambda}_0 = (\bar{\lambda}_{0_1}, \dots, \bar{\lambda}_{0_s}, \dots, \bar{\lambda}_{0_p}) \in \mathbb{R}^p$ é o vetor de multiplicadores de Lagrange associado à restrição $\text{diag}(G_s(\mathbf{x}))_{s=1}^p \mathbf{H}(\mathbf{x}) = \mathbf{0}$; $\bar{\lambda}_1 = (\bar{\lambda}_{1_1}, \dots, \bar{\lambda}_{1_s}, \dots, \bar{\lambda}_{1_p}) \in \mathbb{R}_+^p$ é o vetor de multiplicadores de Lagrange associado à restrição $-\mathbf{G}(\mathbf{x}) \leq \mathbf{0}$; $\bar{\lambda}_2 = (\bar{\lambda}_{2_1}, \dots, \bar{\lambda}_{2_s}, \dots, \bar{\lambda}_{2_p}) \in \mathbb{R}_+^p$ é o vetor de multiplicadores de Lagrange associado à restrição $-\mathbf{H}(\mathbf{x}) \leq \mathbf{0}$. —

Os autores em [Guo, Lin e Ye 2015] não nos apresentam um conjunto restrições equivalentes a A-estacionaridade. Propomos no Teorema 4.1.2 uma forma de gerar tais restrições.

Teorema 4.1.2. *As condições (4.1.9) e (4.1.11), A-estacionaridade, são equivalentes às equações (4.1.17):*

$$\begin{aligned} \text{diag}(\pi_{1_s})_{s=1}^p \mathbf{G}(\mathbf{x}^*) &= \mathbf{0} \\ \text{diag}(\pi_{2_s})_{s=1}^p \mathbf{H}(\mathbf{x}^*) &= \mathbf{0} \\ \max(\pi_{1_s}, \pi_{2_s}) &\geq 0, \forall s = 1, \dots, p. \end{aligned} \quad (4.1.17)$$

Demonstração. A condição (4.1.17) implica imediatamente em (4.1.9) e (4.1.11). Agora, vamos supor que (4.1.9) e (4.1.11) estão asseguradas.

A expressão (4.1.9) implica que $\text{diag}(\pi_{1_s})_{s=1}^p \mathbf{G}(\mathbf{x}^*) = \mathbf{0}$ e que $\text{diag}(\pi_{2_s})_{s=1}^p \mathbf{H}(\mathbf{x}^*) = \mathbf{0}$ (folgas complementares). Vamos mostrar que $\max(\pi_{1_s}, \pi_{2_s}) \geq 0$, para todo s . Se $s \notin \mathbb{A}_2(\mathbf{x}^*)$, então ou $s \in \mathbb{A}_1(\mathbf{x}^*)$ ou $s \in \mathbb{A}_3(\mathbf{x}^*)$. Assim, temos, respectivamente, que ou π_{1_s} é irrestrito e $\pi_{2_s} = 0$ ou π_{2_s} é irrestrito e $\pi_{1_s} = 0$. Agora suponha que $s \in \mathbb{A}_2(\mathbf{x}^*)$, então temos, por hipótese, que $\pi_{1_s} \geq 0$ ou $\pi_{2_s} \geq 0$. Portanto, para essas situações a desigualdade $\max(\pi_{1_s}, \pi_{2_s}) \geq 0$ é sempre satisfeita. Isto completa a demonstração da proposição. ■

Observação 4.1.2. Em [Flegel e Kanzow 2005], os autores chamam a atenção para o fato que a intersecção entre as A- e C- estacionaridades implica na M-estacionaridade. Esta constatação é observável ao compararmos as expressões (4.1.17) e (4.1.14) com (4.1.15). □

Observação 4.1.3. A formulação do MPEC (4.1.1) com as restrições de complementaridade apresentadas em (4.1.2), tem a função lagrangiana relativa ao problema expressa por:

$$L(\omega) = f(\mathbf{x}) + \lambda_0^T \mathbf{g}(\mathbf{x}) + \lambda_1^T \mathbf{h}(\mathbf{x}) + \bar{\lambda}_0^T \text{diag}(G_s(\mathbf{x}))_{s=1}^p \mathbf{h}(\mathbf{x}) - \bar{\lambda}_1^T \mathbf{G}(\mathbf{x}) - \bar{\lambda}_2^T \mathbf{H}(\mathbf{x}),$$

em que $\omega = (\mathbf{x}^T, \lambda_0^T, \lambda_1^T, \bar{\lambda}_0^T, \bar{\lambda}_1^T, \bar{\lambda}_2^T)^T \in \mathbb{R}^{n+m+r+3p}$, com $(\lambda_1, \bar{\lambda}_1, \bar{\lambda}_2) \in \mathbb{R}_+^{r+2p}$. Logo, as condições necessárias de otimalidade de KKT coincidem com os sistemas (4.1.8) e (4.1.16), ou seja, as condições de KKT de primeira ordem coincidem com a S-estacionaridade. □

4.2 Restrições de complementaridade mista

Introduzimos o conceito de problema de complementaridade mista através do seguinte exemplo proposto em [Chuquipoma 2010].

Exemplo 4.2.1. *Seja $f : \mathbb{X} = [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ uma função não-linear e diferenciável sobre $\mathbb{X}$. Se x^* é um ponto de mínimo, então sabemos que três casos podem ocorrer:*

$$\begin{cases} \text{Se } x^* = a & , \text{ então } f'(x^*) \geq 0 \\ \text{Se } a < x^* < b & , \text{ então } f'(x^*) = 0 \\ \text{Se } x^* = b & , \text{ então } f'(x^*) \leq 0 \end{cases} \quad (4.2.1)$$

As condições necessárias de otimalidade podem ser resumidas como o problema de desigualdade variacional dado por:

$$PDV(f', \mathbb{X}) : \text{determine } x^* \in \mathbb{X} \text{ tal que } f'(x^*)(x - x^*) \geq 0 \text{ para todo } x \in \mathbb{X}.$$

□

De acordo com a característica do conjunto $\mathbb{X} \subseteq \mathbb{R} \cup \{-\infty, \infty\}$, o $PDV(f', \mathbb{X})$ pode ser formulado equivalentemente por meio dos seguintes problemas:

Problema 1: Seja $\mathbb{X} = [a, \infty]$, então o $PDV(f', \mathbb{X})$ é equivalente ao problema de complementaridade dado por: encontre $x^* \in \mathbb{X}$ tal que $a \leq x^* \perp f'(x^*) \geq 0$.

Problema 2: Seja $\mathbb{X} = [-\infty, \infty]$, então o $PDV(f', \mathbb{X})$ é equivalente ao problema de encontrar $x^* \in \mathbb{X}$ que seja a solução da equação $f'(x) = 0$.

Problema 3: Seja $\mathbb{X} = [-\infty, b]$, então o $PDV(f', \mathbb{X})$ é equivalente ao problema de complementaridade dado por: encontre $x^* \in \mathbb{X}$ $x^* \leq b \perp f'(x^*) \leq 0$.

Problema 4: Se $\mathbb{X} = [a, b]$, então $PDV(f', \mathbb{X})$ é equivalente ao problema de complementaridade mista dado por: encontre $(x^*, c_1^*, c_2^*) \in (\mathbb{X} \times \mathbb{R}_+^2)$ tal que

$$\begin{aligned} f'(x^*) - c_1^* + c_2^* &= 0 \\ c_1^* &\geq 0 \\ c_2^* &\geq 0 \\ 0 &\leq (x^* - a) \perp c_1^* \geq 0 \\ 0 &\leq (-x^* + b) \perp c_2^* \geq 0. \end{aligned} \tag{4.2.2}$$

As três primeiras expressões em (4.2.2) nos asseguram que a procura da solução x^* para $f'(x)$ é feita de forma irrestrita, ao passo que a deliberação para $f'(x^*) \geq 0$, $f'(x^*) = 0$ ou para $f'(x^*) \leq 0$ é estritamente determinada em função das duas restrições de complementaridade em (4.2.2). Portanto, o problema 4 consiste em encontrar uma solução que resolva pelo menos um dos problemas de complementaridade anteriores.

De uma maneira mais geral, um MPEC cujas restrições essenciais são definidas por um sistema de complementaridade mista é de acordo com a seguinte formulação [Benson, Shanno e Vanderbei 2002]:

$$\begin{aligned} \text{Min } & f(\mathbf{x}) \\ \text{s.a : } & l_{1t} \leq T_t(\mathbf{x}) \leq l_{2t} \perp W_t(\mathbf{x}), \forall t = 1, \dots, q, \end{aligned} \tag{4.2.3}$$

onde: $T_t : \mathbb{R}^n \rightarrow \mathbb{R}$ e $W_t : \mathbb{R}^n \rightarrow \mathbb{R}$ são funções pelo menos de classe C^2 e $l_{1t} < l_{2t}$ são constantes que limitam $T_t(\mathbf{x})$ inferiormente e superiormente respectivamente. Utilizamos os vetores $\mathbf{T}(\mathbf{x}) = (T_1(\mathbf{x}), \dots, T_t(\mathbf{x}), \dots, T_q(\mathbf{x}))^T \in \mathbb{R}^q$, $\mathbf{W}(\mathbf{x}) = (W_1(\mathbf{x}), \dots, W_t(\mathbf{x}), \dots, W_q(\mathbf{x}))^T \in \mathbb{R}^q$, $\mathbf{l}_1 = (l_{11}, \dots, l_{1t}, \dots, l_{1q})^T \in \mathbb{R}^q$ e $\mathbf{l}_2 = (l_{21}, \dots, l_{2t}, \dots, l_{2q})^T \in \mathbb{R}^q$ sempre que for oportuno. Omitimos nesta formulação as restrições de igualdade e de desigualdade usuais por simplicidade. Novamente, as restrições de complementaridade mista em (4.2.3) significam:

$$\left\{ \begin{array}{ll} \text{Se } T_t(\mathbf{x}) = l_{1t} & , \text{ então } W_t(\mathbf{x}) \geq 0 \quad \forall t = 1, \dots, q \\ \text{Se } l_{1t} < T_t(\mathbf{x}) < l_{2t} & , \text{ então } W_t(\mathbf{x}) = 0 \quad \forall t = 1, \dots, q \\ \text{Se } T_t(\mathbf{x}) = l_{2t} & , \text{ então } W_t(\mathbf{x}) \leq 0 \quad \forall t = 1, \dots, q \end{array} \right. \tag{4.2.4}$$

Seja $0,5 \leq T_t(\mathbf{x}) \leq 2 \perp W_t(\mathbf{x})$ um exemplo de restrição de complementaridade mista em que o conjunto imagem de $W_t(\mathbf{x})$ é o conjunto $\mathbb{R}$, então essas restrições são representadas geometricamente de acordo com a Figura 4.2.1.

Figura 4.2.1: Representação geométrica de uma restrição de complementaridade mista.

Na literatura são encontradas pelo menos duas maneiras de *transcrevermos* uma restrição de complementaridade mista para um conjunto de restrições convencionais. A primeira forma consiste em reescrever essas restrições como um par de restrições usuais de complementaridade e acrescentar um conjunto de restrições de igualdade. De acordo com [Dirkse e Ferris 1995, Chen e Mangasarian 1996], segue a formulação (4.2.5).

$$\begin{aligned} \mathbf{W}(\mathbf{x}) - \mathbf{c}_1 + \mathbf{c}_2 &= \mathbf{0} \\ \mathbf{0} \leq \mathbf{c}_1 \perp (\mathbf{T}(\mathbf{x}) - \mathbf{l}_1) &\geq \mathbf{0} \\ \mathbf{0} \leq \mathbf{c}_2 \perp (-\mathbf{T}(\mathbf{x}) + \mathbf{l}_2) &\geq \mathbf{0}, \end{aligned} \quad (4.2.5)$$

onde: $\mathbf{c}_1 = (c_{1_1}, \dots, c_{1_t}, \dots, c_{1_q})^T \in \mathbb{C}_1 \subseteq \mathbb{R}_+^q$ e $\mathbf{c}_2 = (c_{2_1}, \dots, c_{2_t}, \dots, c_{2_q})^T \in \mathbb{C}_2 \subseteq \mathbb{R}_+^q$ são variáveis artificiais positivas. Apresentamos em (4.2.6) outra maneira de expressarmos as restrições de complementaridade mista. A vantagem desta formulação é que as variáveis artificiais $\mathbf{c}_1$ e $\mathbf{c}_2$ não são necessárias e não há restrições extras de igualdade. A formulação (4.2.6) é conforme [Benson, Shanno e Vanderbei 2002, Rosehart, Roman e Schellenberg 2005].

$$\begin{aligned} l_{1_t} &\leq T_t(\mathbf{x}) \leq l_{2_t} \\ (T_t(\mathbf{x}) - l_{1_t}) W_t(\mathbf{x}) &\leq 0, \forall t = 1, \dots, q. \\ (T_t(\mathbf{x}) - l_{2_t}) W_t(\mathbf{x}) &\leq 0 \end{aligned} \quad (4.2.6)$$

Definição 4.2.1. Seja $\mathbf{x}^* \in \mathbb{X}^*$ um minimizador para o MPEC (4.2.3). Definimos os seguintes conjuntos de restrições ativas:

$$\begin{cases} \mathbb{B}_1(\mathbf{x}^*) = \{t; T_t(\mathbf{x}^*) - l_{1_t} = 0, & W_t(\mathbf{x}^*) > 0\} \\ \mathbb{B}_2(\mathbf{x}^*) = \{t; T_t(\mathbf{x}^*) - l_{1_t} = 0, & W_t(\mathbf{x}^*) = 0\} \\ \mathbb{B}_3(\mathbf{x}^*) = \{t; l_{1_t} < T_t(\mathbf{x}^*) < l_{2_t}, & W_t(\mathbf{x}^*) = 0\} \\ \mathbb{B}_4(\mathbf{x}^*) = \{t; T_t(\mathbf{x}^*) - l_{2_t} = 0, & W_t(\mathbf{x}^*) = 0\} \\ \mathbb{B}_5(\mathbf{x}^*) = \{t; T_t(\mathbf{x}^*) - l_{2_t} = 0, & W_t(\mathbf{x}^*) < 0\}. \end{cases} \quad (4.2.7)$$

Os conjuntos $\mathbb{B}_2(\mathbf{x}^*)$ e $\mathbb{B}_4(\mathbf{x}^*)$ são denominados “conjuntos degenerados” ou “conjuntos das restrições bi-ativas”. Se $\mathbb{B}_2(\mathbf{x}^*) = \mathbb{B}_4(\mathbf{x}^*) = \emptyset$, então o vetor $\mathbf{x}^*$ satisfaz as condições de complementaridade estrita. $\square$

Ao considerarmos a formulação (4.2.6) para representar as restrições de complementaridade mista, o problema (4.2.3) é equivalente ao problema de programação não-linear dado por:

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) \\ \text{s.a :} \quad & -\mathbf{T}(\mathbf{x}) + \mathbf{l}_1 \leq \mathbf{0} \\ & \mathbf{T}(\mathbf{x}) - \mathbf{l}_2 \leq \mathbf{0} \\ & \text{diag}(T_t(\mathbf{x}) - l_{1_t})_{t=1}^q \mathbf{W}(\mathbf{x}) \leq \mathbf{0} \\ & \text{diag}(T_t(\mathbf{x}) - l_{2_t})_{t=1}^q \mathbf{W}(\mathbf{x}) \leq \mathbf{0}. \end{aligned} \quad (4.2.8)$$

A função lagrangiana associada ao problema (4.2.8) é conforme (4.2.9) a seguir:

$$\begin{aligned} L(\mathbf{x}, \bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3, \bar{\lambda}_4) = & f(\mathbf{x}) + \bar{\lambda}_1^T (-\mathbf{T}(\mathbf{x}) + \mathbf{l}_1) + \bar{\lambda}_2^T (\mathbf{T}(\mathbf{x}) - \mathbf{l}_2) + \\ & + \bar{\lambda}_3^T (\text{diag}(T_t(\mathbf{x}) - l_{1_t})_{t=1}^q \mathbf{W}(\mathbf{x})) + \bar{\lambda}_4^T (\text{diag}(T_t(\mathbf{x}) - l_{2_t})_{t=1}^q \mathbf{W}(\mathbf{x})), \end{aligned} \quad (4.2.9)$$

em que: $\bar{\lambda}_1 = (\bar{\lambda}_{1_1}, \dots, \bar{\lambda}_{1_t}, \dots, \bar{\lambda}_{1_q})^T \in \mathbb{R}_+^q$ é vetor dual associado à restrição $-\mathbf{T}(\mathbf{x}) + \mathbf{l}_1 \leq \mathbf{0}$; $\bar{\lambda}_2 = (\bar{\lambda}_{2_1}, \dots, \bar{\lambda}_{2_t}, \dots, \bar{\lambda}_{2_q})^T \in \mathbb{R}_+^q$ é vetor dual associado à restrição $\mathbf{T}(\mathbf{x}) - \mathbf{l}_2 \leq \mathbf{0}$; $\bar{\lambda}_3 = (\bar{\lambda}_{3_1}, \dots, \bar{\lambda}_{3_t}, \dots, \bar{\lambda}_{3_q})^T \in \mathbb{R}_+^q$ é vetor dual associado à restrição $\text{diag}(T_t(\mathbf{x}) - l_{1_t})_{t=1}^q \mathbf{W}(\mathbf{x}) \leq \mathbf{0}$ e; $\bar{\lambda}_4 = (\bar{\lambda}_{4_1}, \dots, \bar{\lambda}_{4_t}, \dots, \bar{\lambda}_{4_q})^T \in \mathbb{R}_+^q$ é vetor dual associado à restrição $\text{diag}(T_t(\mathbf{x}) - l_{2_t})_{t=1}^q \mathbf{W}(\mathbf{x}) \leq \mathbf{0}$.

As condições de KKT de primeira ordem vinculados ao problema (4.2.8) são apresentadas no sistema não-linear de equações (4.2.10).

$$\begin{cases} \nabla f(\mathbf{x}) + \nabla \mathbf{T}(\mathbf{x})^T \pi_1 + \nabla \mathbf{W}(\mathbf{x})^T \pi_2 = \mathbf{0} \\ \pi_1 = -\bar{\lambda}_1 + \bar{\lambda}_2 + \text{diag}(W_t(\mathbf{x}))_{t=1}^q (\bar{\lambda}_3 + \bar{\lambda}_4) \\ \pi_2 = \text{diag}(T_t(\mathbf{x}) - l_{1_t})_{t=1}^q \bar{\lambda}_3 + \text{diag}(T_t(\mathbf{x}) - l_{2_t})_{t=1}^q \bar{\lambda}_4 \\ \min(\bar{\lambda}_{1_t}, T_t(\mathbf{x}) - l_{1_t}) = 0 \\ \min(\bar{\lambda}_{2_t}, -T_t(\mathbf{x}) + l_{2_t}) = 0 \\ \min(\bar{\lambda}_{3_t}, (-T_t(\mathbf{x}) + l_{1_t}) W_t(\mathbf{x})) = 0 \\ \min(\bar{\lambda}_{4_t}, (-T_t(\mathbf{x}) - l_{2_t}) W_t(\mathbf{x})) = 0, \quad \forall t = 1, \dots, q. \end{cases} \quad (4.2.10)$$

Verifiquemos que o conjunto de equações (4.2.10) condiz com a condição de S-estacionaridade, Definição 4.1.6, para o MPEC (4.2.3).

Seja $\mathbf{x}^* \in \mathbb{X}$ um minimizador para o problema o (4.2.8) e os vetores $\pi_1 = (\pi_{1_1}, \dots, \pi_{1_t}, \dots, \pi_{1_q})^T \in \mathbb{R}^q$ e $\pi_2 = (\pi_{2_1}, \dots, \pi_{2_t}, \dots, \pi_{2_q})^T \in \mathbb{R}^q$, os quais o par (π_{1_t}, π_{2_t}) é o multiplicador de Lagrange associado à restrição $l_{1_t} \leq T_t(\mathbf{x}) \leq l_{2_t} \perp W_t(\mathbf{x})$. É imediato constatar que:

- se $t \in \mathbb{B}_1$, implica que: $\bar{\lambda}_{1_t} \geq 0$, $\bar{\lambda}_{2_t} = 0$, $\bar{\lambda}_{3_t} \geq 0$ e $\bar{\lambda}_{4_t} = 0$. Portanto, de acordo com o segundo e terceiro conjunto de equações (4.2.10), temos: π_{1_s} é irrestrito e $\pi_{2_s} = 0$;
- se $t \in \mathbb{B}_2$ (conjunto bi-ativo), temos: $\bar{\lambda}_{1_t} \geq 0$, $\bar{\lambda}_{2_t} = 0$, $\bar{\lambda}_{3_t} \geq 0$ e $\bar{\lambda}_{4_t} \geq 0$. Para essas componentes duais do problema (4.2.8), então temos $\pi_{1_s} \leq 0$ e $\pi_{2_s} \leq 0$;
- se $t \in \mathbb{B}_3$, temos: $\bar{\lambda}_{1_t} = 0$, $\bar{\lambda}_{2_t} = 0$, $\bar{\lambda}_{3_t} \geq 0$ e $\bar{\lambda}_{4_t} \geq 0$. Portanto, segue que $\pi_{1_s} = 0$ e π_{2_s} é irrestrito;
- se $t \in \mathbb{B}_4$ (conjunto bi-ativo), temos: $\bar{\lambda}_{1_t} = 0$, $\bar{\lambda}_{2_t} \geq 0$, $\bar{\lambda}_{3_t} \geq 0$ e $\bar{\lambda}_{4_t} \geq 0$. Para essas componentes duais, obtemos: $\pi_{1_s} \geq 0$ e $\pi_{2_s} \geq 0$ e;
- se $t \in \mathbb{B}_5$, temos: $\bar{\lambda}_{1_t} = 0$, $\bar{\lambda}_{2_t} \geq 0$, $\bar{\lambda}_{3_t} = 0$ e $\bar{\lambda}_{4_t} \geq 0$. Portanto, temos: π_{1_s} é irrestrito e $\pi_{2_s} = 0$.

Para as situações em que ou $t \in \mathbb{B}_1$ ou $t \in \mathbb{B}_5$, a função $W_t(\mathbf{x}^*)$ torna-se inativa ao passo que $T_t(\mathbf{x}^*)$ torna-se ativa. Então, de acordo com a condição de folgas complementares do MPCC (expressão (4.1.9)), o multiplicador de Lagrange associado a $W_t(\mathbf{x}^*)$, π_{2_t} , deve ser nulo. Semelhante, para o caso de $t \in \mathbb{B}_3$, o qual temos que $W_t(\mathbf{x}^*)$ é ativa e que $T_t(\mathbf{x}^*)$ é inativa. Portanto, o multiplicador Lagrange associado a $T_t(\mathbf{x}^*)$, π_{1_t} , deve ser nulo.

Por fim, consideremos as situações em que o conjunto bi-ativo $\mathbb{B}_2$ ou $\mathbb{B}_4$ é não-vazio, isto é, ou $t \in \mathbb{B}_2$ ou $t \in \mathbb{B}_4$. A condição de S-estacionaridade, Definição 4.1.6, é satisfeita se ocorre ($\pi_{1_t} \leq 0$ e $\pi_{2_t} \leq 0$), se $t \in \mathbb{B}_2$ e; ($\pi_{1_t} \geq 0$ e $\pi_{2_t} \geq 0$), se $t \in \mathbb{B}_4$.

De fato, o sistema de KKT de primeira ordem (4.2.10) para o problema de programação não-linear (4.2.8) condiz com a condição de S-estacionaridade para o MPCC (4.2.3).

Por fim, consideremos o MPEC (4.2.11).

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) \\ \text{s.a :} \quad & \mathbf{l}_1 \leq \mathbf{T}(\mathbf{x}) \leq \mathbf{l}_2 \perp -\mathbf{W}(\mathbf{x}). \end{aligned} \quad (4.2.11)$$

Definição 4.2.2. Seja $\mathbf{x}^* \in \mathbb{X}^*$ um minimizador para o MPEC (4.2.11). Definimos os seguintes conjuntos de restrições ativas:

$$\left\{ \begin{array}{ll} \mathbb{B}_1(\mathbf{x}^*) = \{t; T_t(\mathbf{x}^*) - l_{1_t} = 0, & W_t(\mathbf{x}^*) < 0\} \\ \mathbb{B}_2(\mathbf{x}^*) = \{t; T_t(\mathbf{x}^*) - l_{1_t} = 0, & W_t(\mathbf{x}^*) = 0\} \\ \mathbb{B}_3(\mathbf{x}^*) = \{t; l_{1_t} < T_t(\mathbf{x}^*) < l_{2_t} & W_t(\mathbf{x}^*) = 0\} \\ \mathbb{B}_4(\mathbf{x}^*) = \{t; T_t(\mathbf{x}^*) - l_{2_t} = 0, & W_t(\mathbf{x}^*) = 0\} \\ \mathbb{B}_5(\mathbf{x}^*) = \{t; T_t(\mathbf{x}^*) - l_{2_t} = 0, & W_t(\mathbf{x}^*) > 0\}. \end{array} \right. \quad (4.2.12)$$

Os conjuntos $\mathbb{B}_2(\mathbf{x}^*)$ e $\mathbb{B}_4(\mathbf{x}^*)$ são denominados “conjuntos degenerados” ou “conjuntos das restrições bi-ativas”. Se $\mathbb{B}_2(\mathbf{x}^*) = \mathbb{B}_4(\mathbf{x}^*) = \emptyset$, então o vetor $\mathbf{x}^*$ satisfaz as condições de complementaridade estrita. $\square$

O problema (4.2.11) é equivalente ao problema de programação não-linear (4.2.13) dado por:

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) \\ \text{s.a :} \quad & -\mathbf{T}(\mathbf{x}) + \mathbf{l}_1 \leq \mathbf{0} \\ & \mathbf{T}(\mathbf{x}) - \mathbf{l}_2 \leq \mathbf{0} \\ & \text{diag}(T_t(\mathbf{x}) - l_{1_t})_{t=1}^q (-\mathbf{W}(\mathbf{x})) \leq \mathbf{0} \\ & \text{diag}(T_t(\mathbf{x}) - l_{2_t})_{t=1}^q (-\mathbf{W}(\mathbf{x})) \leq \mathbf{0}. \end{aligned} \quad (4.2.13)$$

Com procedimentos análogos à formulação do MPEC (4.2.3), temos que a condição de KKT de primeira ordem vinculada ao problema (4.2.13) é apresentada no sistema não-linear de equações (4.2.14).

$$\begin{cases} \nabla f(\mathbf{x}) + \nabla \mathbf{T}(\mathbf{x})^T \pi_1 + \nabla \mathbf{W}(\mathbf{x})^T \pi_2 = \mathbf{0} \\ \pi_1 = -\bar{\lambda}_1 + \bar{\lambda}_2 - \text{diag}(W_t(\mathbf{x}))_{t=1}^q (\bar{\lambda}_3 + \bar{\lambda}_4) \\ \pi_2 = -\text{diag}(T_t(\mathbf{x}) - l_{1_t})_{t=1}^q \bar{\lambda}_3 - \text{diag}(T_t(\mathbf{x}) - l_{2_t})_{t=1}^q \bar{\lambda}_4 \\ \min(\bar{\lambda}_{1_t}, T_t(\mathbf{x}) - l_{1_t}) = 0 \\ \min(\bar{\lambda}_{2_t}, -T_t(\mathbf{x}) + l_{2_t}) = 0 \\ \min(\bar{\lambda}_{3_t}, (-T_t(\mathbf{x}) + l_{1_t})(-W_t(\mathbf{x}))) = 0 \\ \min(\bar{\lambda}_{4_t}, (-T_t(\mathbf{x}) - l_{2_t})(-W_t(\mathbf{x}))) = 0, \quad \forall t = 1, \dots, q; \end{cases} \quad (4.2.14)$$

Seja $\mathbf{x}^* \in \mathbb{X}$ um minimizador para o problema o (4.2.11). Então a configuração dos sinais dos multiplicadores de Lagrange π_1 e π_2 é dada segundo as seguintes condições:

- se $t \in \mathbb{B}_1$, implica que: $\bar{\lambda}_{1_t} \geq 0$, $\bar{\lambda}_{2_t} = 0$, $\bar{\lambda}_{3_t} \geq 0$ e $\bar{\lambda}_{4_t} = 0$. Portanto, de acordo com o segundo e terceiro conjunto de equações (4.2.14), temos: π_{1_s} é irrestrito e $\pi_{2_s} = 0$;
- se $t \in \mathbb{B}_2$ (conjunto bi-ativo), temos: $\bar{\lambda}_{1_t} \geq 0$, $\bar{\lambda}_{2_t} = 0$, $\bar{\lambda}_{3_t} \geq 0$ e $\bar{\lambda}_{4_t} \geq 0$. Para essas componentes duais do problema (4.2.13), então temos $\pi_{1_s} \leq 0$ e $\pi_{2_s} \geq 0$;
- se $t \in \mathbb{B}_3$, temos: $\bar{\lambda}_{1_t} = 0$, $\bar{\lambda}_{2_t} = 0$, $\bar{\lambda}_{3_t} \geq 0$ e $\bar{\lambda}_{4_t} \geq 0$. Portanto, segue que $\pi_{1_s} = 0$ e π_{2_s} é irrestrito;
- se $t \in \mathbb{B}_4$ (conjunto bi-ativo), temos: $\bar{\lambda}_{1_t} = 0$, $\bar{\lambda}_{2_t} \geq 0$, $\bar{\lambda}_{3_t} \geq 0$ e $\bar{\lambda}_{4_t} \geq 0$. Para essas componentes duais, obtemos: $\pi_{1_s} \geq 0$ e $\pi_{2_s} \leq 0$ e;
- se $t \in \mathbb{B}_5$, temos: $\bar{\lambda}_{1_t} = 0$, $\bar{\lambda}_{2_t} \geq 0$, $\bar{\lambda}_{3_t} = 0$ e $\bar{\lambda}_{4_t} \geq 0$. Portanto, temos: π_{1_s} é irrestrito e $\pi_{2_s} = 0$.

O método primal-dual de reescalamiento não-linear integrado para resolver problemas de programação discreta-mista com restrições de equilíbrio

Nos capítulos anteriores, apresentamos: a proposta do método de reescalamiento não-linear integrado, o qual foi desenvolvido para programação convexa; as diversas estratégias, via método de penalidades, para resolvermos um problema de programação não-linear equivalente a um problema de programação discreta-mista e; uma subclasse do problema de programação matemática com restrições de equilíbrio, a qual o conceito de restrições de complementaridade e de complementaridade mista e as possíveis condições de estacionaridade primal-dual foram discutidas.

Embasados nos temas apresentados nos capítulos anteriores, nós propomos neste capítulo um método primal-dual do método de reescalamiento não-linear integrado (PDRNLI) para resolver uma classe de problemas de programação matemática discreta-mista com restrições de equilíbrio, cuja formulação é apresentada em (5.1.1) adiante. A importância prática deste problema no contexto da engenharia elétrica é que sua formulação adéqua-se ao problema de fluxo de potência ótimo reativo com restrições de atuação de dispositivos de controle de tensão proposto em [Lage 2013]. Portanto, o método PDRNLI é uma nova abordagem metodológica para resolver problemas de fluxo de potência ótimo. Este capítulo está organizado da seguinte forma:

- Na Seção 5.1, nós apresentamos a formulação do problema de programação discreta-mista com restrições de equilíbrio. Além disso, apresentamos uma interpretação geométrica das restrições de complementaridade associadas ao problema e verificamos que o conjunto viável é desconexo;
- Por meio das estratégias desenvolvidas no Capítulo 3, apresentamos na Seção 5.2 uma maneira de transformarmos um problema de programação discreta-mista com restrições de equilíbrio

em um MPEC formulado apenas com variáveis contínuas;

- Na Seção 5.3, as restrições de complementaridade do MPEC são reformuladas equivalentemente por meio de restrições de desigualdade usuais. Conforme apresentado no Capítulo 4, isso significa que transformaremos o MPEC em um problema de programação não-linear e não-convexo. A formulação do problema de programação não-linear é o nosso principal objeto de trabalho para o desenvolvimento da abordagem primal-dual de reescalamiento não-linear;
- Na Seção 5.4, apresentamos a formulação do sistema não-linear do método PDRNLI, o qual representa a condição de estacionaridade primal-dual. De acordo com algumas propostas encontradas na literatura, nós apresentamos como esse sistema é abordado. Por fim, nós apresentamos uma nova estratégia de como resolvê-lo;
- Na Seção 5.5, é iniciado o desenvolvimento do método previsor-corretor primal-dual de reescalamiento não-linear integrado. O objetivo desta seção é o desenvolvimento das equações exatas que definem as direções de busca primal e dual;
- Com o vetor de direções de busca primal-dual calculados, então apresentamos na Seção 5.6 o procedimento previsor-corretor do método PDRNLI;
- Na Seção 5.7, mostramos que o método primal-dual de reescalamiento não-linear integrado é um método de gradiente projetado em subespaços do $\mathbb{R}^n$;
- Na Seção 5.8, propomos uma estratégia de convergência global denominada condição de positividade, a qual é embasada na lei de inércia de Sylvester e no estudo de um produto interno que define um pseudo-espaço euclidiano, os quais são apresentados no Apêndice B;
- Na Seção 5.9, apresentamos as estratégias para o cálculo do tamanho do passo primal e dual e para a atualização dos parâmetros;
- Na Seção 5.10, apresentamos uma forma de inicialização do método PDRNLI;
- Na Seção 5.11, nós descrevemos um resumo do método PDRNLI e;
- Na Seção 5.12, nós apresentamos alguns resultados da implementação computacional do método PDRNLI em problemas matemáticos de programação não-linear, MPEC e programação discreta.

5.1 O problema de programação discreta-mista com restrições de equilíbrio

A formulação do problema de programação discreta-mista com restrições de equilíbrio é definido por:

$$\begin{aligned}
 \text{Min} \quad & \bar{f}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \\
 \text{s.a :} \quad & \mathbf{g}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \mathbf{0} \\
 & \mathbf{u}_1 \leq \mathbf{h}_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \leq \mathbf{u}_2 \\
 & \mathbf{u}_3 \leq \mathbf{x}_1 \leq \mathbf{u}_4 \\
 & \mathbf{u}_5 \leq \mathbf{x}_2 \leq \mathbf{u}_6 \perp (\mathbf{x}_3 - \mathbf{u}_7);
 \end{aligned} \tag{5.1.1}$$

onde: $\mathbf{x}_1 \in \mathbb{R}^{n_1}$; $\mathbf{x}_2 \in \mathbb{R}^{n_2}$; $\mathbf{x}_3 \in \mathbb{D} \subset \mathbb{R}^{n_2}$, $\mathbb{D}$ é o conjunto discreto apresentado em (3.1.1), p. 56; $\bar{f} : \mathbb{R}^{n_1+n_2} \times \mathbb{D} \rightarrow \mathbb{R}$, $\mathbf{h}_1 : \mathbb{R}^{n_1+n_2} \times \mathbb{D} \rightarrow \mathbb{R}^{r_1}$ e $\mathbf{g} : \mathbb{R}^{n_1+n_2} \times \mathbb{D} \rightarrow \mathbb{R}^m$, com $m < n_1 + 2n_2$. Além disso, temos que $\mathbf{g} = (g_1, \dots, g_t, \dots, g_m)^T \in \mathbb{R}^m$; $\mathbf{h}_1 = (h_{1_1}, \dots, h_{1_i}, \dots, h_{1_{r_1}})^T \in \mathbb{R}^{r_1}$; $\mathbf{u}_1 = (u_{1_1}, \dots, u_{1_i}, \dots, u_{1_{r_1}})^T \in \mathbb{R}^{r_1}$, $\mathbf{u}_2 = (u_{2_1}, \dots, u_{2_i}, \dots, u_{2_{r_1}})^T \in \mathbb{R}^{r_1}$, com $u_{1_i} < u_{2_i}$; $\mathbf{u}_3 = (u_{3_1}, \dots, u_{3_j}, \dots, u_{3_{n_1}})^T \in \mathbb{R}^{n_1}$, $\mathbf{u}_4 = (u_{4_1}, \dots, u_{4_j}, \dots, u_{4_{n_1}})^T \in \mathbb{R}^{n_1}$, com $u_{3_j} < u_{4_j}$; $\mathbf{u}_5 = (u_{5_1}, \dots, u_{5_j}, \dots, u_{5_{n_2}})^T \in \mathbb{R}^{n_2}$, $\mathbf{u}_6 = (u_{6_1}, \dots, u_{6_j}, \dots, u_{6_{n_2}})^T \in \mathbb{R}^{n_2}$, com $u_{5_j} < u_{6_j}$ e; $\mathbf{u}_7 = (u_{7_1}, \dots, u_{7_j}, \dots, u_{7_{n_3}})^T \in \mathbb{R}^{n_2}$.

Devido a natureza das variáveis $\mathbf{x}_2$ e $\mathbf{x}_3$, as quais são respectivamente contínua e discreta, temos que a restrição de complementaridade mista que relaciona essas variáveis tem uma interessante interpretação geométrica. Suponha que uma restrição de complementaridade mista é dada por $1 \leq x_2 \leq 3 \perp x_3$, em que $x_3 \in \{-3, -2, -1, 1, 2, 3\}$; então a Figura 5.1.1 mostra que o conjunto viável gerado por esta restrição é desconexo. Nós desconhecemos uma literatura que aborda sistematicamente esta categoria de MPEC no sentido teórico e metodológico. Mas, como é apresentado nos resultados numéricos obtidos em [Lage 2013], entendemos que uma miscelânea de métodos clássicos, os quais são surpreendentemente do tipo *gradiente*, é uma opção viável para lidarmos com esta categoria de problemas.

Figura 5.1.1: Exemplo de uma restrição de complementaridade mista definida pela relação entre uma variável contínua e discreta.

5.2 O MPEC discreto-misto como um MPEC formulado em variáveis contínuas

Esta seção consiste nos procedimentos para transformarmos o problema (5.1.1) em um outro MPEC formulado em variáveis estritamente contínuas. Destacamos que na literatura não encontramos metodologias para resolvermos diretamente um MPEC discreto-misto, portanto, esta transformação é necessária. Aqui, as estratégias apresentadas no Capítulo 3 para transformarmos um problema de programação discreta-mista em um problema contínuo equivalente são o cerne desta abordagem. Sejam os vetores $\mathbf{u}_8 = (u_{8_1}, \dots, u_{8_j}, \dots, u_{8_{n_2}})^T \in \mathbb{R}^{n_2}$ e $\mathbf{u}_9 = (u_{9_1}, \dots, u_{9_j}, \dots, u_{9_{n_2}})^T \in \mathbb{R}^{n_2}$, com $u_{8_j} < u_{9_j}$. As componentes u_{8_j} e u_{9_j} significam respectivamente o menor e o maior valor discreto que a variável x_{3_j} pode assumir. De acordo com a estratégia abordada em (3.2.1), p. 57, então a formulação (5.2.1) define um MPEC formulado em variáveis contínuas e equivalente ao problema (5.1.1).

$$\begin{aligned}
 \text{Min} \quad & \bar{f}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \\
 \text{s.a :} \quad & \mathbf{g}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \mathbf{0} \\
 & \mathbf{u}_1 \leq \mathbf{h}_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \leq \mathbf{u}_2 \\
 & \mathbf{u}_3 \leq \mathbf{x}_1 \leq \mathbf{u}_4 \\
 & \mathbf{u}_5 \leq \mathbf{x}_2 \leq \mathbf{u}_6 \perp (\mathbf{x}_3 - \mathbf{u}_7) \\
 & \mathbf{u}_8 \leq \mathbf{x}_3 \leq \mathbf{u}_9 \\
 & \mathbf{m}(\mathbf{x}_3) = \mathbf{0},
 \end{aligned} \tag{5.2.1}$$

onde $\mathbf{m}(\mathbf{x}_3) = (m_1(x_{3_1}), \dots, m_j(x_{3_j}), \dots, m_{n_2}(x_{3_{n_2}}))^T \in \mathbb{R}^{n_2}$ e $\mathbf{x}_3 \in \mathbb{R}^{n_2}$.

Seja $1 \leq x_2 \leq 3 \perp x_3$, em que $x_3 \in \{-3, -2, -1, 1, 2, 3\}$. A Figura 5.2.1 mostra geometricamente como esta restrição de complementaridade discreta-mista pode ser gerada através do conjunto de restrições definido por $\mathbb{U} = \{(x_2, x_3) \in \mathbb{R}^2; 1 \leq x_2 \leq 3 \perp x_3, -3 \leq x_3 \leq 3, m(x_3) = 0\}$.

Figura 5.2.1: Representação geométrica do conjunto de restrições $\mathbb{U}$.

O problema (5.1.1) é equivalente ao MPEC a seguir:

$$\begin{aligned}
 \text{Min} \quad & \bar{f}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + \sum_{j=1}^{n_2} \eta(m_j(x_{3_j})) \\
 \text{s.a :} \quad & \mathbf{g}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \mathbf{0} \\
 & \mathbf{u}_1 \leq \mathbf{h}_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \leq \mathbf{u}_2 \\
 & \mathbf{u}_3 \leq \mathbf{x}_1 \leq \mathbf{u}_4 \\
 & \mathbf{u}_5 \leq \mathbf{x}_2 \leq \mathbf{u}_6 \perp (\mathbf{x}_3 - \mathbf{u}_7) \\
 & \mathbf{u}_8 \leq \mathbf{x}_3 \leq \mathbf{u}_9,
 \end{aligned} \tag{5.2.2}$$

onde $\eta(m_j(x_{3_j}))$ é a função indicatriz tal que:

$$\eta(m_j(x_{3_j})) = \begin{cases} 0 & , \text{ se } m_j(x_{3_j}) = 0 \\ +\infty & , \text{ se } m_j(x_{3_j}) \neq 0. \end{cases} \tag{5.2.3}$$

Sejam $\rho = (\rho_1, \dots, \rho_j, \dots, \rho_{n_2})^T \in \mathbb{R}_+^{n_2}$ o vetor de parâmetros de penalidade e uma função $v \in \mathbb{G}$. Assim, o problema de programação discreta-mista com restrições de equilíbrio (5.1.1) é aproximado pela formulação (5.2.4) a seguir:

$$\begin{aligned}
 \text{Min} \quad & \bar{f}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + \sum_{j=1}^{n_2} \rho_j v(m_j(x_{3_j})) \\
 \text{s.a :} \quad & \mathbf{g}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \mathbf{0} \\
 & \mathbf{u}_1 \leq \mathbf{h}_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \leq \mathbf{u}_2 \\
 & \mathbf{u}_3 \leq \mathbf{x}_1 \leq \mathbf{u}_4 \\
 & \mathbf{u}_5 \leq \mathbf{x}_2 \leq \mathbf{u}_6 \perp (\mathbf{x}_3 - \mathbf{u}_7) \\
 & \mathbf{u}_8 \leq \mathbf{x}_3 \leq \mathbf{u}_9.
 \end{aligned} \tag{5.2.4}$$

5.3 Transformação do MPEC com variáveis contínuas em um problema de programação não-linear

Nesta seção, aplicamos os resultados apresentados no Capítulo 4 para transformarmos um MPEC em um problema usual de programação não-linear. Conforme é apresentado nas expressões (4.2.5) e (4.2.6), podemos reformular uma restrição de complementaridade mista como um conjunto usual de restrições de igualdade e de desigualdade ou somente por meio de restrições de desigualdade. Em [Lage 2013], o autor utilizou a formulação (4.2.5) para resolver sua proposição do problema de fluxo de potência ótimo reativo. Neste trabalho, no entanto, utilizamos a formulação (4.2.6). Isso significa que resolver o problema (5.2.4) equivale resolvermos o problema (5.3.1) definido por:

$$\begin{aligned}
 \text{Min} \quad & \bar{f}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + \sum_{j=1}^{n_2} \rho_j v(m_j(x_{3_j})) \\
 \text{s.a :} \quad & \mathbf{g}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \mathbf{0} \\
 & \mathbf{u}_1 \leq \mathbf{h}_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \leq \mathbf{u}_2 \\
 & \mathbf{u}_3 \leq \mathbf{x}_1 \leq \mathbf{u}_4 \\
 & \mathbf{u}_5 \leq \mathbf{x}_2 \leq \mathbf{u}_6 \\
 & \mathbf{u}_8 \leq \mathbf{x}_3 \leq \mathbf{u}_9 \\
 & (\mathbf{x}_2 - \mathbf{u}_5)(\mathbf{x}_3 - \mathbf{u}_7) \leq 0 \\
 & (\mathbf{x}_2 - \mathbf{u}_6)(\mathbf{x}_3 - \mathbf{u}_7) \leq 0
 \end{aligned} \tag{5.3.1}$$

A formulação (5.3.1) é equivalente à formulação (5.3.2), a qual denominamos “forma padrão”, dada por:

$$\begin{aligned}
 \text{Min} \quad & f_\rho(\mathbf{x}) \\
 \text{s.a :} \quad & \mathbf{g}(\mathbf{x}) = \mathbf{0} \\
 & \mathbf{h}(\mathbf{x}) \leq \mathbf{0},
 \end{aligned} \tag{5.3.2}$$

em que: $\mathbf{x} = (\mathbf{x}_1^T, \mathbf{x}_2^T, \mathbf{x}_3^T)^T \in \mathbb{R}^{n_1+n_2+n_3}$; $f_\rho(\mathbf{x}) = f(\mathbf{x}, \rho) = \bar{f}(\mathbf{x}) + \sum_{j=1}^{n_2} \rho_j v(m_j(x_{3_j}))$ e o funcional $\mathbf{h}(\mathbf{x}) \in \mathbb{R}^{2r_1+2n_1+6n_2}$ é expresso por:

$$\mathbf{h}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \begin{pmatrix} -\mathbf{h}_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + \mathbf{u}_1 \\ \mathbf{h}_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) - \mathbf{u}_2 \\ -\mathbf{x}_1 + \mathbf{u}_3 \\ \mathbf{x}_1 - \mathbf{u}_4 \\ -\mathbf{x}_2 + \mathbf{u}_5 \\ \mathbf{x}_2 - \mathbf{u}_6 \\ -\mathbf{x}_3 + \mathbf{u}_8 \\ \mathbf{x}_3 - \mathbf{u}_9 \\ (\mathbf{x}_2 - \mathbf{u}_5)(\mathbf{x}_3 - \mathbf{u}_7) \\ (\mathbf{x}_2 - \mathbf{u}_6)(\mathbf{x}_3 - \mathbf{u}_7) \end{pmatrix}.$$

O conjunto viável associado ao problema (5.3.2) é dado por $\mathbb{X} = \{\mathbf{x} \in \mathbb{R}^n; \mathbf{g}(\mathbf{x}) = \mathbf{0}, \mathbf{h}(\mathbf{x}) \leq \mathbf{0}\}$. Além disso, assumimos que $\bar{f}$, $\mathbf{g}$ e $\mathbf{h}_1$ são funções pelo menos de classe C^2 em $\mathbb{R}^n$ e, por fim, definimos o conjunto $\mathbb{X}^*(\rho^k) = \{\mathbf{x}^*(\rho^k); \mathbf{x}^*(\rho^k) = \arg \min f(\mathbf{x}, \rho^k), \mathbf{x} \in \mathbb{X}\}$ dos minimizadores do problema (5.3.2).

Seja $\mathbf{x}^{k+1} = \mathbf{x}^* (\rho^k)$. Então, de acordo com resultados apresentados no Capítulo 3 sobre o método de penalidades, resolver o problema de programação discreta-mista com restrições de equilíbrio (5.1.1) consiste em resolvermos uma sequência de problemas de programação não-linear (5.3.2) até que $\lim_{k \rightarrow \infty} \left[\sum_{j=1}^{n_2} \rho_{j_k} v \left(m_j \left(x_{3j_k} \right) \right) \right] = 0$, sempre que $\lim_{k \rightarrow \infty} \rho_{j_k} = \infty$.

Neste capítulo, o desenvolvimento do método primal-dual de reescalamiento não-linear será efetuado sobre a formulação (5.3.2). Por fim, até o final deste capítulo, definimos $n = n_1 + 2n_2$, o qual corresponde à dimensão do vetor $\mathbf{x}$; e por $r = 2r_1 + 2n_1 + 6n_2$, o qual corresponde à dimensão do funcional $\mathbf{h}(\mathbf{x})$.

5.4 O método PDRNLI: abordagem primal-dual

Consideremos a forma padrão (5.3.2). Acrescentamos às restrições de desigualdade de (5.3.2) o vetor de variáveis de folga $\mathbf{z}_1 = (z_{1_1}, \dots, z_{1_i}, \dots, z_{1_r})^T \in \mathbb{Z}_1 \subset \mathbb{R}_+^r$. Assim, temos que (5.3.2) é equivalente à (5.4.1).

$$\begin{aligned} \text{Min} \quad & f_\rho(\mathbf{x}) \\ \text{s.a :} \quad & \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ & \mathbf{h}(\mathbf{x}) + \mathbf{z}_1 = \mathbf{0} \\ & -\mathbf{z}_1 \leq \mathbf{0}. \end{aligned} \quad (5.4.1)$$

Sejam $\psi \in \mathbb{H}$ e o parâmetro de reescala $\mu \in \mathbb{R}_{++}$; o parâmetro dual $\delta_1 = (\delta_{1_1}, \dots, \delta_{1_i}, \dots, \delta_{1_r})^T \in \mathbb{R}_{++}^r$; um $\alpha \in [0, 1]$ e; o fator de transformação $\gamma(\delta_{1_i}, \alpha)$. Então, a função lagrangiana aumentada integrada vinculada ao problema (5.4.1) é conforme apresentada em (5.4.2) a seguir:

$$L_\alpha(\omega) = f_\rho(\mathbf{x}) + \mu \sum_{i=1}^r \gamma(\delta_{1_i}, \alpha) \psi \left(\frac{-\delta_{1_i} z_{1_i}}{\mu \gamma(\delta_{1_i}, \alpha)} \right) + \lambda_0^T \mathbf{g}(\mathbf{x}) + \lambda_1^T (\mathbf{h}(\mathbf{x}) + \mathbf{z}_1), \quad (5.4.2)$$

em que: $\omega = (\mathbf{x}^T, \mathbf{z}_1^T, \lambda_0^T, \lambda_1^T, \delta_1^T, \rho^T, \mu)^T \in \mathbb{R}^{n+3r_1+m+n_2+1}$; $\lambda_0 = (\lambda_{0_1}, \dots, \lambda_{0_t}, \dots, \lambda_{0_m})^T \in \mathbb{L}_0 \subseteq \mathbb{R}^m$ é o vetor de multiplicadores de Lagrange vinculado à restrição $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ e $\lambda_1 = (\lambda_{1_1}, \dots, \lambda_{1_i}, \dots, \lambda_{1_r})^T \in \mathbb{L}_1 \subseteq \mathbb{R}_+^r$ é vetor de multiplicadores de Lagrange vinculado à restrição $\mathbf{h}(\mathbf{x}) + \mathbf{z}_1 = \mathbf{0}$.

A condição de estacionaridade primal-dual, KKT, associado ao problema (5.4.1) consiste em resolvermos, na variável $\mathbf{b} = (\mathbf{x}^T, \mathbf{z}_1^T, \lambda_0^T, \lambda_1^T)^T$, o seguinte sistema não-linear primal-dual (5.4.3) dado por:

$$\begin{cases} \nabla_{\mathbf{x}} L_\alpha(\omega) = \nabla f_\rho(\mathbf{x}) + \nabla \mathbf{g}(\mathbf{x})^T \lambda_0 + \nabla \mathbf{h}(\mathbf{x})^T \lambda_1 = \mathbf{0} \\ \nabla_{\mathbf{z}_1} L_\alpha(\omega) = -\dot{\Psi}(-\mu^{-1} \Delta_\alpha^{-1} \Delta_1 \mathbf{z}_1) \delta_1 + \lambda_1 = \mathbf{0} \\ \nabla_{\lambda_0} L_\alpha(\omega) = \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ \nabla_{\lambda_1} L_\alpha(\omega) = \mathbf{h}(\mathbf{x}) + \mathbf{z}_1 = \mathbf{0} \end{cases} \quad (5.4.3)$$

em que:

$$\nabla \mathbf{g}(\mathbf{x}) = \begin{pmatrix} \nabla g_1(\mathbf{x}) & \cdots & \nabla g_t(\mathbf{x}) & \cdots & \nabla g_m(\mathbf{x}) \end{pmatrix}^T \in \mathbb{R}^{m \times n} \quad (5.4.4)$$

é a matriz jacobiana associada ao funcional $\mathbf{g}(\mathbf{x})$;

$$\nabla \mathbf{h}(\mathbf{x}) = \begin{pmatrix} \nabla h_1(\mathbf{x}) & \cdots & \nabla h_i(\mathbf{x}) & \cdots & \nabla h_r(\mathbf{x}) \end{pmatrix}^T \in \mathbb{R}^{r \times n} \quad (5.4.5)$$

é a matriz jacobiana vinculada ao funcional $\mathbf{h}(\mathbf{x})$;

$$\dot{\Psi}(-\mu^{-1}\Delta_{\alpha}^{-1}\Delta_1\mathbf{z}_1) = \text{diag}\left(\psi'\left(\frac{-\delta_{1_i}z_{1_i}}{\mu\gamma(\delta_{1_i}, \alpha)}\right)\right)_{i=1}^r; \quad (5.4.6)$$

$$\Delta_{\alpha} = \text{diag}(\gamma(\delta_{1_i}, \alpha))_{i=1}^r \quad (5.4.7)$$

e;

$$\Delta_1 = \text{diag}(\delta_{1_i})_{i=1}^r. \quad (5.4.8)$$

Como $\psi'(t) \neq 0$ para todo t que está no domínio de ψ , então, se o argumento $-(\mu\gamma(\delta_{1_i}, \alpha))^{-1}\delta_{1_i}z_{1_i}$, para todo $i = 1, \dots, r$, está no domínio de ψ , então a matriz

$$\dot{\Psi}^{-1}(-\mu^{-1}\Delta_{\alpha}^{-1}\Delta_1\mathbf{z}_1) = \text{diag}\left(\frac{1}{\psi'\left(\frac{-\delta_{1_i}z_{1_i}}{\mu\gamma(\delta_{1_i}, \alpha)}\right)}\right)_{i=1}^r \quad (5.4.9)$$

está bem definida.

Embasados em [Pinheiro et al. 2015], modificamos a formulação do sistema não-linear (5.4.3) para a forma equivalente apresentada em (5.4.10) adiante. Esta formulação é o nosso principal instrumento de trabalho com a abordagem predictor-corrector junto ao método PDRNLI.

$$\begin{cases} \nabla_{\mathbf{x}} L_{\alpha}(\omega) = \nabla f_{\rho}(\mathbf{x}) + \nabla \mathbf{g}(\mathbf{x})^T \lambda_0 + \nabla \mathbf{h}(\mathbf{x})^T \lambda_1 = \mathbf{0} \\ \nabla_{\mathbf{z}_1} L_{\alpha}(\omega) = \dot{\Psi}^{-1}(-\mu^{-1}\Delta_{\alpha}^{-1}\Delta_1\mathbf{z}_1) \lambda_1 - \delta_1 = \mathbf{0} \\ \nabla_{\lambda_0} L_{\alpha}(\omega) = \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ \nabla_{\lambda_1} L_{\alpha}(\omega) = \mathbf{h}(\mathbf{x}) + \mathbf{z}_1 = \mathbf{0} \end{cases} \quad (5.4.10)$$

Consideremos o caso particular em que definimos $\alpha = 1$, o qual implica que $\gamma(\delta_{1_i}, \alpha) = \delta_{1_i}$; assumimos que o vetor ρ seja nulo e; escolhemos a função barreira logarítmica modificada dada por $\psi(t) = -\ln(-t+1)$. Então, sob essas condições, é fácil mostrar que o sistema não-linear (5.4.10) do método PDRNLI torna-se o sistema não-linear primal-dual apresentado em [Pinheiro et al. 2015, eq. 33].

De fato, exceto pela segunda equação de (5.4.10), as equações do sistema não-linear do método PDRNLI são idênticas às três primeiras de (33) em [Pinheiro et al. 2015]. Ao definirmos $\alpha = 1$ temos que $\Delta_{\alpha}^{-1}\Delta_1 = \text{diag}\left((\delta_{1_i})^{-1}\delta_{1_i}\right)_{i=1}^r = I_r$, onde I_r é a matriz identidade de ordem r . Como $\psi(t) = -\ln(-t+1)$, então $\psi'(t) = \frac{1}{-t+1}$. Isso significa que $\psi'\left(\frac{-\delta_{1_i}z_{1_i}}{\mu\delta_{1_i}}\right) = \frac{\mu}{z_{1_i}+\mu}$, logo, por (5.4.9), temos $\dot{\Psi}^{-1}(-\mu^{-1}\Delta_{\alpha}^{-1}\Delta_1\mathbf{z}_1) = \mu^{-1}\text{diag}(z_{1_i}+\mu)_{i=1}^r$. Assim, a segunda equação de (5.4.10) é expressa por $\mu^{-1}\text{diag}(z_{1_i}+\mu)_{i=1}^r \lambda_1 - \delta_1 = \mathbf{0}$, a qual é exatamente a quarta equação de (33) em [Pinheiro et al. 2015].

5.5 O método PDRNLI: dedução do vetor de direções de busca primal-dual

Neste trabalho, nós assumimos que o vetor de parâmetros $\sigma^{\mathbf{k}} = (\delta_1^{\mathbf{k}^T}, \rho^{\mathbf{k}^T}, \mu_k)^T$ seja fixado enquanto ocorre um ciclo completo do método PDRNLI. Seja o vetor $\mathbf{b}^{\mathbf{k}} = (\mathbf{x}^{\mathbf{k}^T}, \mathbf{z}_1^{\mathbf{k}^T}, \lambda_0^{\mathbf{k}^T}, \lambda_1^{\mathbf{k}^T})^T$

uma boa estimativa primal-dual para resolvermos o sistema (5.4.10). Um ciclo completo do método PDRNLI consiste em encontrar $\mathbf{b}^{k+1} = \left(\mathbf{x}^{k+1^T}, \mathbf{z}_1^{k+1^T}, \lambda_0^{k+1^T}, \lambda_1^{k+1^T} \right)^T$ tal que:

$$\begin{cases} \nabla_{\mathbf{x}} L_{\alpha} \left(\mathbf{b}^{k+1^T}, \sigma^{\mathbf{k}^T} \right) = \nabla f_{\rho^k}(\mathbf{x}^{k+1}) + \nabla \mathbf{g}(\mathbf{x}^{k+1})^T \lambda_0^{k+1} + \nabla \mathbf{h}(\mathbf{x}^{k+1})^T \lambda_1^{k+1} = \mathbf{0} \\ \nabla_{\mathbf{z}_1} L_{\alpha} \left(\mathbf{b}^{k+1^T}, \sigma^{\mathbf{k}^T} \right) = \dot{\Psi}^{-1} \left(-\mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \mathbf{z}_1^{k+1} \right) \lambda_1^{k+1} - \delta_1^{\mathbf{k}} = \mathbf{0} \\ \nabla_{\lambda_0} L_{\alpha} \left(\mathbf{b}^{k+1^T}, \sigma^{\mathbf{k}^T} \right) = \mathbf{g}(\mathbf{x}^{k+1}) = \mathbf{0} \\ \nabla_{\lambda_1} L_{\alpha} \left(\mathbf{b}^{k+1^T}, \sigma^{\mathbf{k}^T} \right) = \mathbf{h}(\mathbf{x}^{k+1}) + \mathbf{z}_1^{k+1} = \mathbf{0}. \end{cases}$$

Nós vamos aplicar o método de Newton para resolvermos o sistema (5.4.10) para obtermos $\mathbf{b}^{k+1}$ usando o ponto $\mathbf{b}^k$ como ponto inicial. Assumamos que $\mathbf{b}^{k+1} = \mathbf{b}^k + \mathbf{d}_b^k$, onde $\mathbf{d}_b^k$ é o vetor de direções de busca primal-dual definido por $\mathbf{d}_b^k = \left(\mathbf{d}_x^k, \mathbf{d}_{z_1}^k, \mathbf{d}_{\lambda_0}^k, \mathbf{d}_{\lambda_1}^k \right)^T$. Vamos linearizar cada uma das equações do sistema (5.4.10) para obtermos o sistema de Newton que nos permitirá explicitar cada uma das componentes do vetor de direções de busca $\mathbf{d}_b^k$, o qual é o principal objetivo desta seção. Iniciamos o procedimento da dedução das direções de busca com a linearização da expressão (5.5.1), a qual é dada por

$$\nabla_{\mathbf{x}} L_{\alpha} \left(\mathbf{b}^T, \sigma^{\mathbf{k}^T} \right) = \nabla f_{\rho^k}(\mathbf{x}) + \nabla \mathbf{g}(\mathbf{x})^T \lambda_0 + \nabla \mathbf{h}(\mathbf{x})^T \lambda_1. \quad (5.5.1)$$

Por assumirmos que

$$\nabla f_{\rho^k}(\mathbf{x}^{k+1}) + \nabla \mathbf{g}(\mathbf{x}^{k+1})^T \lambda_0^{k+1} + \nabla \mathbf{h}(\mathbf{x}^{k+1})^T \lambda_1^{k+1} = \mathbf{0} \quad (5.5.2)$$

e que $\mathbf{b}^{k+1} = \mathbf{b}^k + \mathbf{d}_b^k$, então ao desprezarmos os termos de ordem superior, temos:

$$\begin{aligned} \nabla f_{\rho^k}(\mathbf{x}^k) + \nabla^2 f_{\rho^k}(\mathbf{x}^k) \mathbf{d}_x^k + \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^k + \nabla \mathbf{g}(\mathbf{x}^k)^T \mathbf{d}_{\lambda_0}^k + \sum_{t=1}^m \lambda_{0_{t_k}} \nabla^2 g_t(\mathbf{x}^k) \mathbf{d}_x^k + \\ + \nabla \mathbf{h}(\mathbf{x}^k)^T \lambda_1^k + \nabla \mathbf{h}(\mathbf{x}^k)^T \mathbf{d}_{\lambda_1}^k + \sum_{i=1}^r \lambda_{1_{i_k}} \nabla^2 h_i(\mathbf{x}^k) \mathbf{d}_x^k = \mathbf{0}. \end{aligned}$$

Seja:

$$\mathbb{R}^{n \times n} \ni \mathcal{K}_k = \nabla^2 f_{\rho^k}(\mathbf{x}^k) + \sum_{t=1}^m \lambda_{0_{t_k}} \nabla^2 g_t(\mathbf{x}^k) + \sum_{i=1}^r \lambda_{1_{i_k}} \nabla^2 h_i(\mathbf{x}^k), \quad (5.5.3)$$

então a linearização de (5.5.2) é expressa conforme (5.5.4) a seguir:

$$\mathcal{K}_k \mathbf{d}_x^k + \nabla \mathbf{g}(\mathbf{x}^k)^T \mathbf{d}_{\lambda_0}^k + \nabla \mathbf{h}(\mathbf{x}^k)^T \mathbf{d}_{\lambda_1}^k = -\nabla f_{\rho^k}(\mathbf{x}^k) - \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^k - \nabla \mathbf{h}(\mathbf{x}^k)^T \lambda_1^k. \quad (5.5.4)$$

Prosseguimos, agora, com a linearização da expressão (5.5.5) dada por:

$$\nabla_{\mathbf{z}_1} L_{\alpha} \left(\mathbf{b}^T, \sigma^{\mathbf{k}^T} \right) = \dot{\Psi}^{-1} \left(-\mu^{-1} \Delta_{\alpha}^{-1} \Delta_1 \mathbf{z}_1 \right) \lambda_1 - \delta_1 = \mathbf{0}. \quad (5.5.5)$$

Novamente, por assumirmos que

$$\dot{\Psi}^{-1} \left(-\mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \mathbf{z}_1^{k+1} \right) \lambda_1^{k+1} - \delta_1^{\mathbf{k}} = \mathbf{0} \quad (5.5.6)$$

e que $\mathbf{b}^{k+1} = \mathbf{b}^k + \mathbf{d}_b^k$, então ao desprezarmos os termos de ordem superior, temos:

$$\left(\dot{\Psi}^{-1} \left(-\mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \mathbf{z}_1^k \right) + \mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \dot{\Psi}^{-2} \left(-\mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \mathbf{z}_1^k \right) \ddot{\Psi} \left(-\mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \mathbf{z}_1^k \right) D_{z_1^k} \right) (\lambda_1^k + \mathbf{d}_{\lambda_1}^k) = \delta_1^{\mathbf{k}}, \quad (5.5.7)$$

onde:

$$\ddot{\Psi} \left(-\mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \mathbf{z}_1^k \right) = \text{diag} \left(\psi'' \left(\frac{\delta_{1_{i_k}} z_{1_{i_k}}}{\mu_k \gamma \left(\delta_{1_{i_k}}, \alpha \right)} \right) \right)_{i=1}^r \quad (5.5.8)$$

e que

$$D_{z_1^k} = \text{diag} \left(d_{z_{1_{i_k}}} \right)_{i=1}^r. \quad (5.5.9)$$

Por simplicidade, abreviamos a notação da matriz $\dot{\Psi}^{-1} \left(-\mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \mathbf{z}_1^k \right)$ para $\dot{\Psi}^{-1}(\cdot)$ e da matriz $\ddot{\Psi} \left(-\mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \mathbf{z}_1^k \right)$ para $\ddot{\Psi}(\cdot)$. Portanto, ao reorganizarmos os termos de (5.5.7), a linearização de (5.5.6) é expressa em (5.5.10), a qual é definida por:

$$\begin{aligned} \mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \dot{\Psi}^{-2}(\cdot) \ddot{\Psi}(\cdot) \Lambda_{1_k} \mathbf{d}_{z_1}^k + \dot{\Psi}^{-1}(\cdot) \mathbf{d}_{\lambda_1}^k = \\ = -\dot{\Psi}^{-1}(\cdot) \lambda_1^k + \delta_1^k - \mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \dot{\Psi}^{-2}(\cdot) \ddot{\Psi}(\cdot) D_{z_1^k} \mathbf{d}_{\lambda_1}^k, \end{aligned} \quad (5.5.10)$$

em que:

$$\Lambda_{1_k} = \text{diag} \left(\lambda_{1_{i_k}} \right)_{i=1}^r. \quad (5.5.11)$$

Continuamos, agora, com a linearização da expressão (5.5.12) dada por:

$$\nabla_{\lambda_0} L_{\alpha} \left(\mathbf{b}^T, \sigma^{\mathbf{k}^T} \right) = \mathbf{g}(\mathbf{x}) = \mathbf{0}. \quad (5.5.12)$$

Mais uma vez, por assumirmos que

$$\mathbf{g} \left(\mathbf{x}^{\mathbf{k}+1} \right) = \mathbf{0} \quad (5.5.13)$$

e que $\mathbf{b}^{\mathbf{k}+1} = \mathbf{b}^{\mathbf{k}} + \mathbf{d}_{\mathbf{b}}^{\mathbf{k}}$, então ao desprezarmos os termos de ordem superior, obtemos imediatamente que a linearização de (5.5.13) é de acordo com a expressão (5.5.14) a seguir:

$$\nabla \mathbf{g} \left(\mathbf{x}^{\mathbf{k}} \right) \mathbf{d}_{\mathbf{x}}^{\mathbf{k}} = -\mathbf{g} \left(\mathbf{x}^{\mathbf{k}} \right). \quad (5.5.14)$$

Finalizamos o procedimento de linearização do sistema (5.4.10) com a linearização da expressão (5.5.15) dada por:

$$\nabla_{\lambda_1} L_{\alpha} \left(\mathbf{b}^T, \sigma^{\mathbf{k}^T} \right) = \mathbf{h}(\mathbf{x}) + \mathbf{z}_1 = \mathbf{0}. \quad (5.5.15)$$

Por assumirmos que

$$\mathbf{h} \left(\mathbf{x}^{\mathbf{k}+1} \right) + \mathbf{z}_1^{\mathbf{k}+1} = \mathbf{0} \quad (5.5.16)$$

e que $\mathbf{b}^{\mathbf{k}+1} = \mathbf{b}^{\mathbf{k}} + \mathbf{d}_{\mathbf{b}}^{\mathbf{k}}$, então ao desprezarmos os termos de ordem superior, obtemos imediatamente que a linearização de (5.5.16) é de acordo com a expressão (5.5.17) a seguir:

$$\nabla \mathbf{h} \left(\mathbf{x}^{\mathbf{k}} \right) \mathbf{d}_{\mathbf{x}}^{\mathbf{k}} + \mathbf{d}_{z_1}^{\mathbf{k}} = -\mathbf{h} \left(\mathbf{x}^{\mathbf{k}} \right) - \mathbf{z}_1^{\mathbf{k}}. \quad (5.5.17)$$

As equações (5.5.4), (5.5.10), (5.5.14) e (5.5.17) definem o sistema de Newton (5.5.18) do método PDRNLI.

$$A_k \mathbf{d}_{\mathbf{b}}^{\mathbf{k}} = \mathbf{a}^{\mathbf{k}}, \quad (5.5.18)$$

em que:

$$A_k = \begin{pmatrix} \mathcal{K}_k & 0 & \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}})^T & \nabla \mathbf{h}(\mathbf{x}^{\mathbf{k}})^T \\ 0 & \mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \dot{\Psi}^{-2}(\cdot) \ddot{\Psi}(\cdot) \Lambda_{1_k} & 0 & \dot{\Psi}^{-1}(\cdot) \\ \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}) & 0 & 0 & 0 \\ \nabla \mathbf{h}(\mathbf{x}^{\mathbf{k}}) & I_r & 0 & 0 \end{pmatrix} \quad (5.5.19)$$

é a matriz dos coeficientes do sistema (5.5.19) e

$$\mathbf{a}^k = \begin{pmatrix} -\nabla f_{\rho^k}(\mathbf{x}^k) - \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^k - \nabla \mathbf{h}(\mathbf{x}^k)^T \lambda_1^k \\ -\dot{\Psi}^{-1}(\cdot) \lambda_1^k + \delta_1^k - \mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \dot{\Psi}^{-2}(\cdot) \ddot{\Psi}(\cdot) D_{z_1^k} \mathbf{d}_{\lambda_1}^k \\ -\mathbf{g}(\mathbf{x}^k) \\ -\mathbf{h}(\mathbf{x}^k) - \mathbf{z}_1^k \end{pmatrix} \quad (5.5.20)$$

é o vetor dos termos independentes.

O cálculo para explicitarmos o vetor de direções de busca primal-dual em função dos dados de entrada é desenvolvido a seguir. Ao menos duas questões nos motivam para este cálculo: nos permite explorar a esparsidade da matriz dos coeficientes A_k e a identificação de estruturas algébricas, as quais nos permitem uma interpretação geométrica e o desenvolvimento de estratégias adicionais que visam a robustez do método PDRNLI.

Iniciamos o cálculo do vetor das direções pela expressão (5.5.17). É imediato constatarmos que a direção de busca primal $\mathbf{d}_{z_1}^k$ é expressa por:

$$\mathbf{d}_{z_1}^k = -\nabla \mathbf{h}(\mathbf{x}^k) \mathbf{d}_x^k - \mathbf{h}(\mathbf{x}^k) - \mathbf{z}_1^k. \quad (5.5.21)$$

Para explicitarmos a direção de busca dual $\mathbf{d}_{\lambda_1}^k$ é necessário um procedimento algébrico intermediário. Considere a expressão (5.5.10). Ao multiplicarmos ambos os lados desta equação pela matriz $\dot{\Psi}(\cdot)$, então obtemos (5.5.22) dada por:

$$\mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \dot{\Psi}^{-1}(\cdot) \ddot{\Psi}(\cdot) \Lambda_{1_k} \mathbf{d}_{z_1}^k + \mathbf{d}_{\lambda_1}^k = -\lambda_1^k + \dot{\Psi}(\cdot) \delta_1^k - \mu_k^{-1} \Delta_{\alpha_k}^{-1} \Delta_{1_k} \dot{\Psi}^{-1}(\cdot) \ddot{\Psi}(\cdot) D_{z_1^k} \mathbf{d}_{\lambda_1}^k. \quad (5.5.22)$$

Definição 5.5.1 (Matriz afim-escala principal). *Definimos por matriz afim-escala principal à matriz diagonal $\bar{Z}_{1_k} \in \mathbb{R}^{r \times r}$ expressa por:*

$$\bar{Z}_{1_k} = \Delta_{\alpha_k}^{-1} \Delta_{1_k} \dot{\Psi}^{-1}(\cdot) \ddot{\Psi}(\cdot). \quad (5.5.23)$$

□

Ao substituírmos (5.5.21) em (5.5.22) e reorganizarmos os termos algébricos, então a direção de busca dual $\mathbf{d}_{\lambda_1}^k$ é conforme (5.5.24) a seguir:

$$\mathbf{d}_{\lambda_1}^k = \mu_k^{-1} \bar{Z}_{1_k} \Lambda_{1_k} \nabla \mathbf{h}(\mathbf{x}^k) \mathbf{d}_x^k - \mu_k^{-1} \bar{Z}_{1_k} \Lambda_{1_k} (-\mathbf{h}(\mathbf{x}^k) - \mathbf{z}_1^k) - \lambda_1^k + \dot{\Psi}(\cdot) \delta_1^k - \mu_k^{-1} \bar{Z}_{1_k} D_{z_1^k} \mathbf{d}_{\lambda_1}^k. \quad (5.5.24)$$

O próximo passo consiste em substituírmos a expressão (5.5.24) em (5.5.4). O resultado desta operação é expressa em (5.5.25) adiante.

$$\theta_k \mathbf{d}_x^k + \nabla \mathbf{g}(\mathbf{x}^k)^T \mathbf{d}_{\lambda_0}^k = \mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^k, \quad (5.5.25)$$

onde:

$$\theta_k = \nabla^2 f_{\rho^k}(\mathbf{x}^k) + \sum_{t=1}^m \lambda_{0_{t_k}} \nabla g_t(\mathbf{x}^k) + \sum_{i=1}^r \lambda_{1_{i_k}} \nabla h_i(\mathbf{x}^k) + \mu_k^{-1} \nabla \mathbf{h}(\mathbf{x}^k)^T \bar{Z}_{1_k} \Lambda_{1_k} \nabla \mathbf{h}(\mathbf{x}^k) \quad (5.5.26)$$

e

$$\mathbf{v}_\mu^k = -\nabla f_{\rho^k}(\mathbf{x}^k) - \nabla \mathbf{h}(\mathbf{x}^k)^T \left(\dot{\Psi}(\cdot) \delta_1^k - \mu_k^{-1} \bar{Z}_{1_k} \Lambda_{1_k} \left(-\mathbf{h}(\mathbf{x}^k) - \mathbf{z}_1^k \right) - \mu_k^{-1} \bar{Z}_{1_k} D_{z_1^k} \mathbf{d}_{\lambda_1}^k \right). \quad (5.5.27)$$

As expressões (5.5.25) e (5.5.14) definem o seguinte sistema aumentado de KKT¹ dado por:

$$\begin{pmatrix} \theta_k & \nabla \mathbf{g}(\mathbf{x}^k)^T \\ \nabla \mathbf{g}(\mathbf{x}^k) & 0 \end{pmatrix} \begin{pmatrix} \mathbf{d}_x^k \\ \mathbf{d}_{\lambda_0}^k \end{pmatrix} = \begin{pmatrix} \mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^k \\ -\mathbf{g}(\mathbf{x}^k) \end{pmatrix}. \quad (5.5.28)$$

Para podermos resolver o sistema (5.5.28), assumamos que θ_k admite inversa e que posto $(\nabla \mathbf{g}(\mathbf{x}^k)) = m$. Com esses pressupostos, os procedimentos algébricos para explicitarmos a direção de busca dual $\mathbf{d}_{\lambda_0}^k$ são efetuados a partir de (5.5.25).

$$\begin{aligned} \theta_k \mathbf{d}_x^k + \nabla \mathbf{g}(\mathbf{x}^k)^T \mathbf{d}_{\lambda_0}^k &= \mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^k \\ \nabla \mathbf{g}(\mathbf{x}^k)^T \mathbf{d}_{\lambda_0}^k &= \mathbf{v}_\mu^k - \theta_k \mathbf{d}_x^k - \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^k. \end{aligned}$$

Como a matriz θ_k admite inversa, então:

$$\begin{aligned} \nabla \mathbf{g}(\mathbf{x}^k)^T \mathbf{d}_{\lambda_0}^k &= \mathbf{v}_\mu^k - \theta_k \mathbf{d}_x^k - \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^k \\ \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \mathbf{d}_{\lambda_0}^k &= \theta_k^{-1} \mathbf{v}_\mu^k - \mathbf{d}_x^k - \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^k \\ \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \mathbf{d}_{\lambda_0}^k &= \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k) \mathbf{d}_x^k - \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^k. \end{aligned}$$

Agora, por (5.5.14), temos:

$$\begin{aligned} \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \mathbf{d}_{\lambda_0}^k &= \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k) \mathbf{d}_x^k - \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^k \\ \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \mathbf{d}_{\lambda_0}^k &= \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \mathbf{v}_\mu^k + \mathbf{g}(\mathbf{x}^k) - \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^k \end{aligned}$$

A inversa da matriz $\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T$ está bem definida, pelo fato de que posto $(\nabla \mathbf{g}(\mathbf{x}^k)) = m$. Assim, a direção de busca dual $\mathbf{d}_{\lambda_0}^k$ é expressa conforme (5.5.29) a seguir:

$$\mathbf{d}_{\lambda_0}^k = \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right)^{-1} \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \mathbf{v}_\mu^k + \mathbf{g}(\mathbf{x}^k) \right) - \lambda_0^k. \quad (5.5.29)$$

Por fim, segue imediatamente a partir de (5.5.25) que a direção de busca primal $\mathbf{d}_x^k$ é expressa por:

$$\mathbf{d}_x^k = \theta_k^{-1} \left(\mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k)^T \left(\lambda_0^k + \mathbf{d}_{\lambda_0}^k \right) \right). \quad (5.5.30)$$

Em resumo, as direções de busca primal-dual $\mathbf{d}_b^k$ do método PDRNLI é conforme (5.5.31):

$$\begin{cases} \mathbf{d}_{\lambda_0}^k = \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right)^{-1} \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \mathbf{v}_\mu^k + \mathbf{g}(\mathbf{x}^k) \right) - \lambda_0^k \\ \mathbf{d}_x^k = \theta_k^{-1} \left(\mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k)^T \left(\lambda_0^k + \mathbf{d}_{\lambda_0}^k \right) \right) \\ \mathbf{d}_{z_1}^k = -\nabla \mathbf{h}(\mathbf{x}^k) \mathbf{d}_x^k - \mathbf{h}(\mathbf{x}^k) - \mathbf{z}_1^k \\ \mathbf{d}_{\lambda_1}^k = -\mu_k^{-1} \bar{Z}_{1_k} \Lambda_{1_k} \mathbf{d}_{z_1}^k + \dot{\Psi}(\cdot) \delta_1^k - \mu_k^{-1} \bar{Z}_{1_k} D_{z_1^k} \mathbf{d}_{\lambda_1}^k - \lambda_1^k \end{cases} \quad (5.5.31)$$

¹Compare-o com o sistema linear apresentado em (A.3.6).

5.6 O método PDRNLI: uma abordagem previsor-corretor

A principal dificuldade para o cálculo do vetor de direções de busca primal-dual $\mathbf{d}_b^k$ do método PDRNLI é a presença do termo quadrático $D_{z_1^k} \mathbf{d}_{\lambda_1}^k$ em (5.5.31), cujas componentes são desconhecidas *a priori*. Esta dificuldade é a principal motivação para uma abordagem previsor-corretor junto ao método PDRNLI. Neste trabalho, seguimos a abordagem proposta em [Pinheiro et al. 2015], a qual é uma variante das abordagens previsor-corretor desenvolvidas em [Mehrotra 1992, Lustig, Marsten e Shanno 1992, Lustig, Marsten e Shanno 1994, Wu, Debs e Marsten 1993] para o tradicional método de pontos interiores e da abordagem proposta em [Sousa, Baptista e Costa 2009] para o método da função barreira modificada.

O procedimento previsor do método PDRNLI consiste em calcularmos o vetor de direções de busca primal-dual com o pressuposto de que $D_{z_1^k} \mathbf{d}_{\lambda_1}^k = \mathbf{0}$ com o parâmetro de reescala $\mu_k \neq 0$. Assim, as direções de busca-primal-dual do método PDRNLI são dadas por:

$$\begin{cases} \mathbf{d}_{\lambda_0}^k = \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right)^{-1} \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \bar{\mathbf{v}}_\mu^k + \mathbf{g}(\mathbf{x}^k) \right) - \lambda_0^k \\ \mathbf{d}_x^k = \theta_k^{-1} \left(\bar{\mathbf{v}}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k)^T (\lambda_0^k + \mathbf{d}_{\lambda_0}^k) \right) \\ \mathbf{d}_{z_1}^k = -\nabla \mathbf{h}(\mathbf{x}^k) \mathbf{d}_x^k - \mathbf{h}(\mathbf{x}^k) - \mathbf{z}_1^k \\ \mathbf{d}_{\lambda_1}^k = -\mu_k^{-1} \bar{Z}_{1k} \Lambda_{1k} \mathbf{d}_{z_1}^k + \dot{\Psi}(\cdot) \delta_1^k - \lambda_1^k, \end{cases} \quad (5.6.1)$$

onde

$$\bar{\mathbf{v}}_\mu^k = -\nabla f_{\rho^k}(\mathbf{x}^k) - \nabla \mathbf{h}(\mathbf{x}^k)^T \left(\dot{\Psi}(\cdot) \delta_1^k - \mu_k^{-1} \bar{Z}_{1k} \Lambda_{1k} (-\mathbf{h}(\mathbf{x}^k) - \mathbf{z}_1^k) \right). \quad (5.6.2)$$

O procedimento corretor tem o propósito de corrigir todas as componentes do vetor de direção de busca primal-dual $\mathbf{d}_b^k$ apresentada em (5.6.1). O procedimento é viável pois, desta vez, conhecemos o termo quadrático $D_{z_1^k} \mathbf{d}_{\lambda_1}^k$. As componentes do vetor de direções de busca primal-dual obtidas pelo procedimento corretor são definidas por $\tilde{\mathbf{d}}_b^k$, a qual é dada por:

$$\begin{cases} \tilde{\mathbf{d}}_{\lambda_0}^k = \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right)^{-1} \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \mathbf{v}_\mu^k + \mathbf{g}(\mathbf{x}^k) \right) - \lambda_0^k \\ \tilde{\mathbf{d}}_x^k = \theta_k^{-1} \left(\mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k)^T (\lambda_0^k + \mathbf{d}_{\lambda_0}^k) \right) \\ \tilde{\mathbf{d}}_{z_1}^k = -\nabla \mathbf{h}(\mathbf{x}^k) \mathbf{d}_x^k - \mathbf{h}(\mathbf{x}^k) - \mathbf{z}_1^k \\ \tilde{\mathbf{d}}_{\lambda_1}^k = -\mu_k^{-1} \bar{Z}_{1k} \Lambda_{1k} \mathbf{d}_{z_1}^k + \dot{\Psi}(\cdot) \delta_1^k - \mu_k^{-1} \bar{Z}_{1k} D_{z_1^k} \mathbf{d}_{\lambda_1}^k - \lambda_1^k. \end{cases} \quad (5.6.3)$$

Após o cálculo das direções de busca $\tilde{\mathbf{d}}_b^k$, então o novo ponto $\mathbf{b}^{k+1}$ é definido por $\mathbf{b}^k + \tilde{\mathbf{d}}_b^k$.

5.7 O método PDRNLI como um método de gradiente projetado

Para mostrarmos que o método PDRNLI é um método de gradiente projetado, utilizamos os resultados apresentados no Apêndice A.2. Nesta subseção assumimos que a matriz simétrica θ_k , expressão (5.5.26), é apenas não-singular e que *posto* $(\nabla \mathbf{g}(\mathbf{x}^k)) = m$.

Além disso, se θ_k possui a_{1k} autovalores positivos e $0 \neq a_{2k}$ autovalores negativos tal que $a_{1k} + a_{2k} = n$ (com todas as multiplicidades algébricas), então a função $\phi_{\theta_k}(\mathbf{y}) = \mathbf{y}^T \theta_k \mathbf{y}$ é denominada

forma quadrática isotrópica regular [Lam 2004], a qual pode ser interpretada como uma generalização do conceito de “quadrado da magnitude do vetor $\mathbf{y}$ ”. Caso contrário, isto é, se $a_{1_k} = n$ ou se $a_{2_k} = n$, então $\phi_{\theta_k}(\mathbf{y})$ é denominada forma quadrática anisotrópica regular. Observe que se $a_{1_k} = n$, então a função $\sqrt{\phi_{\theta_k}(\mathbf{y})}$ define uma norma em $\mathbb{R}^n$.

É fácil constatar-mos que a partir do vetor $\tilde{\mathbf{d}}_{\lambda_0}^k$, expresso em (5.6.3), nós obtemos λ_0^{k+1} através de (5.7.1):

$$\lambda_0^{k+1} = \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right)^{-1} \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \mathbf{v}_\mu^k + \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right)^{-1} \mathbf{g}(\mathbf{x}^k). \quad (5.7.1)$$

De acordo com (A.3.9), página 217, se identificarmos N por $\nabla \mathbf{g}(\mathbf{x}^k)$, M por θ_k e $\mathbf{v}$ por $\mathbf{v}_\mu^k$, então, a menos do vetor residual $\left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right) \mathbf{g}(\mathbf{x}^k)$, temos que a expressão (5.7.1) nos permite determinar um vetor $\mathbf{w}^k \in \text{Im} \left(\nabla \mathbf{g}(\mathbf{x}^k)^T \right) \subset \mathbb{R}^n$ tal que o quadrado da magnitude do vetor $\left(\mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^{k+1} \right)$, o qual é definido pela forma quadrática isotrópica regular

$$\phi_{\theta_k^{-1}} \left(\mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^{k+1} \right) = \left(\mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^{k+1} \right)^T \theta_k^{-1} \left(\mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^{k+1} \right),$$

com assinatura (a_{1_k}, a_{2_k}) , é a menor possível. De fato, a expressão $\nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^{k+1}$ é explícita por:

$$\begin{aligned} \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^{k+1} &= \nabla \mathbf{g}(\mathbf{x}^k)^T \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right)^{-1} \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \mathbf{v}_\mu^k + \\ &\quad + \nabla \mathbf{g}(\mathbf{x}^k)^T \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right)^{-1} \mathbf{g}(\mathbf{x}^k). \end{aligned}$$

De acordo com a Proposição A.3.2 e com a Definição A.3.3, identificamos o operador linear que nos permite projetar obliquamente o $\mathbf{v}_\mu^k$ em $\mathbf{w}^k \in \text{Im} \left(\nabla \mathbf{g}(\mathbf{x}^k)^T \right)$. Este operador é explicitado em (5.7.2) a seguir:

$$P_{\text{Im} \left(\nabla \mathbf{g}(\mathbf{x}^k)^T \right)_{\theta_k^{-1}}} = \nabla \mathbf{g}(\mathbf{x}^k)^T \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right)^{-1} \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1}, \quad (5.7.2)$$

Logo, é imediato constatar-mos que

$$\nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^{k+1} = P_{\text{Im} \left(\nabla \mathbf{g}(\mathbf{x}^k)^T \right)_{\theta_k^{-1}}} \mathbf{v}_\mu^k + \nabla \mathbf{g}(\mathbf{x}^k)^T \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right)^{-1} \mathbf{g}(\mathbf{x}^k). \quad (5.7.3)$$

É interessante que a direção $\tilde{\mathbf{d}}_{\mathbf{x}}^k$ contenha informações que condizem com o $\ker(\nabla \mathbf{g}(\mathbf{x}^k)) \subset \mathbb{R}^n$ para nos assegurarmos que esta direção de busca primal é, de fato, uma direção viável. Para compreendermos a importância deste subespaço no contexto da otimização, consideremos a segunda linha do sistema de KKT (5.5.28), a qual é dada por:

$$\nabla \mathbf{g}(\mathbf{x}^k) \mathbf{d}_{\mathbf{x}}^k = -\mathbf{g}(\mathbf{x}^k).$$

Suponha que $\mathbf{x}^k$ seja um ponto viável, isto é, $\mathbf{g}(\mathbf{x}^k) = \mathbf{0}$. Então, a condição de trasladarmos de $\mathbf{x}^k$ para $\mathbf{x}^{k+1}$ sem abandonarmos a região viável é que $\nabla \mathbf{g}(\mathbf{x}^k) \mathbf{d}_{\mathbf{x}}^k = \mathbf{0}$, ou seja, $\mathbf{d}_{\mathbf{x}}^k \in \ker(\nabla \mathbf{g}(\mathbf{x}^k))$.

Como conhecemos $\tilde{\mathbf{d}}_{\mathbf{x}}^k$ através de (5.6.3) e que λ_0^{k+1} é conforme (5.7.1), temos:

$$\begin{aligned} \tilde{\mathbf{d}}_{\mathbf{x}}^k &= \theta_k^{-1} \left(\mathbf{v}_\mu^k - \nabla \mathbf{g}(\mathbf{x}^k)^T \lambda_0^{k+1} \right) \Rightarrow \\ \tilde{\mathbf{d}}_{\mathbf{x}}^k &= \theta_k^{-1} \mathbf{v}_\mu^k - \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right)^{-1} \nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \mathbf{v}_\mu^k - \\ &\quad - \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \left(\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T \right)^{-1} \mathbf{g}(\mathbf{x}^k) \end{aligned}$$

Reajustando os termos algébricos, temos:

$$\tilde{\mathbf{d}}_{\mathbf{x}}^{\mathbf{k}} = \left(I_n - \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}})^T \left(\nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}})^T \right)^{-1} \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}) \right) \theta_k^{-1} \mathbf{v}_{\mu}^{\mathbf{k}} + \\ + \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}})^T \left(\nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}})^T \right)^{-1} (-\mathbf{g}(\mathbf{x}^{\mathbf{k}})).$$

De acordo com a Proposição (A.2.3) e com a Definição (A.1.6), identificamos o operador linear que nos permite projetar obliquamente o $\theta_k^{-1} \mathbf{v}_{\mu}^{\mathbf{k}}$ em $(\mathbf{w}^{\mathbf{k}})^{\perp} \in \ker(\nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}))$. Este operador é explicitado em (5.7.4) a seguir:

$$P_{\ker(\nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}))_{\theta_k}} = I_n - \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}})^T \left(\nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}})^T \right)^{-1} \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}). \quad (5.7.4)$$

Assim, explicitamos o vetor de direção de busca primal $\tilde{\mathbf{d}}_{\mathbf{x}}^{\mathbf{k}}$ em (5.7.5):

$$\tilde{\mathbf{d}}_{\mathbf{x}}^{\mathbf{k}} = P_{\ker(\nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}))_{\theta_k}} \theta_k^{-1} \mathbf{v}_{\mu}^{\mathbf{k}} + \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}})^T \left(\nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}})^T \right)^{-1} (-\mathbf{g}(\mathbf{x}^{\mathbf{k}})). \quad (5.7.5)$$

Uma interessante observação é que $\mathbf{u}^* = \tilde{\mathbf{d}}_{\mathbf{x}}^{\mathbf{k}}$ é a solução ótima do problema de programação quadrática (5.7.6), o qual é interpretado como um problema de projeção de um vetor em um subespaço do $\mathbb{R}^n$.

$$\begin{aligned} \text{Min} \quad & \frac{1}{2} \phi_{\theta_k}(\theta_k^{-1} \mathbf{v}_{\mu}^{\mathbf{k}} - \mathbf{u}) = \frac{1}{2} (\theta_k^{-1} \mathbf{v}_{\mu}^{\mathbf{k}} - \mathbf{u})^T \theta_k (\theta_k^{-1} \mathbf{v}_{\mu}^{\mathbf{k}} - \mathbf{u}) \\ \text{s.a :} \quad & \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}) \mathbf{u} = -\mathbf{g}(\mathbf{x}^{\mathbf{k}}) \end{aligned} \quad (5.7.6)$$

De acordo com o problema (A.3.4), página 215, identificamos que $P_{\ker(\nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}))_{\theta_k}} \theta_k^{-1} \mathbf{v}_{\mu}^{\mathbf{k}}$ é a componente do vetor $\tilde{\mathbf{d}}_{\mathbf{x}}^{\mathbf{k}}$ que nos assegura que esta direção de busca primal é uma direção viável, isto é, nos assegura que $\nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}) \tilde{\mathbf{d}}_{\mathbf{x}}^{\mathbf{k}} = \mathbf{0}$.

Entretanto, a componente

$$\mathbf{q}^{\mathbf{k}} = \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}})^T \left(\nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}})^T \right)^{-1} (-\mathbf{g}(\mathbf{x}^{\mathbf{k}})) \quad (5.7.7)$$

não é nula em geral, pois é pressuposto que $\mathbf{x}^{\mathbf{k}}$ é uma *boa* estimativa primal para resolvermos o sistema (5.4.10). Nossa proposta é apresentarmos uma interpretação geométrica a respeito desta componente. Para isto, consideremos o problema (A.3.13), página 218.

É imediato constatar que se o termo independente associado à restrição de igualdade do problema (A.3.13) é $-\mathbf{g}(\mathbf{x}^{\mathbf{k}})$, então, de acordo com a expressão (A.3.14), a componente $\mathbf{q}^{\mathbf{k}}$ é o vetor que satisfaz a restrição $\nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}) \mathbf{q}^{\mathbf{k}} = -\mathbf{g}(\mathbf{x}^{\mathbf{k}})$, cujo quadrado da magnitude dado por:

$$\phi_{\theta_k}(\mathbf{q}^{\mathbf{k}}) = (\mathbf{q}^{\mathbf{k}})^T \theta_k \mathbf{q}^{\mathbf{k}}$$

é a menor possível. Portanto, a interpretação geométrica do vetor $\mathbf{q}^{\mathbf{k}}$ é que este vetor é o mais próximo de satisfazer $\nabla \mathbf{g}(\mathbf{x}^{\mathbf{k}}) \tilde{\mathbf{d}}_{\mathbf{x}}^{\mathbf{k}} = \mathbf{0}$. Logo, $\mathbf{q}^{\mathbf{k}}$ é denominado “vetor de desvio quadrático mínimo” do método PDRNLI.

5.8 O método PDRNLI: A estratégia de convergência global denominada condição de positividade

A matriz hessiana do método PDRNLI é a matriz $\frac{\partial L(\omega^k)}{\partial \mathbf{x} \partial \mathbf{x}} = \theta_k$, expressão (5.5.26). Mostramos que essa matriz (e sua inversa) está no cerne de muitas transformações lineares pelas quais nos permitem afirmar que o método PDRNLI determina, da melhor forma possível, um vetor de direção de busca $\tilde{\mathbf{d}}_x^k$ tal que $\mathbf{x}^{k+1}$ torna-se ou permanece viável.

As matrizes das transformações lineares são: $P_{\ker(\nabla \mathbf{g}(\mathbf{x}^k))_{\theta_k}}$, $P_{\text{Im}(\nabla \mathbf{g}(\mathbf{x}^k)^T)_{\theta_k^{-1}}}$ e a pseudo-inversa de $\nabla \mathbf{g}(\mathbf{x}^k)$ dada por $\theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T (\nabla \mathbf{g}(\mathbf{x}^k) \theta_k^{-1} \nabla \mathbf{g}(\mathbf{x}^k)^T)^{-1}$. A interpretação geométrica dessas transformações é relatada por meio dos problemas de otimização apresentados em (A.3.4), (A.3.8) e (A.3.13). O que há de comum em todos esses problemas é que a respectiva função objetivo é uma forma quadrática cuja interpretação é o quadrado da magnitude de um vetor.

Nesta seção, apresentamos a proposta do algoritmo denominado “Condição de Positividade”, a qual consiste em uma sequência de perturbações múltiplas e positivas da matriz identidade adicionadas à matriz θ_k até que o quadrado da magnitude do vetor $\tilde{\mathbf{d}}_x^{k-1}$ mediante a forma quadrática isotrópica regular $\phi_{\theta_k}(\tilde{\mathbf{d}}_x^{k-1}) = (\tilde{\mathbf{d}}_x^{k-1})^T \theta_k \tilde{\mathbf{d}}_x^{k-1}$ seja um número positivo.

Em nosso algoritmo, não é necessário aplicarmos o procedimento da decomposição de Cholesky [Franco 2007, p. 141] sobre a matriz θ_k para predizermos se uma direção de busca primal será de descida. O motivo é que não nos interessa saber se θ_k é definida positiva em todo $\mathbb{R}^n$ como é feito no algoritmo de estratégia de convergência global em [Pinheiro et al. 2015], mas se esta é definida positiva na vizinhança de um ponto $\mathbf{x}^k$.

No Capítulo 4, nós apresentamos na Definição 4.1.6 a condição S-estacionaridade de primeira ordem associado ao MPEC (4.1.1). Além disso, se o respectivo MPEC é formulado como um problema de programação não-linear, então, de acordo com o Teorema 4.1.1 e com as discussões feitas sobre o sistema não-linear de equações (4.2.10), temos que as condições de KKT de primeira ordem associadas com este problema coincidem com a S-estacionaridade.

Em [Guo, Lin e Ye 2012], os autores investigaram os três tipos de condições de KKT de segunda ordem para problemas de programação não-linear e também estudaram as condições S-/M-/C-estacionaridade de segunda ordem para o MPEC (4.1.1). Logo, por exemplo, é possível afirmarmos que um ponto $\mathbf{x}^*$ S-estacionário é, de fato, um minimizador local.

No Teorema 3.1 em [Guo, Lin e Ye 2012], os autores apresentam uma conexão entre as condições S-estacionaridade e de KKT de segunda ordem. Além disso, eles estabelecem no Corolário 3.1 deste teorema que se as usuais condições suficientes (necessárias) de S-estacionaridade de segunda ordem estão satisfeitas, então as condições suficientes (necessárias) KKT estão satisfeitas. Todavia, a recíproca da proposição não é verdadeira. Portanto, nesta tese, nos referimos apenas às condições necessárias (suficientes) de KKT de segunda ordem para assegurarmos pelo menos as condições necessárias de segunda ordem para a S-estacionaridade.

Iniciamos nosso estudo a partir do conceito de condições suficientes de KKT de segunda ordem.

Para este fim, consideremos o problema de programação não-linear (5.8.1).

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) \\ \text{s.a :} \quad & \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ & \mathbf{h}(\mathbf{x}) \leq \mathbf{0}, \end{aligned} \quad (5.8.1)$$

em que a função objetivo e os funcionais vinculados às restrições de igualdade e de desigualdade são pelo menos de classe C^2 em $\mathbb{R}^n$. A função lagrangiana associada à (5.8.1) é dada por: $L(\mathbf{x}, \lambda_0, \lambda_1) = f(\mathbf{x}) + \lambda_0^T \mathbf{g}(\mathbf{x}) + \lambda_1^T \mathbf{h}(\mathbf{x})$ e a matriz hessiana associada à variável $\mathbf{x}$ é dada por: $\nabla_{\mathbf{xx}}^2 L(\mathbf{x}, \lambda_0, \lambda_1) = \nabla^2 f(\mathbf{x}) + \sum_{t=1}^m \lambda_{0t} \nabla^2 g_t(\mathbf{x}) + \sum_{i=1}^r \lambda_{1i} \nabla^2 h_i(\mathbf{x})$.

De acordo com o Teorema 4.4.2 (condições suficientes de KKT de segunda ordem) em [Bazaraa, Sherali e Shetty 2006, p. 213], ao considerarmos: $\mathbf{x}^*$ um ponto de KKT de primeira ordem; os multiplicadores de Lagrange ótimos λ_0^* e λ_1^* , os quais estão respectivamente vinculados às restrições $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ e $\mathbf{h}(\mathbf{x}) \leq \mathbf{0}$ e; o cone das direções viáveis de primeira ordem em $\mathbf{x}^*$ dado por:

$$\mathcal{L}(\mathbf{x}^*) = \left\{ \mathbf{0} \neq \mathbf{u} \in \mathbb{R}^n; \begin{aligned} & \nabla \mathbf{g}(\mathbf{x}^*) \mathbf{u} = \mathbf{0}, \\ & \nabla h_i(\mathbf{x}^*)^T \mathbf{u} = 0, \forall i \in \mathbb{I}(\mathbf{x}^*) \text{ e } \lambda_{1_i}^* > 0 \\ & \nabla h_i(\mathbf{x}^*)^T \mathbf{u} \leq 0, \forall i \in \mathbb{I}(\mathbf{x}^*) \text{ e } \lambda_{1_i}^* = 0 \end{aligned} \right\},$$

então $\mathbf{u}^T \nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*) \mathbf{u} > 0$ para **todo** $\mathbf{u} \in \mathcal{L}(\mathbf{x}^*)$ nos assegura que $\mathbf{x}^*$ é um minimizador local estrito para o problema (5.8.1). Em particular, está claro que se $\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)$ é não-singular e definida positiva, então temos que $\phi_{\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)}(\mathbf{u}) = \mathbf{u}^T \nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*) \mathbf{u} > 0$ para todo $\mathbf{u} \in \mathcal{L}(\mathbf{x}^*)$.

Consideramos importante constataremos que não é necessário que $\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)$ seja definida positiva em todo o $\mathbb{R}^n$ para assegurarmos que $\phi_{\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)}(\mathbf{u}) > 0$. De fato, consideremos o seguinte problema não-convexo proposto:

$$\begin{aligned} \text{Min} \quad & -x_1 x_2 \\ \text{s.a :} \quad & x_1 + x_2^2 - 8 = 0. \end{aligned} \quad (5.8.2)$$

A função lagrangiana vinculada a este problema é dada por:

$$L(x_1, x_2, \lambda) = -x_1 x_2 + \lambda (x_1 + x_2^2 - 8). \quad (5.8.3)$$

Os dois pontos que satisfazem as condições de KKT de primeira ordem são $\left(\frac{16}{3}, \sqrt{\frac{8}{3}}, \sqrt{\frac{8}{3}}\right)^T$ e $\left(\frac{16}{3}, -\sqrt{\frac{8}{3}}, -\sqrt{\frac{8}{3}}\right)^T$. Vamos mostrar que $\mathbf{x}^* = \left(\frac{16}{3}, \sqrt{\frac{8}{3}}\right)^T$ satisfaz as condições suficientes de segunda ordem. A função primal-lagrangiana é $L(x_1, x_2, \lambda^*) = -x_1 x_2 + \sqrt{\frac{8}{3}}(x_1 + x_2^2 - 8)$. Logo, $\nabla_{\mathbf{xx}} L(x_1^*, x_2^*, \lambda^*) = \begin{pmatrix} 0 & -1 \\ -1 & 2\sqrt{\frac{8}{3}} \end{pmatrix}$. Seja $\nabla g(x_1^*, x_2^*) = \begin{pmatrix} 1 & 2\sqrt{\frac{8}{3}} \end{pmatrix}^T$, então o cone das direções viáveis de primeira ordem em $\mathbf{x}^*$ é dado por $\mathcal{L}(\mathbf{x}^*) = \left\{ \mathbf{u} \in \mathbb{R}^2; \mathbf{u} = \kappa \begin{pmatrix} -2\sqrt{\frac{8}{3}} & 1 \end{pmatrix}^T, \forall \kappa \in \mathbb{R} - \{0\} \right\}$. Por fim, temos $\mathbf{u}^T \nabla_{\mathbf{xx}}^2 L(x_1^*, x_2^*, \lambda^*) \mathbf{u} = 2\sqrt{\frac{8}{3}} \kappa^2 > 0$. Logo, as condições suficientes de KKT de segunda ordem estão satisfeitas em $\mathbf{x}^*$ para todo $\mathbf{u} \in \mathcal{L}(\mathbf{x}^*)$.

Neste exemplo, a matriz $\nabla_{\mathbf{xx}}^2 L(x_1^*, x_2^*, \lambda_0^*)$ não é definida positiva em $\mathbb{R}^2$. Com efeito, pode-se mostrar que seus autovalores são $v_1 = \frac{2\sqrt{6}+\sqrt{33}}{3} > 0$ e $v_2 = \frac{2\sqrt{6}-\sqrt{33}}{3} < 0$. Em termos geométricos,

concluimos que a melhor aproximação quadrática da função primal lagrangiana $L(x_1, x_2, \lambda^*)$ em $\mathbf{x}^*$ é uma hipérbole.

Este exemplo mostrou que não é necessário que a matriz $\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)$ seja definida positiva em todo o $\mathbb{R}^n$ para assegurarmos as condições suficientes de KKT de segunda ordem em $\mathbf{x}^*$. Por este motivo e por induzir uma inadequada aproximação quadrática na vizinhança do $\mathbf{x}^*$, métodos do tipo Levenberg-Marquardt, os quais exigem que $\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)$ seja decomponível no sentido de Cholesky, podem ser ineficazes ou exigir um esforço computacional desnecessário em programação não-linear e não-convexa.

Nosso estudo procede com o conceito de um pseudo-espço euclidiano, em que um breve estudo é apresentado no Apêndice B. Sejam a matriz simétrica $\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)$ tal que esta seja não-singular e possui $a_1 \neq 0$ autovalores positivos e $a_2 \neq 0$ autovalores negativos (contando todas as multiplicidades algébricas e $a_1 + a_2 = n$) e os vetores $\mathbf{a}$ e $\mathbf{b}$ do $\mathbb{R}^n$. Então, de acordo com a Definição B.3.4, página 228; ao munirmos o $\mathbb{R}^n$ com a forma bilinear dada por:

$$\langle \mathbf{a}, \mathbf{b} \rangle_{\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)} = \frac{1}{2} \left(\phi_{\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)}(\mathbf{a} + \mathbf{b}) - \phi_{\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)}(\mathbf{a}) - \phi_{\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)}(\mathbf{b}) \right),$$

cuja forma quadrática $\langle \mathbf{a}, \mathbf{a} \rangle_{\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)} = \phi_{\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)}(\mathbf{a})$ é isotrópica regular de assinatura (a_1, a_2) , então a natureza geométrica deste espaço é não-euclidiana.

A forma bilinear $\langle \mathbf{a}, \mathbf{b} \rangle_{\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)}$ é o produto interno entre dois vetores neste espaço. A forma quadrática $\langle \mathbf{a}, \mathbf{a} \rangle_{\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)}$, ao contrário do espaço euclidiano, nos permite encontrar regiões do $\mathbb{R}^n$ em que: $\langle \mathbf{a}, \mathbf{a} \rangle_{\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)}$ é positivo ($\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)$ é definida positiva em $\mathbf{a} \neq \mathbf{0}$), é negativo ($\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)$ é definida negativa em $\mathbf{a} \neq \mathbf{0}$) ou nulo ($\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)$ é isotrópica em $\mathbf{a} \neq \mathbf{0}$).

Em particular, as condições suficientes de KKT de segunda ordem exigem apenas que para todo $\mathbf{u} \in \mathcal{L}(\mathbf{x}^*)$, então $\mathbf{u}$ deve estar em uma região do $\mathbb{R}^n$ para o qual o produto interno $\langle \mathbf{u}, \mathbf{u} \rangle_{\nabla_{\mathbf{xx}}^2 L(\mathbf{x}^*, \lambda_0^*, \lambda_1^*)}$ no pseudo-espço euclidiano é sempre positivo para podermos afirmar que um ponto $\mathbf{x}^*$ de KKT é um minimizador estrito. Este produto interno é o cerne do algoritmo “Condição de Positividade” do método PDRNLI proposto neste trabalho.

A matriz θ_k é rerepresentada a seguir:

$$\theta_k = \nabla^2 f_{\rho^k}(\mathbf{x}^k) + \sum_{t=1}^m \lambda_{0t_k} \nabla g_t(\mathbf{x}^k) + \sum_{i=1}^r \lambda_{1i_k} \nabla h_i(\mathbf{x}^k) + \mu_k^{-1} \nabla \mathbf{h}(\mathbf{x}^k)^T \bar{Z}_{1k} \Lambda_{1k} \nabla \mathbf{h}(\mathbf{x}^k).$$

Além disso, nos convém explicitar a matriz $\nabla^2 f_{\rho^k}(\mathbf{x}^k)$ conforme (5.8.4) a seguir:

$$\nabla^2 f_{\rho^k}(\mathbf{x}^k) = \nabla^2 \bar{f}(\mathbf{x}^k) + M_{\rho^k}, \quad (5.8.4)$$

onde:

$$M_{\rho^k}(\mathbf{x}^k) = \begin{pmatrix} 0 & & & 0 \\ & \text{diag} \left(\rho_{i_k} v''(m_i(x_{3_{i_k}})) \left(\frac{\partial m_i(x_{3_{i_k}})}{\partial x_{n_1+n_2+i}} \right)^2 + \rho_{i_k} v'(m_i(x_{3_{i_k}})) \frac{\partial^2 m_i(x_{3_{i_k}})}{\partial x_{n_1+n_2+i}^2} \right)_{i=1}^{n_2} & & \\ 0 & & & \end{pmatrix}. \quad (5.8.5)$$

A condição $\rho_{i_k} v''(m_i(x_{3_{i_k}})) \left(\frac{\partial m_i(x_{3_{i_k}})}{\partial x_{n_1+n_2+i}} \right)^2 + \rho_{i_k} v'(m_i(x_{3_{i_k}})) \frac{\partial^2 m_i(x_{3_{i_k}})}{\partial x_{n_1+n_2+i}^2} > 0$, para todo $i = 1, \dots, n_2$, não é inserida à formulação (5.3.1), mas, de forma indireta, nós a induzimos por

meio da matriz $M_{\rho^k, l}(\mathbf{x}^k)$, a qual é resultado de uma sequência de perturbações diagonais sobre matriz $M_{\rho^k}(\mathbf{x}^k)$ até todas as componentes diagonais sejam estritamente positivas. Assim, tal restrição sempre será atendida e esta estratégia nos permite gerar uma sequência de pontos $\mathbf{x}_3^k$ tal que $\lim_{k \rightarrow \infty} \left[\sum_{j=1}^{n_2} \rho_{j_k} v \left(m_j \left(x_{3_{j_k}} \right) \right) \right] = 0$, sempre que $\lim_{k \rightarrow \infty} \rho_{j_k} = \infty$. Isso implica que esta estratégia nos permite encontrar ao menos um ponto discreto na região viável.

Com a matriz $M_{\rho^k, l}(\mathbf{x}^k)$ definida positiva em $\mathbb{R}^n$, nós construímos a matriz $\theta_{k, l}$ segundo (5.8.6):

$$\begin{aligned} \theta_{k, l} = \nabla^2 \bar{f}(\mathbf{x}^k) + M_{\rho^k, l}(\mathbf{x}^k) + \sum_{t=1}^m \lambda_{0_{t_k}} \nabla g_t(\mathbf{x}^k) + \sum_{i=1}^r \lambda_{1_{i_k}} \nabla h_i(\mathbf{x}^k) + \\ + \mu_k^{-1} \nabla \mathbf{h}(\mathbf{x}^k)^T \bar{Z}_{1_k} \nabla \mathbf{h}(\mathbf{x}^k) + \beta_{k, l} I_n, \end{aligned} \quad (5.8.6)$$

onde $\beta \in \mathbb{R}_{++}$ é denominado “parâmetro de amortecimento”.

Assumamos que o método PDRNLI está em uma iteração $k > 1$ e, novamente, que o ponto que $\mathbf{x}^k$ é uma *boa* estimativa primal para resolvermos o sistema (5.4.10). Assim, esperamos que a direção de busca primal $\tilde{\mathbf{d}}_{\mathbf{x}}^{k-1}$ é tal que a distância $\|\tilde{\mathbf{d}}_{\mathbf{x}}^k - \tilde{\mathbf{d}}_{\mathbf{x}}^{k-1}\|$ deva ser um número suficientemente pequeno, mesmo que a direção $\tilde{\mathbf{d}}_{\mathbf{x}}^k$ seja desconhecida *a priori*. Nessas condições, podemos predizer que se o produto interno $\langle \tilde{\mathbf{d}}_{\mathbf{x}}^{k-1}, \tilde{\mathbf{d}}_{\mathbf{x}}^{k-1} \rangle_{\theta_{k, l}} = \phi_{\theta_{k, l}}(\tilde{\mathbf{d}}_{\mathbf{x}}^{k-1})$ em um pseudo-espço euclidiano é positivo, então $\tilde{\mathbf{d}}_{\mathbf{x}}^k$ deverá ser uma direção de busca primal de descida.

Logo, se para algum k obtivermos que a sequência primal-dual $\left\{ \left(\mathbf{x}^{k^T}, \mathbf{z}_1^{k^T}, \lambda_0^{k^T}, \lambda_1^{k^T} \right)^T \right\}$ é de acumulação em torno da solução ótima primal-dual, temos que $\tilde{\mathbf{d}}_{\mathbf{x}}^{k-1}$ está no cone $\mathcal{L}(\mathbf{x}^*)$ e que satisfaz $\phi_{\theta_{k, l}}(\tilde{\mathbf{d}}_{\mathbf{x}}^{k-1}) > 0$, portanto $\mathbf{x}^k$ é um minimizador para o problema (5.2.1).

Por fim, para o caso particular em que $k = 0$, definimos um vetor $\mathbf{x}^0$ e assumimos que $\tilde{\mathbf{d}}_{\mathbf{x}}^{k-1}$ é definido por:

$$\tilde{\mathbf{d}}_{\mathbf{x}}^{k-1} = 0,0005\mathbf{x}^0. \quad (5.8.7)$$

O Algoritmo 5.8.1 descreve os procedimentos computacionais da Condição de Positividade.

Algoritmo 5.8.1 : Condição de Positividade.

Procedimento Inicial:

Em uma iteração k do método PDRNLI, defina um contador $l = 0$, faça $\beta_{k,l} = \beta_k$ e $M_{\rho^k,l} = M_{\rho^k} + \beta_{k,l}I_n$ e siga para o procedimento intermediário.

Procedimento Intermediário:

Enquanto $M_{\rho^k,l}$ não for definida positiva no $\mathbb{R}^n$, faça:

Passo 1.1: $\beta_{k,l+1} = \alpha_\beta \beta_{k,l}$, $\alpha_\beta > 1$;

Passo 1.2: $M_{\rho^k,l+1} = M_{\rho^k,l} + \beta_{k,l+1}I_n$;

Passo 1.3: $l = l + 1$.

fim

Seja $M_{\rho^k,l}$ definida positiva. Construa $\theta_{k,l}$ conforme (5.8.6) e vá para o procedimento principal.

Procedimento Principal

Enquanto $\phi_{\theta_{k,l}}(\tilde{\mathbf{d}}_x^{k-1}) = (\tilde{\mathbf{d}}_x^{k-1})^T \theta_{k,l} \tilde{\mathbf{d}}_x^{k-1} \leq 0$, faça:

Passo 2.1: $\beta_{k,l+1} = \alpha_\beta \beta_{k,l}$, $\alpha_\beta > 1$;

Passo 2.2: $\theta_{k,l+1} = \theta_{k,l} + \beta_{k,l+1}I_n$;

Passo 2.3: $l = l + 1$.

fim

Faça $\theta_k = \theta_{k,l}$ e $\beta_k = \beta_{k,l}$.

A escolha para o fator α_β é de acordo com a heurística apresentada em [Pinheiro et al. 2015] e é conforme (5.8.8):

$$\alpha_\beta = \frac{1 + \sqrt{(\sqrt{5} - 1)^2 \alpha_\mu^2 + 1}}{2}, \quad (5.8.8)$$

onde α_μ é um número conhecido e que está no intervalo $(0, 1)$. Após o cálculo do vetor $\mathbf{b}^{k+1}$, o fator β_{k+1} pode ser aumentado, reduzido ou não ser passível de modificação de acordo com as seguintes condições:

1. Se $L_\alpha(\omega^k) - L_\alpha(\omega^{k+1}) < 0,25$, então a redução do valor de β_{k+1} é conforme (5.8.9):

$$\beta_{k+1} = \frac{4\beta_k}{(\sqrt{5} + 1) + \sqrt{16\alpha_\mu^2 + (\sqrt{5} + 1)^2}}; \quad (5.8.9)$$

2. Se $L_\alpha(\omega^k) - L_\alpha(\omega^{k+1}) > 0,75$, então:

$$\beta_{k+1} = \alpha_\beta \beta_k; \quad (5.8.10)$$

3. Se $0,25 \leq L_\alpha(\omega^k) - L_\alpha(\omega^{k+1}) \leq 0,75$, então:

$$\beta_{k+1} = \beta_k. \quad (5.8.11)$$

5.9 O método PDRNLI: Cálculo do tamanho do passo e atualização dos parâmetros

O cálculo do tamanho do passo é de acordo com [Granville 1994, Pinheiro et al. 2015]. Uma vez calculado o vetor $\mathbf{d}_b^k = (\mathbf{d}_x^k, \mathbf{d}_{z_1}^k, \mathbf{d}_{\lambda_0}^k, \mathbf{d}_{\lambda_1}^k)^T$ por meio do procedimento corretor, então o cálculo do tamanho do passo primal $\bar{\alpha}_p$ e dual $\bar{\alpha}_d$ é conforme (5.9.1):

$$\bar{\alpha}_p = \min \left\{ 1, \min_{-z_{1_i}^k \leq 0 \text{ e } \tilde{d}_{z_{1_i}}^k < 0} \left\{ \frac{-z_{1_i}^k}{\tilde{d}_{z_{1_i}}^k} \right\} \right\} \quad \text{e} \quad \bar{\alpha}_d = \min \left\{ 1, \min_{\lambda_{1_i}^k \geq 0 \text{ e } \tilde{d}_{\lambda_{1_i}}^k < 0} \left\{ \frac{-\lambda_{1_i}^k}{\tilde{d}_{\lambda_{1_i}}^k} \right\} \right\}. \quad (5.9.1)$$

Neste trabalho, o novo ponto $\mathbf{b}^{k+1}$ é obtido por meio de (5.9.2):

$$\mathbf{b}^{k+1} = \begin{pmatrix} \mathbf{x}^k \\ \mathbf{z}_1^k \\ \lambda_0^k \\ \lambda_1^k \end{pmatrix} + \begin{pmatrix} \bar{\alpha}_p \tilde{\mathbf{d}}_x^k \\ \bar{\alpha}_p \tilde{\mathbf{d}}_{z_1}^k \\ \tilde{\mathbf{d}}_{\lambda_0}^k \\ 0,9995 \bar{\alpha}_d \tilde{\mathbf{d}}_{\lambda_1}^k \end{pmatrix} \quad (5.9.2)$$

Para a atualização do parâmetro dual, utilizamos a expressão (5.9.3) a seguir [Pinheiro et al. 2015]:

$$\delta_1^{k+1} = \lambda_1^{k+1}, \quad (5.9.3)$$

isto é, cada componente δ_{1_i} de δ_1 é substituído pelo respectiva componente λ_{1_i} de λ_1 .

A atualização do parâmetro de reescala procede de acordo com (5.9.4).

$$\mu_{k+1} = \alpha_\mu \mu_k, \quad (5.9.4)$$

para algum $\alpha_\mu \in (0, 1)$ previamente estabelecido. Após a atualização, efetuamos o seguinte teste heurístico: Se $z_{1_{i_{k+1}}} > \frac{-\mu_{k+1} \gamma(\delta_{1_{i_{k+1}}}, \alpha)}{\delta_{1_{i_{k+1}}}}$, então (5.9.4) permanece. Caso contrário, μ^{k+1} é atualizado de acordo com (5.9.5).

$$\mu_{k+1} = -(1 - \alpha_\mu) \min \left\{ z_{1_1}^{k+1}, \dots, z_{1_i}^{k+1}, \dots, z_{1_r}^{k+1} \right\}. \quad (5.9.5)$$

Obtemos o vetor ρ^{k+1} conforme a estratégia de controle de viabilidade descrita em [Martínez 2009, seção 2.2]. Para isto, consideramos para todo $i = 1, \dots, n_2$ um número $\alpha_{\rho_i} > 1$ e um número $\tau \in [0, 1)$. Então:

$$\rho_{i_{k+1}} = \rho_{i_k}, \quad (5.9.6)$$

se $|m_i(x_{3_{i_{k+1}}})| \leq \tau |m_i(x_{3_{i_k}})|$. Caso contrário, faça:

$$\rho_{i_{k+1}} = \alpha_{\rho_i} \rho_{i_k}. \quad (5.9.7)$$

5.10 Inicialização do método PDRNLI

Para a inicialização do método PDRNLI, entramos com os valores de μ_0 , $\mathbf{x}^0$ e δ_1^0 e escolhemos um $\alpha \in [0, 1]$ e uma função $\psi \in \mathbb{H}$. Para o cálculo das demais variáveis, utilizamos o sistema (5.4.3).

Para o cálculo das componentes do vetor $\mathbf{z}_1^0$, utilizamos (5.10.1):

$$z_{1_i}^0 = \max \{ -h_i(\mathbf{x}^0); 0, 1 \} \quad , \quad \forall i = 1, \dots, r. \quad (5.10.1)$$

Calculamos λ_1^0 por meio de (5.10.2):

$$\lambda_1^0 = \dot{\Psi} \left(-\mu_0^{-1} \Delta_{\alpha_0}^{-1} \Delta_{1_0} \mathbf{z}_1^0 \right) \delta_1^0 \quad (5.10.2)$$

Vamos supor que $\text{posto}(\nabla \mathbf{g}(\mathbf{x}^0))$ seja m , logo, o procedimento para cálculo de λ_0^0 é feito da seguinte forma:

$$\begin{aligned} \nabla f(\mathbf{x}^0) + \nabla \mathbf{g}(\mathbf{x}^0)^T \lambda_0^0 + \nabla \mathbf{h}(\mathbf{x}^0)^T \lambda_1^0 &= \mathbf{0} \Rightarrow \\ \nabla \mathbf{g}(\mathbf{x}^0) \nabla \mathbf{g}(\mathbf{x}^0)^T \lambda_0^0 &= -\nabla \mathbf{g}(\mathbf{x}^0) \left(\nabla f(\mathbf{x}^0) + \nabla \mathbf{h}(\mathbf{x}^0)^T \lambda_1^0 \right). \end{aligned}$$

Como $\text{posto}(\nabla \mathbf{g}(\mathbf{x}^0)) = m$, então a matriz simétrica e definida positiva $\nabla \mathbf{g}(\mathbf{x}^0) \nabla \mathbf{g}(\mathbf{x}^0)^T \in \mathbb{R}^{m \times m}$ possui inversa. Assim, temos:

$$\lambda_0^0 = \left(\nabla \mathbf{g}(\mathbf{x}^0) \nabla \mathbf{g}(\mathbf{x}^0)^T \right)^{-1} \nabla \mathbf{g}(\mathbf{x}^0) \left(-\nabla f(\mathbf{x}^0) - \nabla \mathbf{h}(\mathbf{x}^0)^T \lambda_1^0 \right). \quad (5.10.3)$$

5.11 Algoritmo computacional do método PDRNLI

O detalhamento do método PDRNLI é descrito no Algoritmo 5.11.1.

Algoritmo 5.11.1 Método predictor-corretor primal-dual de reescalamiento não-linear integrado

Procedimento Inicial:

Defina uma função $\psi \in \mathbb{H}$ e uma função $v \in \mathbb{G}$ e um $\alpha \in [0, 1]$. Defina $\varepsilon > 0$ suficientemente pequeno e um contador $k = 0$. Defina os parâmetros $\mu_0 > 0$, $\rho^0 \in \mathbb{R}_{++}^{n_2}$, $\delta_1^0 \in \mathbb{R}_{++}^r$, $\beta_0 > 0$ e $\tau \in [0, 1)$. Defina os pesos $\alpha_\mu \in (0, 1)$ e $\alpha_{\rho_i} > 1$. Defina um vetor $\mathbf{x}^0$ e determine $\tilde{\mathbf{d}}_{\mathbf{x}}^{k-1}$ de acordo com (5.8.7) e os vetores $\mathbf{z}_1^0$, λ_0^0 e λ_1^0 conforme (5.10.1), (5.10.2) e (5.10.3) respectivamente. Calcule o vetor de residual $\mathbf{a}^0$ em (5.5.20) e siga para o procedimento principal.

Procedimento Principal:

Enquanto $\|\mathbf{a}^k\|_\infty > \varepsilon$, faça:

Procedimento Predictor

Passo 1.1: Aplique a condição de positividade de acordo com Algoritmo 5.8.1.

Passo 1.2: Determine as direções de busca primal-dual $\mathbf{d}_b^k$ conforme apresentado em (5.6.1).

Procedimento Corretor

Passo 2.1: Obtenha as direções de busca primal-dual $\tilde{\mathbf{d}}_b^k$ através de (5.6.3).

Cálculo do tamanho dos passos primal e dual e atualização da solução

Passo 3.1: Calcule o tamanho dos passos primal $\bar{\alpha}_p$ e dual $\bar{\alpha}_d$ de acordo com (5.9.1).

Passo 3.2: Utilize (5.9.2) para obter o vetor $\mathbf{b}^{k+1}$.

Atualização dos parâmetros

Passo 4.1: Verifique as condições (5.8.9) – (5.8.11) para obter β_{k+1} .

Passo 4.2: Obtenha o parâmetro dual δ_1^{k+1} por meio de (5.9.3).

Passo 4.3: Determine o parâmetro de reescala μ^{k+1} utilizando (5.9.4) e use (5.9.5) se necessário.

Passo 4.4: Use a estratégia de controle de viabilidade (5.9.6) e (5.9.7) para obter novo parâmetro de penalidade ρ^{k+1} .

Passo 4.5: Faça $k = k + 1$.

Cálculo do vetor residual

Passo 5.1: Calcule $\mathbf{a}^k$ utilizando (5.5.20).

5.12 Aplicação do método PDRNLI em exemplos matemáticos

Nesta seção são apresentados os resultados da implementação computacional do método PDRNLI apresentado no Capítulo 5 em exemplos matemáticos de programação não-linear, MPEC e de programação discreta. A implementação do Algoritmo 5.11.1 foi desenvolvida em MATLAB®R2012b.

5.12.1 Programação não-linear convexo

Nesta seção, consideramos o seguinte problema de otimização convexo (5.12.1) [Lage 2013, p. 82].

$$\begin{aligned}
 \text{Min} \quad & (x_1 - 2)^4 + (x_1 - 2x_2)^2 \\
 \text{s.a :} \quad & x_1 + x_2 - 3 = 0 \\
 & x_1^2 - x_2 \leq 0 \\
 & -x_2 + 1,5 \leq 0 \\
 & x_2 - 2 \leq 0.
 \end{aligned} \tag{5.12.1}$$

As curvas de nível da função $f(\mathbf{x})$ são apresentadas na Figura 5.12.1.

Figura 5.12.1: Curvas de nível da função objetivo do problema (5.12.1).

A região viável é apresentada na Figura 5.12.2.

Figura 5.12.2: Conjunto de soluções viáveis do problema (5.12.1).

A Tabela 5.12.1 apresenta as informações sobre os parâmetros iniciais, o ponto inicial (inviável), a precisão ε e a função ψ para tratarmos com as restrições de desigualdade. Esclarecemos que o problema (5.12.1), assim como qualquer problema de programação não-linear padrão, é compatível com a formulação (5.3.2) desde que o parâmetro de penalidade ρ^k seja sempre nulo para todo k . Além disso, tornar $\rho^k = 0$ para todo k implica que a matriz $M_{\rho^k}(\mathbf{x}^k)$, expressão (5.8.5), é nula em todas as iterações. Este resultado afeta em nada a ideia central do algoritmo da condição de positividade, pois é evidente que qualquer perturbação diagonal sobre a matriz nula $M_{\rho^k}(\mathbf{x}^k)$ a tornará definida positiva.

Tabela 5.12.1: Dados de entrada para resolvermos o problema (5.12.1).

$\mathbf{x}^0$	$(-1, 5; -0, 5)^T$
μ_0	0, 01
α_μ	0, 2
$\delta_{1_i}^0, i = 1, 2, 3.$	0, 1
$\gamma(\delta_{1_i}^0, \alpha), i = 1, 2, 3.$	$(\delta_{1_i}^0)^\alpha, \alpha \in [0, 1].$
β_0	0, 1
$\psi(t)$	$\frac{1}{2}(\arg \sinh(t) + t^2 + t\sqrt{t^2 + 1})$
ε	10^{-7}

Para qualquer $\alpha \in [0, 1]$ dado, temos que a solução ótima primal $(\mathbf{x}^{*T}, \mathbf{z}_1^{*T})^T$ determinada pelo método PDRNLI para este problema é dada por:

$$\begin{pmatrix} x_1^* \\ x_2^* \\ z_{1_1}^* \\ z_{1_2}^* \\ z_{1_3}^* \end{pmatrix} = \begin{pmatrix} 1,3027756 \\ 1,6972244 \\ 0,0000000 \\ 0,1972244 \\ 0,3027756 \end{pmatrix}.$$

A solução dual $(\lambda_0^*, \lambda_1^{*T})^T$ ótima é dada por:

$$\begin{pmatrix} \lambda_0^* \\ \lambda_{1_1}^* \\ \lambda_{1_2}^* \\ \lambda_{1_3}^* \end{pmatrix} = \begin{pmatrix} -4,5099222 \\ 3,8567702 \\ 0,0000000 \\ 0,0000000 \end{pmatrix}.$$

A Tabela 5.12.2 sintetiza a sequência numérica dos principais resultados obtidos pelo método PDRNLI ao impormos $\alpha = 0$: função objetivo, função lagrangiana aumentada, quão satisfeita está as condições de KKT, parâmetro de reescala, parâmetro de amortecimento, tamanho dos passos primal e dual.

Tabela 5.12.2: Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_1$ de funções penalidades aplicado ao problema (5.12.1).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	150,3125	-276,2802502	8,72E+01	0,1	0,1	-	-
1	26,36786554	15675,13407	1,77E+01	2,00E-03	0,1	0,83102	1
2	29,19134968	30,04696536	6,84E+03	7,09E-01	6,08E-02	1	1
3	11,49455348	10,22047684	1,62E+02	1,41E-01	6,16E-02	1	1
4	3,963171769	4,362455837	2,07E+01	2,83E-02	6,23E-02	0,98454	1
5	4,594117091	4,558628245	1,11E-01	5,67E-03	6,31E-02	1	1
6	4,626147742	4,607523409	1,49E-02	1,13E-03	3,84E-02	1	1
7	4,611931557	4,609135906	1,09E-01	2,26E-04	2,34E-02	1	1
8	4,611482085	4,610946012	3,35E-03	4,53E-05	1,42E-02	1	1
9	4,611422124	4,611317091	4,26E-04	9,07E-06	8,67E-03	1	1
10	4,611412639	4,611391919	6,52E-05	1,81E-06	5,28E-03	1	1
11	4,611411047	4,611406948	1,06E-05	3,63E-07	3,21E-03	1	1
12	4,611410774	4,611409962	1,76E-06	7,26E-08	1,95E-03	1	1
13	4,611410727	4,611410566	2,91E-07	1,45E-08	1,19E-03	1	1
14	4,611410719	4,611410687	6,50E-08	2,90E-09	7,26E-04	1	1

A Figura 5.12.3 apresenta a sequência de soluções $\mathbf{x}^k$ sobre o plano $\mathbb{R}^2$ desenvolvida pelo método PDRNLI juntamente com as curvas de nível da função objetivo e com a região viável do problema (5.12.1).

Figura 5.12.3: A trajetória dos pontos $\mathbf{x}^k$ gerada pelo método PDRNLI para a classe $\mathbb{P}_1$ de funções penalidades para resolver o problema (5.12.1).

Com precisão de 7 casas decimais, a Tabela 5.12.3 mostra as sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ geradas pelo método PDRNLI aplicado ao problema (5.12.1).

Tabela 5.12.3: As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_1$ de funções penalidades aplicado ao problema (5.12.1).

Ciclo k	x_{1k}	x_{2k}	λ_{0k}	λ_{11k}	λ_{12k}	λ_{13k}
0	-1,5	-0,5	85,3525538	0,0414214	0,0414214	0,0019992
1	0,1551081	2	35,2659651	49,2598076	4,3888653	0,0085223
2	0,4089455	2,5910545	-4,8718926	13,3778942	0,9978005	0,0992545
3	0,9349974	2,0650026	-9,6945895	2,5825929	0,5764236	0,100384
4	1,3513952	1,6486048	-4,4963966	2,7877399	0,4771169	0,0828848
5	1,3040206	1,6959794	-4,4902621	3,7970026	0,1049137	0,0377752
6	1,3017168	1,6982832	-4,5097403	3,8645173	0,0141226	0,0088963
7	1,3027382	1,6972618	-4,5093401	3,856813	0,0027721	0,0018013
8	1,3027705	1,6972295	-4,5097995	3,8567613	0,0005536	0,0003606
9	1,3027748	1,6972252	-4,509897	3,8567666	0,0001107	0,0000721
10	1,3027755	1,6972245	-4,5099171	3,8567692	0,0000221	0,0000144
11	1,3027756	1,6972244	-4,5099212	3,8567699	0,0000044	0,0000029
12	1,3027756	1,6972244	-4,509922	3,8567701	0,0000009	0,0000006
13	1,3027756	1,6972244	-4,5099222	3,8567701	0,0000002	0,0000001
14	1,3027756	1,6972244	-4,5099222	3,8567701	0	0

A Tabela 5.12.4 sintetiza a sequência numérica dos principais resultados obtidos pelo método PDRNLI ao impormos $\alpha = 1$: função objetivo, função lagrangiana aumentada, quão satisfeita está as condições de KKT, parâmetro de reescala, parâmetro de amortização, tamanho dos passos primal e dual.

Tabela 5.12.4: Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades aplicado ao problema (5.12.1).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	150,3125	-275,9815104	8,72E+01	0,1	0,1	-	-
1	50,28614008	9012,424976	8,08E+01	2,00E-03	0,1	0,72717	1
2	9,983587222	11,938329	7,25E+01	4,00E-04	6,08E-02	1	1
3	9,956302912	9,923033778	2,02E+01	8,00E-05	6,16E-02	1	0,72604
4	4,203328989	9,61467182	4,49E+02	1,60E-05	6,23E-02	0,01313	1
5	4,608034953	4,53126067	5,26E+00	3,20E-06	6,23E-02	1	1
6	4,61142966	4,611247515	4,92E-03	6,40E-07	6,31E-02	1	1
7	4,611411799	4,611410902	3,60E-05	1,28E-07	3,84E-02	1	1
8	4,611410869	4,611410738	1,77E-06	2,56E-08	2,34E-02	1	1
9	4,611410742	4,61141072	2,32E-07	5,12E-09	1,42E-02	1	1
10	4,611410722	4,611410718	9,38E-08	1,02E-09	8,67E-03	1	1

A Figura 5.12.4 apresenta a sequência de soluções $\mathbf{x}^k$ sobre o plano $\mathbb{R}^2$ desenvolvida pelo método PDRNLI juntamente com as curvas de nível da função objetivo e com a região viável do problema (5.12.1).

Figura 5.12.4: A trajetória dos pontos $\mathbf{x}^k$ gerada pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades para resolver o problema (5.12.1).

Com precisão de 7 casas decimais, a Tabela 5.12.5 mostra as sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ geradas pelo método PDRNLI aplicado ao problema (5.12.1).

Tabela 5.12.5: As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades aplicado ao problema (5.12.1).

Ciclo k	x_{1k}	x_{2k}	λ_{0k}	λ_{11k}	λ_{12k}	λ_{13k}
0	-1,5	-0,5	85,2623689	0,0049876	0,0049876	0,0002
1	-0,3641682	2	-10,5418447	12,0324853	1,0198647	0,0008848
2	1,0007464	1,9992536	-64,8222843	1,107557	0,4786687	54,3845711
3	1,0019889	1,9980111	9,0177702	0,6459599	0,131953	0,0271923
4	1,3328938	1,6671062	-575,6753121	49,4193881	6,6583649	923,3326228
5	1,3030184	1,6969816	-9,0444083	6,5966186	0,2161555	7,9248281
6	1,3027743	1,6972257	-4,5095759	3,8585363	0,0001112	0,0039839
7	1,3027756	1,6972244	-4,5099225	3,8567716	0,0000001	0,000002
8	1,3027756	1,6972244	-4,5099223	3,8567702	0	0
9	1,3027756	1,6972244	-4,5099222	3,8567701	0	0
10	1,3027756	1,6972244	-4,5099222	3,8567702	0	0

Apresentamos na Tabela 5.12.6 a influência do parâmetro α sobre o desempenho do método PDRNLI proposto quando este parâmetro assume alguns valores que estão no intervalo $[0, 1]$. Para um dado valor para α , esta influência é observada nos seguintes quesitos: número de ciclos; função objetivo no ponto ótimo; função lagrangiana aumentada integrada no ponto ótimo e precisão sobre as condições de KKT. Em todos os casos, as condições iniciais dos parâmetros e das variáveis; o tipo de fator de transformação $\gamma(\delta_{1_i}^0, \alpha)$, $i = 1, \dots, 3$, e a função $\psi(t)$ são aquelas apresentadas na Tabela 5.12.1.

Tabela 5.12.6: Influência do parâmetro α sobre o desempenho do método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades aplicado ao problema (5.12.1).

α	Ciclos k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$
0	14	4,611410719	4,611410687	6,50E-08
0,1	9	4,611410718	4,611365952	3,47E-08
0,2	14	4,611410719	4,611410717	4,67E-08
0,3	14	4,61141072	4,611410718	4,56E-08
0,4	14	4,61141072	4,611410718	4,04E-08
0,5	13	4,611410718	4,611410718	2,21E-08
0,6	11	4,611410718	4,611410717	1,24E-09
0,7	9	4,611410718	4,611410713	1,52E-09
0,8	14	4,611410719	4,611410718	2,92E-08
0,9	14	4,61141072	4,611410718	2,12E-08
1	10	4,611410722	4,611410718	9,39E-08

Apresentamos na Figura 5.12.5 as diferentes trajetórias de pontos $\mathbf{x}^k$ geradas pelo método PDRNLI ao atribuímos os valores de α definidos na Tabela 5.12.6. Nesta figura, a sequência de pontos em amarelo é a mesma apresentada na Figura 5.12.3 ao passo que a sequência de pontos em verde é a mesma apresentada na Figura 5.12.4.

Figura 5.12.5: Trajetória dos pontos $\mathbf{x}^k$ gerados pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades em função da escolha de α .

5.12.2 Programação não-linear não-convexo

Não nos preocupamos em mostrar na Seção 5.12.1 os efeitos do parâmetro de amortecimento β junto à estratégia de convergência global (condição de positividade), pois o problema (5.12.1) é convexo. O próximo exemplo nos mostra que em problemas de programação não-linear não-convexo, a aplicação da estratégia de convergência global pode mudar significativamente a sequência $\{\mathbf{x}^k\}$ até um ponto de KKT de primeira ordem. Para constatarmos isto, seja o problema proposto em (5.12.2) a seguir:

$$\begin{aligned}
 \text{Min} \quad & \sin(x_2^2 - x_1^2) \\
 \text{s.a :} \quad & x_1 + x_2 - 3 = 0 \\
 & x_1^2 - x_2 \leq 0 \\
 & -x_2 + 1,5 \leq 0 \\
 & x_2 - 2 \leq 0.
 \end{aligned} \tag{5.12.2}$$

Resolvemos este problema sob dois aspectos:

Aspecto 1: Ignoramos a aplicação do condição de positividade no Algoritmo 5.11.1 e;

Aspecto 2: Consideramos integralmente o Algoritmo 5.11.1.

A Tabela 5.12.7 apresenta os dados de entrada para resolvermos este problema.

Tabela 5.12.7: Dados de entrada para resolvermos o problema (5.12.2).

$\mathbf{x}^0$	$(0, 5; 1, 8)^T$
μ_0	0, 1
α_μ	0, 2
$\delta_{1_i}^0, i = 1, 2, 3.$	0, 1
$\gamma(\delta_{1_i}^0, \alpha), i = 1, 2, 3.$	$1 + (\delta_{1_i}^0 - 1) \alpha, \alpha \in [0, 1]$
β_0	1
$\psi(t)$	$-\ln(-t + 1)$
ε	10^{-7}

As Tabelas 5.12.8 e 5.12.9 sintetizam respectivamente para esses dois aspectos a sequência numérica dos principais resultados em um ciclo k obtidos pelo método PDRNLI ao impormos $\alpha = 0, 1$: função objetivo, função lagrangiana aumentada, quão satisfeita estão as condições de KKT, parâmetro de reescala, parâmetro de amortecimento (para o segundo aspecto), tamanho dos passos primal e dual.

Tabela 5.12.8: Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_{int}$ sem aplicação da estratégia de convergência global.

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	0,151012712	-0,79591588	2,27E+00	0,1	-	-
1	0,982513221	346,3907336	4,64E+00	2,00E-03	0,18693	1
2	0,996434204	0,980976421	3,53E-01	4,00E-04	1	1
3	0,99999088	0,994489264	4,61E-03	8,00E-05	1	0,99657
4	0,99999999	0,999397468	2,72E-05	1,60E-05	1	1
5	1	0,999937369	3,34E-07	3,20E-06	1	1
6	1	0,999986819	5,28E-08	6,40E-07	1	1

Tabela 5.12.9: Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_{int}$ com a estratégia de convergência global.

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	0,151012712	-0,79591588	2,27E+00	0,1	0,1	-	-
1	-0,659970823	-340,3922304	4,86E+00	2,00E-03	1,37E-01	0,1014	1
2	0,149363313	0,132961892	3,97E-01	4,00E-04	1,38E-01	1	1
3	0,14114614	0,13941048	3,33E-02	8,00E-05	8,44E-02	1	1
4	0,141120045	0,140847925	1,03E-04	1,60E-05	5,14E-02	1	1
5	0,141120009	0,141072275	1,43E-07	3,20E-06	3,13E-02	1	1
6	0,141120008	0,141110675	9,43E-09	6,40E-07	1,90E-02	1	1

As Figuras 5.12.6 e 5.12.7 apresentam respectivamente sob os dois aspectos a sequência $\{\mathbf{x}^k\}$ desenvolvidas pelo método PDRNLI sobre o plano $\mathbb{R}^2$ juntamente com as curvas de nível da função objetivo e com a região viável do problema (5.12.2).

Figura 5.12.6: A trajetória dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ sem a estratégia de convergência global.

Figura 5.12.7: A trajetória dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ com a estratégia de convergência global.

Com precisão de 7 casas decimais, as Tabelas 5.12.10 e 5.12.11 mostram respectivamente para esses dois aspectos as sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ geradas pelo método PDRNLI aplicado ao problema (5.12.2).

Tabela 5.12.10: As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ sem a estratégia de convergência global.

Ciclo k	x_{1_k}	x_{2_k}	λ_{0_k}	$\lambda_{1_{1_k}}$	$\lambda_{1_{2_k}}$	$\lambda_{1_{3_k}}$
0	0,5	1,8	1,2810925	0,0055454	0,0232737	0,0312715
1	0,9308542	1,5	-4,4078535	0,0574423	0,6215676	1,0786193
2	1,2241216	1,7758784	-0,2542924	0,0866009	0,1533175	0,6630815
3	1,2389124	1,7610876	-0,0259797	0,0128912	0,0000767	0,0236775
4	1,2382236	1,7617764	-0,0004184	0,000318	0,00006	0,0003097
5	1,2382051	1,7617949	-0,0000638	0,0000527	0,0000287	0,0000507
6	1,2382015	1,7617985	-0,0000113	0,0000102	0,000008	0,0000098

Tabela 5.12.11: As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ com a estratégia de convergência global.

Ciclo k	x_{1_k}	x_{2_k}	λ_{0_k}	$\lambda_{1_{1_k}}$	$\lambda_{1_{2_k}}$	$\lambda_{1_{3_k}}$
0	0,5	1,8	1,2810925	0,0055454	0,0232737	0,0312715
1	0,3709817	2	4,2914051	0,0257427	0,7075858	2,1877527
2	1,0013886	1,9986114	-1,8584431	0,0028019	0,0087185	6,1203453
3	1,0000044	1,9999956	-1,9809042	0,000486	0,0006692	5,9421058
4	1	2	-1,9801102	0,0000629	0,0001188	5,9402621
5	1	2	-1,9800084	0,0000117	0,0000232	5,9400134
6	1	2	-1,9799896	0,0000023	0,0000046	5,9399666

As Tabelas 5.12.12 e 5.12.13 sintetizam respectivamente a influência do parâmetro α sobre o desempenho do método PDRNLI proposto quando este parâmetro assume alguns valores que estão no intervalo $[0, 1]$ para os casos em que a estratégia de convergência global é ignorada e considerada no Algoritmo 5.11.1.

Para um dado valor para α , esta influência é observada nos seguintes quesitos: número de ciclos; função objetivo no ponto ótimo; função lagrangiana aumentada integrada no ponto ótimo e precisão sobre as condições de KKT. Em todos os casos, as condições iniciais dos parâmetros e das variáveis; o tipo de fator de transformação $\gamma(\delta_{1_i}^0, \alpha)$, $i = 1, \dots, 3$, e a função $\psi(t)$ são aquelas apresentadas na Tabela 5.12.7.

Tabela 5.12.12: Influência do parâmetro α sobre o desempenho do método PDRNLI para a classe $\mathbb{P}_{int}$ sem a estratégia de convergência global.

α	Ciclos k	$f(\mathbf{x}^k)$	$L_\alpha(\mathbf{b}^{k^T}, \sigma^{k^T})$	$\ \mathbf{a}^k\ _\infty$
0	6	0,99999999997952	0,99999999997952	5,81E-08
0,1	6	0,99999999998438	0,99998681912226	5,28E-08
0,2	6	0,99999999998729	0,99998745016709	4,74E-08
0,3	14	0,99999999999361	0,99999509755480	2,54E-08
0,4	10	0,14112000805986	0,14111999757139	5,05E-10
0,5	7	0,14112000806179	0,14111897701947	1,04E-09
0,6	7	0,14112000806162	0,14111918439078	8,25E-10
0,7	7	0,14112000806136	0,14111938813602	6,32E-10
0,8	7	0,14112000806100	0,14111959291484	4,38E-10
0,9	7	0,14112000806052	0,14111980026000	4,54E-10
1	7	0,14112000806024	0,14112000805987	3,77E-09

Tabela 5.12.13: Influência do parâmetro α sobre o desempenho do método PDRNLI para a classe $\mathbb{P}_{int}$ com a estratégia de convergência global.

α	Ciclos k	$f(\mathbf{x}^k)$	$L_\alpha(\mathbf{b}^{k^T}, \sigma^{k^T})$	$\ \mathbf{a}^k\ _\infty$
0	6	0,14112000808793	0,14110963958016	1,05E-08
0,1	6	0,14112000809762	0,14111067451395	9,43E-09
0,2	6	0,14112000810454	0,14111170980794	8,39E-09
0,3	6	0,14112000810870	0,14111274544921	7,35E-09
0,4	6	0,14112000811008	0,14111378142947	6,30E-09
0,5	6	0,14112000810869	0,14111481774623	5,26E-09
0,6	6	0,14112000810452	0,14111585440792	4,21E-09
0,7	6	0,14112000809757	0,14111689145622	3,16E-09
0,8	6	0,14112000808781	0,14111792907653	2,11E-09
0,9	6	0,14112000807490	0,14111897135343	1,31E-09
1	6	0,14112000805987	0,14112000805987	4,01E-09

Apresentamos na Figura 5.12.7 as diferentes trajetórias de pontos $\mathbf{x}^k$ geradas pelo método PDRNL ao atribuirmos os valores de α definidos na Tabela 5.12.12. Nesta figura, a sequência de pontos $\mathbf{x}^k$ em amarelo é obtida ao assumirmos $\alpha = 0$ e a sequência de pontos $\mathbf{x}^k$ em verde é obtida ao assumirmos $\alpha = 1$.

Figura 5.12.8: As trajetórias dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI em função da escolha de α sem a estratégia de convergência global.

Apresentamos na Figura 5.12.9 as diferentes trajetórias de pontos $\mathbf{x}^k$ geradas pelo método PDRNL ao atribuímos os valores de α definidos na Tabela 5.12.13. Nesta figura, a sequência de pontos $\mathbf{x}^k$ em amarelo é obtida ao assumirmos $\alpha = 0$ e a sequência de pontos $\mathbf{x}^k$ em verde é obtida ao assumirmos $\alpha = 1$.

Figura 5.12.9: As trajetórias dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI em função da escolha de α com a estratégia de convergência global.

5.12.3 Programação discreta

Consideremos o problema de programação discreta dado por:

$$\begin{aligned}
 \text{Min} \quad & (x_1 - 2)^4 + (x_1 - 2x_2)^2 \\
 \text{s.a :} \quad & x_1 + x_2 - 3 = 0 \\
 & x_1^2 - x_2 \leq 0 \\
 & x_1 \in \mathbb{D} \\
 & 1,5 \leq x_2 \leq 2,
 \end{aligned} \tag{5.12.3}$$

onde $\mathbb{D} = \{0,75; 0,9; 1,05; 1,2; 1,35; 1,5; 1,65\}$. As curvas de nível da função objetivo e a região viável do problema são apresentadas na Figura 5.12.10.

Figura 5.12.10: Curvas de nível da função objetivo e região viável do problema (5.12.3)

A Tabela 5.12.14 apresenta as informações sobre os parâmetros iniciais, o ponto inicial (inviável), a precisão ε e a função ψ para tratarmos com as restrições de desigualdade. O fato de que a distribuição dos elementos do $\mathbb{D}$ ser uma progressão aritmética de razão $r = 1,5$, então, de acordo com a expressão (3.2.2), p. 57, uma adequada função $m(x_1)$ é conforme apresentada nesta tabela.

Tabela 5.12.14: Dados de entrada para resolvermos o problema (5.12.3).

$\mathbf{x}^0$	$(0; 0)^T$
μ_0	0,01
α_μ	0,25
β_0	0, 1
ρ_0	10^{-3}
α_ρ	20
τ	0,1
$\delta_{1_i}^0, i = 1, \dots, 5$	0,1
$\gamma(\delta_{1_i}^0, \alpha), i = 1, \dots, 5$	$\delta_{1_i}^0 (\alpha + (1 - \alpha) \delta_{1_i}^0)^{-1}$
$\psi(t)$	$\frac{1}{2} (\arg \sinh(t) + t^2 + t\sqrt{t^2 + 1})$
$m(x_1)$	$\frac{0,15}{\pi} \sin\left(\frac{\pi}{0,15} (x_1 - 0,75)\right)$
$v(y)$	$y \ln\left(\frac{y}{2} + 1\right)$
ε	10^{-7}

Resolvemos o problema (5.12.3) por meio de (5.12.4) dado por:

$$\begin{aligned}
 \text{Min} \quad & (x_1 - 2)^4 + (x_1 - 2x_2)^2 + \rho \frac{0,15}{\pi} \sin\left(\frac{\pi}{0,15} (x_1 - 0,75)\right) \ln\left(\frac{0,15}{2\pi} \sin\left(\frac{\pi}{0,15} (x_1 - 0,75)\right) + 1\right) \\
 \text{s.a :} \quad & x_1 + x_2 - 3 = 0 \\
 & x_1^2 - x_2 \leq 0 \\
 & -x_1 + 0,75 \leq 0 \\
 & -x_2 + 1,5 \leq 0 \\
 & x_1 - 1,65 \leq 0 \\
 & x_2 - 2 \leq 0.
 \end{aligned}
 \tag{5.12.4}$$

A Tabela 5.12.15 resume a sequência numérica dos principais resultados em um ciclo k obtidos pelo método PDRNLI ao impormos $\alpha = 0,7$: função objetivo, função lagrangiana aumentada, quão satisfeita está as condições de KKT, parâmetro de reescala, parâmetro de amortecimento, parâmetro de penalidade e tamanho dos passos primal e dual.

Tabela 5.12.15: Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades para resolvermos o problema (5.12.3).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	ρ^k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	16	-32,02702844	1,60E+01	0,1	0,1	1,00E-03	-	-
1	10,03953557	1225,666811	1,56E+01	2,50E-03	0,1	2,00E-02	0,99869	1
2	2,8019495	4,871289786	2,01E+01	6,25E-04	6,27E-02	4,00E-01	0,64863	1
3	4,534543172	4,60536849	5,77E+00	1,56E-04	6,63E-02	8,00E+00	1	1
4	4,617730119	4,616475279	1,67E-01	3,91E-05	1,26E-01	1,60E+02	1	1
5	4,740008199	4,74001078	3,38E+00	9,77E-06	1,27E+00	3,20E+03	0,97263	1
6	7,183420045	7,183424124	6,77E+01	2,44E-06	2,47E+01	6,40E+04	1	1
7	56,05629665	56,05628722	1,35E+03	6,10E-07	4,90E+02	1,28E+06	1	0,99999
8	1054,50762	1054,50783	2,67E+04	1,53E-07	1,05E+04	2,56E+07	1	0,996
9	910,6941017	1102,539329	2,10E+05	3,81E-08	6,34E+03	5,12E+08	1	0,62056
10	39,37542587	39,39121398	7,67E+05	9,54E-09	3,83E+03	5,12E+08	1	0,99992
11	6,169599924	6,169599937	1,27E+01	2,38E-09	3,90E+03	5,12E+08	1	1
12	6,169599736	6,169599732	2,32E-04	5,96E-10	3,97E+03	5,12E+08	1	1
13	6,169599736	6,169599736	3,15E-08	1,49E-10	2,40E+03	5,12E+08	1	1

A Figura 5.12.11 apresenta a sequência $\{\mathbf{x}^k\}$ desenvolvidas pelo método PDRNLI sobre o plano $\mathbb{R}^2$ juntamente com as curvas de nível da função objetivo e com a região viável do problema (5.12.3).

Figura 5.12.11: A trajetória dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades para resolvermos o problema (5.12.3).

Com precisão de 7 casas decimais, a Tabelas 5.12.16 mostra as sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ geradas pelo método PDRNLI aplicado à sequência de subproblemas (5.12.4).

Tabela 5.12.16: As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades para resolvermos o problema (5.12.3).

Ciclo k	x_{1k}	x_{2k}	λ_{0k}	λ_{11k}	λ_{12k}	λ_{13k}	λ_{14k}	λ_{15k}
0	0	0	16,0098474	0,0068175	0,0068175	0,0068175	0,0004151	0,0003425
1	0,9960574	2,0000000	-9,9536509	2,2085910	0,0193860	0,0694869	0,0008286	0,0010161
2	1,4502786	1,5483361	-13,1422024	3,6970541	0,0325042	0,0871234	0,0006996	13,3903167
3	1,3083494	1,6916506	-5,0049795	3,9306813	0,0013821	0,7329600	0,0003536	1,3571292
4	1,3027835	1,6972165	-4,5075732	3,8599860	0,0000011	0,0012995	0,0000004	0,0014366
5	1,3027759	1,6972241	-4,4605570	3,9061376	0,0000000	0,0000008	0,0000000	0,0000009
6	1,3027759	1,6972241	-3,5223815	4,8438367	0,0000000	0,0000000	0,0000000	0,0000000
7	1,3027727	1,6972273	15,2378485	23,5991386	0,0000000	0,0000000	0,0000000	0,0000000
8	1,3020308	1,6979692	376,7087167	402,5607654	0,0000000	0,0000000	0,0000000	0,0000000
9	1,1915413	1,8084587	-36202,9101068	15714,8272254	0,0000000	0,0000000	0,0000000	0,0000000
10	1,2003602	1,7996398	218,2207989	203,0054204	0,0000000	0,0000000	0,0000000	0,0000000
11	1,2000001	1,7999999	-11,0026973	0,1035668	0,0000000	0,0000000	0,0000000	0,0000000
12	1,2000000	1,8000000	-9,6001072	0,0000518	0,0000000	0,0000000	0,0000000	0,0000000
13	1,2000000	1,8000000	-9,5999996	0,0000000	0,0000000	0,0000000	0,0000000	0,0000000

A Tabela 5.12.17 sintetiza a influência do parâmetro α sobre o desempenho do método PDRNLI proposto quando este parâmetro assume alguns valores que estão no intervalo $[0, 1]$. Para um dado valor para α , esta influência é observada nos seguintes quesitos: número de ciclos; função objetivo no ponto ótimo; função lagrangiana aumentada integrada no ponto ótimo, precisão sobre as condições de KKT e o valor do parâmetro de penalidade. Em todos os casos, as condições iniciais dos parâmetros e das variáveis; o tipo de fator de transformação $\gamma(\delta_{1_i}^0, \alpha)$, $i = 1, \dots, 5$, as funções $\psi(t)$, $m(x_1)$ e $v(y)$ são aquelas apresentadas na Tabela 5.12.14.

Tabela 5.12.17: Influência do parâmetro α sobre o desempenho do método PDRNLI para a classe $\mathbb{P}_{int}$ sem a estratégia de convergência global.

α	Ciclos k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	ρ^k
0	15	6,1695999867902	6,1695999866184	8,13E-08	1,02E+10
0,1	14	6,1695999867902	6,1695999867902	8,30E-08	1,02E+10
0,2	14	6,1695999867902	6,1695999867902	8,15E-04	1,02E+10
0,3	13	6,1695997358040	6,1695997358032	2,25E-08	5,12E+08
0,4	15	6,1695999867902	6,1695999867902	8,12E-08	1,02E+10
0,5	16	6,1695999867902	6,1695999867902	8,12E-08	1,02E+10
0,6	15	6,1695999867902	6,1695999867902	8,12E-08	1,02E+10
0,7	13	6,1695997358040	6,1695997358033	3,15E-08	5,12E+08
0,8	13	6,1695997358040	6,1695997358035	2,35E-08	5,12E+08
0,9	15	6,1695999867902	6,1695999867902	8,12E-08	1,02E+10
1	50	8,9370062500000	8,9370062500000	1,81E+02	1,31E+19

Observamos que ao impormos $\alpha = 1$, então o método PDRNLI convergiu para a solução viável $\mathbf{x}^* = (1,05 ; 1,95)^T$ em 50 ciclos. As condições de folgas complementares e viabilidade primal estão satisfeitas com precisão superior a 10^{-17} , todavia, a condição de viabilidade dual (primeiro conjunto de equações em (5.4.3)) é violada neste caso.

5.12.4 Programação matemática com restrições de equilíbrio

Consideremos o MPEC (5.12.5) dado por:

$$\begin{aligned}
 \text{Min} \quad & (x_1 - 2)^4 + (x_1 - 2x_2)^2 \\
 \text{s.a :} \quad & x_1 + x_2 - 3 = 0 \\
 & 1,5 \leq x_2 \leq 2 \quad \perp \quad x_1^2 - x_2
 \end{aligned} \tag{5.12.5}$$

A região viável do problema é apresentada na Figura 5.12.12. O traço em vermelho representa geometricamente as restrições de complementaridade mista deste problema.

Figura 5.12.12: Região viável e representação geométrica das restrições de complementaridade mista do problema (5.12.5).

Ao identificarmos $\mathbf{T}(\mathbf{x})$ por x_2 e $\mathbf{W}(\mathbf{x})$ por $-x_1^2 + x_2$, representamos as restrições de complementaridade mista do problema (5.12.5) conforme a formulação apresentada em (4.2.8), página 79. Assim, o MPEC (5.12.5) é equivalente ao problema de programação não-linear definido por:

$$\begin{aligned}
 \text{Min} \quad & (x_1 - 2)^4 + (x_1 - 2x_2)^2 \\
 \text{s.a :} \quad & x_1 + x_2 - 3 = 0 \\
 & -x_2 + 1,5 \leq 0 \\
 & x_2 - 2 \leq 0 \\
 & (x_2 - 1,5)(x_1^2 - x_2) \leq 0 \\
 & (x_2 - 2)(x_1^2 - x_2) \leq 0.
 \end{aligned} \tag{5.12.6}$$

A Tabela 5.12.18 apresenta as informações sobre os parâmetros iniciais, o ponto inicial (inviável), a precisão ε e a função ψ para tratarmos com as restrições de desigualdade.

Tabela 5.12.18: Dados de entrada para resolvermos o problema (5.12.6).

$\mathbf{x}^0$	$(0, 5; 0, 5)^T$
μ_0	0,1
α_μ	0,2
β_0	0,1
$\delta_{1_i}^0, i = 1, \dots, 4$	0,1
$\gamma(\delta_{1_i}^0, \alpha), i = 1, \dots, 4$	$\delta_{1_i}^0(\alpha + (1 - \alpha)\delta_{1_i}^0)^{-1}$
$\psi(t)$	$\frac{1}{2}(\arg \sinh(t) + t^2 + t\sqrt{t^2 + 1})$
ε	10^{-7}

A Tabela 5.12.19 resume a sequência numérica dos principais resultados em um ciclo k obtidos pelo método PDRNLI ao impormos $\alpha = 0,6$: função objetivo, função lagrangiana aumentada, quão

satisfeita está as condições de KKT, parâmetro de reescala, parâmetro de amortecimento e tamanho dos passos primal e dual.

Tabela 5.12.19: Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades para resolvermos o problema (5.12.6).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	5,3125000000	-7,2083865757	8,35E+00	0,1	0,1	-	-
1	3,1785084839	1450,5151566990	1,82E+01	2,00E-02	0,1	0,25344	1
2	1,5770605345	4,5944376871	9,64E+00	4,00E-03	6,09E-02	0,59093	0,38046
3	2,6386551067	3,1626389537	1,52E+01	8,00E-04	6,16E-02	0,089552	1
4	4,3818310713	4,2553251518	1,59E+02	1,60E-04	6,24E-02	0,9741	1
5	4,3549463589	4,3520504701	8,68E+01	3,20E-05	3,80E-02	1	0,55116
6	2,3125000000	129,3704934157	3,65E+03	1,39E-01	2,31E-02	0,0044926	4,62E-05
7	2,4212966157	2,2766163954	9,59E-01	2,78E-02	1,41E-02	1	1
8	2,3309866538	2,3145042746	7,67E-01	5,57E-03	1,43E-02	0,98449	1
9	2,3154891257	2,3128285210	1,26E-01	1,11E-03	8,68E-03	1	1
10	2,3130049462	2,3125490813	1,95E-02	2,23E-04	5,28E-03	1	1
11	2,3125865832	2,3125073251	3,17E-03	4,45E-05	3,22E-03	1	1
12	2,3125147989	2,3125010807	5,27E-04	8,90E-06	1,96E-03	1	1
13	2,3125024788	2,3125001530	8,73E-05	1,78E-06	1,19E-03	1	1
14	2,3125003978	2,3125000199	1,41E-05	3,56E-07	7,26E-04	1	1
15	2,3125000589	2,3125000022	2,18E-06	7,12E-08	4,42E-04	1	1
16	2,3125000075	2,3125000002	3,07E-07	1,42E-08	2,69E-04	1	1
17	2,3125000007	2,3125000000	3,69E-08	2,85E-09	1,64E-04	1	1

A Figura 5.12.13 apresenta a sequência $\{\mathbf{x}^k\}$ desenvolvida pelo método PDRNLI sobre o plano $\mathbb{R}^2$ juntamente com as curvas de nível da função objetivo e com a região viável do problema (5.12.5).

Figura 5.12.13: A trajetória dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades para resolvermos o problema (5.12.6).

Com precisão de 7 casas decimais, a Tabela 5.12.20 mostra as sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ geradas pelo método PDRNLI aplicado ao problema (5.12.6).

Tabela 5.12.20: As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades para resolvermos o problema (5.12.6).

Ciclo k	x_{1_k}	x_{2_k}	λ_{0_k}	λ_{1_1k}	λ_{1_2k}	λ_{1_3k}	λ_{1_4k}
0	0,5000000	0,5000000	6,2884478	0,0547266	0,0051943	0,0547266	0,0547266
1	0,7378170	0,7690716	-7,4504287	0,0664226	0,0115328	0,2134307	0,5028211
2	1,2251444	1,1640654	-4,9276784	0,0860407	0,0086532	0,0001067	0,3164698
3	1,1455720	1,2983352	28,3417647	53,4305575	0,0142377	0,0008948	5,6977910
4	1,3108346	1,6747620	20,2021460	34,6302287	0,0043486	0,0009743	11,6812331
5	1,3215184	1,6784816	-3,1558218	16,4472908	0,0018950	0,3983330	0,0058406
6	1,5000000	1,5000000	3657,9174159	16,6411122	0,0018847	0,0001992	0,0060048
7	1,4886894	1,5113106	3,6053033	9,7478564	0,0004337	0,0001517	0,0024200
8	1,4980582	1,5019418	3,5158002	9,5377658	0,0000203	0,0001926	0,0001608
9	1,4996855	1,5003145	3,5028052	9,5067331	0,0000002	0,0001979	0,0000021
10	1,4999469	1,5000531	3,5004765	9,5012682	0,0000000	0,0002034	0,0000000
11	1,4999909	1,5000091	3,5000819	9,5003475	0,0000000	0,0002080	0,0000000
12	1,4999984	1,5000016	3,5000140	9,5001917	0,0000000	0,0002119	0,0000000
13	1,4999997	1,5000003	3,5000023	9,5001668	0,0000000	0,0002151	0,0000000
14	1,5000000	1,5000000	3,5000004	9,5001642	0,0000000	0,0002177	0,0000000
15	1,5000000	1,5000000	3,5000001	9,5001649	0,0000000	0,0002197	0,0000000
16	1,5000000	1,5000000	3,5000000	9,5001657	0,0000000	0,0002209	0,0000000
17	1,5000000	1,5000000	3,5000000	9,5001661	0,0000000	0,0002215	0,0000000

As informações descritas na Figura 5.12.13 e na Tabela 5.12.20 mostram que a função $\mathbf{T}(\mathbf{x}) = x_2$ tornou-se ativa ao passo que $\mathbf{W}(\mathbf{x}) = x_1^2 - x_2$ tornou-se inativa em $\mathbf{x}^*$, pois temos $\mathbf{T}(\mathbf{x}^*) = 1,5$ e $\mathbf{W}(\mathbf{x}^*) = 0,75$. Nós constatamos que de acordo com a Definição 4.2.1, página 78, o minimizador $\mathbf{x}^*$ explicita a restrição de complementaridade mista $1,5 \leq x_2 \leq 2 \perp x_1^2 - x_2$ como um elemento associado ao conjunto $\mathbb{B}_1(\mathbf{x}^*)$ de restrições ativas.

Além disso, a Tabela 5.12.20 mostra que os multiplicadores de Lagrange associados ao problema equivalente (5.12.6) têm a seguinte distribuição: $\lambda_{1_1}^* \geq 0$, $\lambda_{1_2}^* = 0$, $\lambda_{1_3}^* \geq 0$ e $\lambda_{1_4}^* = 0$. Logo, de acordo com o sistema (4.2.10), página 79, concluímos que os multiplicadores de Lagrange associados respectivamente às funções $\mathbf{T}(\mathbf{x})$ e $\mathbf{W}(\mathbf{x})$ em $\mathbf{x}^*$, isto é, os vetores $\pi_1^*(\mathbf{x}^*, \lambda_{1_1}^*, \lambda_{1_2}^*, \lambda_{1_3}^*, \lambda_{1_4}^*)$ e $\pi_2^*(\mathbf{x}^*, \lambda_{1_3}^*, \lambda_{1_4}^*)$; devem ser respectivamente irrestrito e nulo.

A Tabela 5.12.21 mostra que este resultado teórico também é verificado na prática.

Tabela 5.12.21: As sequência de multiplicadores de Lagrange $\{\pi_{1_k}\}$ e $\{\pi_{2_k}\}$ associadas à restrição de complementaridade mista do problema (5.12.5) determinadas pelo método PDRNLI para a classe $\mathbb{P}_{\text{int}}$ de funções penalidades.

Ciclos k	$\pi_1(\mathbf{x}^k, \lambda_{1_{1_k}}, \lambda_{1_{2_k}}, \lambda_{1_{3_k}}, \lambda_{1_{4_k}})$	$\pi_2(\mathbf{x}^k, \lambda_{1_{3_k}}, \lambda_{1_{4_k}})$
0	-0,0221690	-0,1368164
1	0,1060503	-0,7749393
2	-0,1840464	-0,2645839
3	-53,4961022	-3,9981200
4	-35,1343521	-3,7990105
5	-16,4728510	0,0692173
6	-16,6438805	-0,0030024
7	-9,7492355	-0,0011809
8	-9,5380077	-0,0000797
9	-9,5068826	-0,0000010
10	-9,5014207	0,0000000
11	-9,5005035	0,0000000
12	-9,5003506	0,0000000
13	-9,5003282	0,0000000
14	-9,5003275	0,0000000
15	-9,5003297	0,0000000
16	-9,5003314	0,0000000
17	-9,5003322	0,0000000

A Tabela 5.12.22 resume a sequência numérica dos principais resultados em um ciclo k obtidos pelo método PDRNLI ao impormos $\alpha = 1$: função objetivo, função lagrangiana aumentada, quão satisfeita está as condições de KKT, parâmetro de reescala, parâmetro de amortecimento e tamanho dos passos primal e dual.

Tabela 5.12.22: Evolução dos principais resultados obtidos pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades para resolvermos o problema (5.12.6).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	5,3125	-7,205508855	8,32E+00	0,1	0,1	-	-
1	3,09654859	1440,205857	1,83E+01	2,00E-02	0,1	0,26457	1
2	1,954468604	3,028911177	8,53E+00	4,00E-03	6,09E-02	0,84261	0,39159
3	3,246016614	5,259347373	6,02E+01	8,00E-04	6,16E-02	0,02445	1
4	3,967062333	3,361225929	1,31E+01	1,60E-04	3,75E-02	0,87396	1
5	4,106847063	4,060773134	4,20E+00	2,63E-02	3,80E-02	1	0,94828
6	4,554891296	3,96126899	2,22E+01	5,26E-03	2,31E-02	0,41696	0,92464
7	4,526383484	4,569062876	9,85E-01	1,05E-03	1,41E-02	0,76195	0,99981
8	4,608900211	4,611411695	1,16E-01	2,10E-04	8,57E-03	0,94719	1
9	4,611441676	4,611410809	8,75E-04	4,21E-05	5,22E-03	1	1
10	4,611410727	4,611410717	5,54E-05	8,41E-06	3,18E-03	0,9957	1
11	4,611410718	4,611410718	1,48E-08	1,68E-06	1,94E-03	1	1

A Figura 5.12.14 apresenta a sequência $\{\mathbf{x}^k\}$ desenvolvidas pelo método PDRNLI sobre o plano $\mathbb{R}^2$ juntamente com as curvas de nível da função objetivo e com a região viável do problema (5.12.5).

Figura 5.12.14: A trajetória dos pontos $\mathbf{x}^k$ geradas pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades para resolvermos o problema (5.12.6).

Com precisão de 7 casas decimais, a Tabela 5.12.23 mostra as sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ geradas pelo método PDRNLI aplicado ao problema (5.12.6).

Tabela 5.12.23: As sequências $\{\mathbf{x}^k\}$, $\{\lambda_0^k\}$ e $\{\lambda_1^k\}$ determinadas pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades para resolvermos o problema (5.12.6).

Ciclo k	x_{1k}	x_{2k}	λ_{0k}	λ_{11k}	λ_{12k}	λ_{13k}	λ_{14k}
0	0,5000000	0,5000000	6,2794012	0,0414214	0,0033296	0,0414214	0,0414214
1	0,7498584	0,7792916	-7,5375551	0,0548539	0,0074013	0,2279934	0,5897812
2	1,3958269	1,3726774	-4,7476296	0,0647969	0,0055808	0,0001140	0,1133149
3	1,2750107	1,4991537	63,0159258	133,4796339	0,0470807	0,0023141	5,2801527
4	1,3337998	1,6377351	32,1271841	53,4789462	0,0041047	0,0031083	22,5910096
5	1,3402533	1,6597467	-15,1723974	0,0267395	0,0003116	72,2953725	9,9743139
6	1,3068543	1,6931457	8,7588211	0,0000134	0,0000709	20,2849922	27,0688584
7	1,3089225	1,6910775	-5,2756900	0,0000004	0,0000000	35,2515374	8,3750507
8	1,3029562	1,6970438	-4,5443719	0,0000000	0,0000000	36,2034161	10,7464495
9	1,3027734	1,6972266	-4,5097413	0,0000000	0,0000000	36,2557945	10,8788857
10	1,3027756	1,6972244	-4,5099230	0,0000000	0,0000000	36,2835768	10,8966319
11	1,3027756	1,6972244	-4,5099222	0,0000000	0,0000000	36,2835755	10,8966327

As informações descritas na Figura 5.12.14 e na Tabela 5.12.23 mostram que a solução primal $\mathbf{x}^*$ e dual λ_0^* é a mesma obtida pelo método PDRNLI para resolvermos o problema (5.12.1) conforme mostra a Tabela 5.12.3. Todavia, a solução dual λ_1^* apresentada na Tabela 5.12.23 parece ser *confusa* quando comparamos diretamente com as demais soluções duais apresentadas na Tabela 5.12.3.

Vamos esclarecer que as soluções obtidas pelo método PDRNLI para resolvermos os problemas (5.12.1) e (5.12.6) são as mesmas quando explicitarmos os multiplicadores de Lagrange π_1^* e π_2^* associados à restrição de complementaridade mista do problema (5.12.6).

Os dados da Tabela 5.12.23 mostram que a função $\mathbf{T}(\mathbf{x}) = x_2$ tornou-se inativa ao passo que $\mathbf{W}(\mathbf{x}) = x_1^2 - x_2$ tornou-se ativa em $\mathbf{x}^*$, pois temos $\mathbf{T}(\mathbf{x}^*) = 1,6972244$ e $\mathbf{W}(\mathbf{x}^*) = 0$. Nós constatamos que de acordo com a Definição 4.2.1, página 78, o minimizador $\mathbf{x}^*$ explicita a restrição

de complementaridade mista $1, 5 \leq x_2 \leq 2 \perp x_1^2 - x_2$ como um elemento associado ao conjunto $\mathbb{B}_3(\mathbf{x}^*)$ de restrições ativas.

Além disso, a Tabela 5.12.23 mostra que os multiplicadores de Lagrange associados ao problema equivalente (5.12.6) têm a seguinte distribuição: $\lambda_{1_1}^* = 0$, $\lambda_{1_2}^* = 0$, $\lambda_{1_3}^* \geq 0$ e $\lambda_{1_4}^* \geq 0$. Logo, de acordo com o sistema (4.2.10), página 79, concluímos que os multiplicadores de Lagrange associados respectivamente às funções $\mathbf{T}(\mathbf{x})$ e $\mathbf{W}(\mathbf{x})$ em $\mathbf{x}^*$, isto é, os vetores $\pi_1^*(\mathbf{x}^*, \lambda_{1_1}^*, \lambda_{1_2}^*, \lambda_{1_3}^*, \lambda_{1_4}^*)$ e $\pi_2^*(\mathbf{x}^*, \lambda_{1_3}^*, \lambda_{1_4}^*)$; devem ser respectivamente nulo e irrestrito.

A Tabela 5.12.24 mostra que este resultado teórico também é verificado na prática e, mais ainda, mostra que os multiplicadores de Lagrange ótimos associados à restrição de complementaridade mista do problema (5.12.5) são congruentes aos multiplicadores de Lagrange ótimos vinculados às restrições de desigualdade do problema (5.12.1).

Tabela 5.12.24: As sequência de multiplicadores de Lagrange $\{\pi_{1_k}\}$ e $\{\pi_{2_k}\}$ associadas à restrição de complementaridade mista do problema (5.12.5) determinadas pelo método PDRNLI para a classe $\mathbb{P}_2$ de funções penalidades.

Ciclos k	$\pi_1(\mathbf{x}^k, \lambda_{1_{1_k}}, \lambda_{1_{2_k}}, \lambda_{1_{3_k}}, \lambda_{1_{4_k}})$	$\pi_2(\mathbf{x}^k, \lambda_{1_{3_k}}, \lambda_{1_{4_k}})$
0	-0,0173810	-0,1035534
1	0,1300077	-0,8842677
2	-0,1245121	-0,0710995
3	-134,1007781	-2,6445470
4	-56,6670932	-8,1835022
5	-11,2589026	8,1551505
6	-0,6971137	-4,3882382
7	-0,9685363	4,1485343
8	-0,0305676	3,8779550
9	0,0003783	3,8567693
10	0,0000001	3,8567706
11	0,0000000	3,8567701

A Tabela 5.12.25 sintetiza a influência do parâmetro α sobre o desempenho do método PDRNLI proposto quando este parâmetro assume alguns valores que estão no intervalo $[0, 1]$. Para um dado valor para α , esta influência é observada nos seguintes quesitos: número de ciclos; função objetivo no ponto ótimo; função lagrangiana aumentada integrada no ponto ótimo, precisão sobre as condições de KKT e o valor do parâmetro de penalidade. Em todos os casos, as condições iniciais dos parâmetros e das variáveis; o tipo de fator de transformação $\gamma(\delta_{1_i}^0, \alpha)$, $i = 1, \dots, 4$, a função $\psi(t)$ é aquela apresentadas na Tabela 5.12.18.

Tabela 5.12.25: Influência do parâmetro α sobre o desempenho do método PDRNLI para a classe $\mathbb{P}_{int}$ de funções penalidades.

α	Ciclos k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	λ_{0_k}	π_{1_k}	π_{2_k}
0	12	2,3125000	2,3124942	0,0000000	3,5000053	-9,5001037	-0,0000018
0,1	21	2,3125000	2,3125000	0,0000001	3,5000000	-9,5329624	0,0000000
0,2	16	2,3125000	2,3125000	0,0000001	3,5000000	-9,5000006	0,0000000
0,3	14	2,3125000	2,3125000	0,0000001	3,5000000	-9,5000518	0,0000000
0,4	17	2,3125000	2,3125000	0,0000000	3,5000000	-9,5008602	0,0000000
0,5	-	-	-	-	-	-	-
0,6	17	2,3125000	2,3125000	0,0000000	3,5000000	-9,5003322	0,0000000
0,7	17	2,3125000	2,3125000	0,0000000	3,5000000	-9,5013206	0,0000000
0,8	16	4,6114107	4,6114107	0,0000000	-4,5099222	0,0000000	3,8567701
0,9	16	4,6114107	4,6114107	0,0000000	-4,5099222	0,0000000	3,8567701
1	11	4,6114107	4,6114107	0,0000000	-4,5099222	0,0000000	3,8567701

De acordo com a Tabela 5.12.25, o método PDRNLI divergiu ao atribuírmos $\alpha = 0, 5$.

Fluxo de Potência Ótimo Reativo

De acordo com [Rosehart, Roman e Schellenberg 2005, Platbrood et al. 2014], o Fluxo de Potência Ótimo Reativo (FPOR) é um problema de otimização não-linear, não-convexo, estático, de grande porte com variáveis contínuas e discretas e com restrições de equilíbrio. Neste trabalho, a finalidade do FPOR é a minimização das perdas de potência ativa mediante o ajuste ótimo das variáveis de controle associadas à potência reativa, de modo a satisfazer restrições físicas e técnico-operacionais do sistema elétrico. No FPOR, as variáveis de controle associadas à potência ativa são fixas.

As restrições físicas e técnico-operacionais são: as equações de balanço de potência ativa nas barras de carga e de geração, exceto a barra *slack*; as equações de balanço de potência reativa nas barras de carga; limites preestabelecidos nas magnitudes de tensão em todas as barras do sistema elétrico, nos *taps* dos transformadores em-fase, nas susceptâncias *shunt* de barra vinculadas aos bancos de capacitores e reatores e na geração de potência reativa em todas as barras de geração.

As variáveis de controle vinculadas à potência reativa são: as magnitudes de tensão das barras de geração, os *taps* dos transformadores em-fase e as susceptâncias *shunt* de barra vinculadas aos bancos de capacitores e reatores. Além das variáveis de controle, temos também as variáveis dependentes. Essas variáveis estão associadas às tensões fasoriais nas barras de carga e aos ângulos de fase de tensão nas barras de geração; exceto o ângulo de fase na barra *slack* (este é pré-especificado). As variáveis de controle e dependentes do FPOR definem o estado do sistema elétrico.

Devido ao fato do método PDRNLI proposto neste trabalho ser de gradiente projetado, Seção 5.7, temos que o atendimento das restrições de balanço de potência ativa e reativa do sistema elétrico (e das outras restrições) é realizado de forma implícita através do ajuste das variáveis de estado. Logo, não separamos variáveis de controle das variáveis dependentes para o cálculo das direções de busca primal como é apresentado pelos autores em [Dommel e Tinney 1968], os quais utilizaram o método de gradiente reduzido para atender tais restrições do problema.

6.1 Modelos de FPOR

Os modelos de FPOR explorados neste trabalho são apresentados nas subseções que se seguem. Na Seção 6.1.1, descrevemos o modelo geral, incluindo variáveis de controle contínuas e discretas e com restrições de complementaridade mista. Na Seção 6.1.2, o modelo geral é relaxado de modo

que as variáveis discretas são modeladas como contínuas. Na Seção 6.1.3, descrevemos o modelo de FPOR clássico em que as variáveis de estado são contínuas e as restrições de complementaridade não são consideradas.

6.1.1 FPOR com variáveis de controle discretas e com restrições de atuação de dispositivos de controle de tensão

Este modelo geral de FPOR é proposto em [Lage 2013, p. 146]. Sua definição é conforme (6.1.1) a seguir:

$$\begin{aligned}
 \text{Min} \quad & \sum_{k,m \in \mathcal{L}} g_{km} (V_k^2 + V_m^2 - 2V_k V_m \cos(\gamma_k - \gamma_m)) + \\
 & + \sum_{k,m \in \mathcal{T}} g_{km} \left(\frac{1}{(t_{km})^2} V_k^2 + V_m^2 - 2 \frac{1}{t_{km}} V_k V_m \cos(\gamma_k - \gamma_m) \right) \\
 \text{s.a.} : \quad & P_{G_k} - P_{C_k} - \sum_{m \in \mathcal{V}_k} P_{km} (\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{sh}) = 0 \quad \forall k \in \mathcal{C} \cup \mathcal{G}' \\
 & Q_{G_k} - Q_{C_k} + b_k^{sh} V_k^2 - \sum_{m \in \mathcal{V}_k} Q_{km} (\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{sh}) = 0 \quad \forall k \in \mathcal{C} \\
 & Q_{Ger_k}^{\min} \leq Q_{C_k} - b_k^{sh} V_k^2 + \sum_{m \in \mathcal{V}_k} Q_{km} (\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{sh}) \leq Q_{Ger_k}^{\max} \quad \forall k \in \mathcal{G} \\
 & V_k^{\min} \leq V_k \leq V_k^{\max}, \quad \forall k \in \mathcal{B} \\
 & V_k^{\min} \leq V_k \leq V_k^{\max} \perp \left(b_k^{sh} - b_{k_{espec}}^{sh} \right) \quad \forall k \in \mathcal{B}^{sh} \\
 \text{ou} \quad & V_k^{\min} \leq V_k \leq V_k^{\max} \perp (t_{km} - t_{km_{espec}}) \quad \forall k, m \in \mathcal{T}; k \in \mathcal{T}_{tap} \\
 \text{ou} \quad & V_k^{\min} \leq V_k \leq V_k^{\max} \perp -(t_{km} - t_{km_{espec}}) \quad \forall k, m \in \mathcal{T}; k \in \mathcal{T}_z \\
 & t_{km} \in \mathbb{T}_{tap} \quad \forall k, m \in \mathcal{T} \\
 & b_k^{sh} \in \mathbb{B}^{sh} \quad \forall k \in \mathcal{B}^{sh},
 \end{aligned} \tag{6.1.1}$$

onde:

$\mathcal{B}$ é o conjunto de todas as barras do sistema elétrico;

$\mathcal{C}$ é o conjunto de todas as barras de carga;

$\mathcal{G}$ é o conjunto de todas as barras de geração;

$\mathcal{G}'$ é o conjunto de todas as barras de geração, exceto a barra *slack*;

$\mathcal{B}^{sh}$ é o conjunto de todas as barras com controle de magnitude de tensão por meio de bancos de capacitores e reatores;

$\mathcal{L}$ é o conjunto de todos os ramos $k - m$ que representam as linhas de transmissão;

$\mathcal{T}$ é o conjunto de todos os ramos $k - m$ que possuem transformadores em-fase;

$\mathcal{T}_{tap}$ é o conjunto de barras conectadas a transformadores em-fase, em que ocorre a comutação do *tap* do transformador (denominado, de forma simples, de conjunto de “barras do *tap*”);

$\mathcal{T}_z$ é o conjunto de barras conectadas a transformadores em-fase em que não ocorre a comutação do *tap* do transformador (denominado, de forma simples, de conjunto de “barras da impedância”);

$\mathcal{V}_k$ é o conjunto de todas as barras vizinhas à barra k ;

g_{km} é a condutância série do ramo $k - m$ em pu ;

b_{km} é a susceptância série do ramo $k - m$ em pu ;

b_{km}^{sh} é a susceptância *shunt* do ramo $k - m$ em pu ;

P_{G_k} é a potência ativa gerada na barra k em pu ;

P_{C_k} é a potência ativa consumida na barra k em pu ;

$P_{km}(\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{sh})$ é o fluxo de potência ativa que flui da barra k para a barra m em pu ;

Q_{G_k} é a potência reativa gerada na barra k em pu ;

Q_{C_k} é a potência reativa consumida na barra k em pu ;

$Q_{km}(\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{sh})$ é o fluxo de potência reativa que flui da barra k para a barra m em pu ;

$Q_{G_k}^{\min}$ é o limitante inferior de geração de potência reativa na barra k em pu ;

$Q_{G_k}^{\max}$ é o limitante superior de geração de potência reativa na barra k em pu ;

V_k é a magnitude de tensão da barra k em pu ;

V_m é a magnitude de tensão da barra m em pu ;

$V_k^{\min}$ é o limitante inferior para a magnitude de tensão da barra k em pu ;

$V_k^{\max}$ é o limitante superior para a magnitude de tensão da barra k em pu ;

γ_k é o ângulo de fase de tensão da barra k em radianos;

γ_m é o ângulo de fase de tensão da barra m em radianos;

t_{km} é o *tap* do transformador em-fase do ramo $k - m$ em pu ;

$\mathbb{T}_{tap}$ é o conjunto discreto cujos elementos são as possibilidades para t_{km} ;

b_k^{sh} é a susceptância *shunt* do banco de capacitores e reatores conectado à barra k em pu ;

$\mathbb{B}^{sh}$ é o conjunto discreto cujos elementos são as possibilidades para b_k^{sh} ;

$t_{km_{espec}}$ é o valor especificado para o *tap* do transformador do ramo $k - m$ (geralmente fornecido no banco de dados do sistema elétrico) em pu e;

$b_{k_{espec}}^{sh}$ é o valor especificado para a susceptância *shunt* associada aos bancos de capacitores e reatores na barra k (geralmente fornecido no banco de dados do sistema elétrico) em pu .

Para uma linha de transmissão, os fluxos de potência ativa e reativa que fluem da barra k para a barra m e vice e versa são definidos respectivamente pelas expressões (6.1.2) e (6.1.3):

$$P_{km}(\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{sh}) = g_{km} V_k^2 - V_k V_m (g_{km} \cos(\gamma_k - \gamma_m) + b_{km} \sin(\gamma_k - \gamma_m)). \quad (6.1.2)$$

$$Q_{km}(\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{sh}) = - (b_{km} + b_{km}^{sh}) V_k^2 + V_k V_m (b_{km} \cos(\gamma_k - \gamma_m) - g_{km} \sin(\gamma_k - \gamma_m)). \quad (6.1.3)$$

Para um transformador em-fase do ramo $k - m$, os fluxos de potência ativa e reativa que fluem da barra k para a barra m são definidos respectivamente pelas expressões gerais (6.1.4) e (6.1.5). Estas expressões são deduzidas a partir de análise do circuito mostrado na Figura 6.1.1, na qual são definidas a barra de comutação do tap e a barra da impedância, respectivamente.

$$P_{km}(\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{sh}) = \begin{cases} g_{km} \frac{1}{t_{km}^2} V_k^2 - \frac{1}{t_{km}} V_k V_m (g_{km} \cos(\gamma_k - \gamma_m) + b_{km} \sin(\gamma_k - \gamma_m)) & , \forall k \in \mathcal{T}_{tap} \\ g_{km} V_k^2 - \frac{1}{t_{km}} V_k V_m (g_{km} \cos(\gamma_k - \gamma_m) + b_{km} \sin(\gamma_k - \gamma_m)) & , \forall k \in \mathcal{T}_z. \end{cases} \quad (6.1.4)$$

$$Q_{km}(\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{sh}) = \begin{cases} -b_{km} \frac{1}{t_{km}^2} V_k^2 + \frac{1}{t_{km}} V_k V_m (b_{km} \cos(\gamma_k - \gamma_m) - g_{km} \sin(\gamma_k - \gamma_m)) & , \forall k \in \mathcal{T}_{tap} \\ -b_{km} V_k^2 + \frac{1}{t_{km}} V_k V_m (b_{km} \cos(\gamma_k - \gamma_m) - g_{km} \sin(\gamma_k - \gamma_m)) & , \forall k \in \mathcal{T}_z. \end{cases} \quad (6.1.5)$$

Figura 6.1.1: Modelo de circuito adotado para o transformador em-fase.

As restrições de complementaridade mista em (6.1.1) são denominadas restrições de atuação de dispositivos de tensão. Sua finalidade é tornar as variáveis de controle t_{km} e b_k^{sh} passíveis de ajustes somente se a respectiva magnitude de tensão V_k na barra não puder ser mantida em seus limitantes inferior ou superior. Logo, a inserção de restrições de complementaridade modela uma mudança de comportamento do sistema, a qual está sujeita a regras diferentes sob diferentes circunstâncias.

Por exemplo, a interpretação física de $V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{sh} - b_{k_{espec}}^{sh})$ é que se $V_k = V_k^{\min}$, então a susceptância b_k^{sh} é ajustada de forma a aumentar a injeção ou diminuir a absorção de potência reativa nessa barra, ou seja, b_k^{sh} deve ou se manter em $b_{k_{espec}}^{sh}$ ou ser incrementada em relação a este valor. Se $V_k = V_k^{\max}$, então a susceptância b_k^{sh} é ajustada de forma a diminuir a injeção, ou aumentar a absorção, de potência reativa nessa barra, ou seja, b_k^{sh} ou deve se manter em $b_{k_{espec}}^{sh}$ ou ser decrementada em relação a este valor. Por fim, se $V_k^{\min} < V_k < V_k^{\max}$; então a susceptância b_k^{sh} assume somente o valor $b_{k_{espec}}^{sh}$.

Por fim, a formulação matemática do modelo (6.1.1) é adequada à formulação geral definida em (5.1.1), página 85. Isso significa que podemos aplicar o método PDRNLI com todas as estratégias metodológicas apresentadas no Capítulo 5 para resolvermos esta classe de problemas de FPOR.

6.1.2 FPOR com variáveis de controle contínuas e com restrições de atuação de dispositivos de controle de tensão

Este modelo de FPOR é a versão relaxada do modelo (6.1.1), pois assumimos que todas as variáveis de controle são contínuas e estão restritas a um intervalo. Assim, a definição do modelo de FPOR de Lage em variáveis contínuas é dada por:

$$\begin{aligned}
 \text{Min} \quad & \sum_{k,m \in \mathcal{L}} g_{km} (V_k^2 + V_m^2 - 2V_k V_m \cos(\gamma_k - \gamma_m)) + \\
 & + \sum_{k,m \in \mathcal{T}} g_{km} \left(\frac{1}{(t_{km})^2} V_k^2 + V_m^2 - 2\frac{1}{t_{km}} V_k V_m \cos(\gamma_k - \gamma_m) \right) \\
 \text{s.a :} \quad & P_{G_k} - P_{C_k} - \sum_{m \in \mathcal{V}_k} P_{km} (\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{sh}) = 0 \quad \forall k \in \mathcal{C} \cup \mathcal{G}' \\
 & Q_{G_k} - Q_{C_k} + b_k^{sh} V_k^2 - \sum_{m \in \mathcal{V}_k} Q_{km} (\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{sh}) = 0 \quad \forall k \in \mathcal{C} \\
 & Q_{Ger_k}^{\min} \leq Q_{C_k} - b_k^{sh} V_k^2 + \sum_{m \in \mathcal{V}_k} Q_{km} (\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{sh}) \leq Q_{Ger_k}^{\max} \quad \forall k \in \mathcal{G} \\
 & V_k^{\min} \leq V_k \leq V_k^{\max} \quad \forall k \in \mathcal{B} \\
 & V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{sh} - b_{k_{espec}}^{sh}) \quad \forall k \in \mathcal{B}^{sh} \\
 \text{ou} \quad & V_k^{\min} \leq V_k \leq V_k^{\max} \perp (t_{km} - t_{km_{espec}}) \quad \forall k, m \in \mathcal{T}; k \in \mathcal{T}_{tap} \\
 \text{ou} \quad & V_k^{\min} \leq V_k \leq V_k^{\max} \perp -(t_{km} - t_{km_{espec}}) \quad \forall k, m \in \mathcal{T}; k \in \mathcal{T}_z \\
 & t_{km}^{\min} \leq t_{km} \leq t_{km}^{\max} \quad \forall k, m \in \mathcal{T} \\
 & b_k^{sh\min} \leq b_k^{sh} \leq b_k^{sh\max} \quad \forall k \in \mathcal{B}^{sh},
 \end{aligned} \tag{6.1.6}$$

onde:

$t_{km}^{\min}$ é o limitante inferior para t_{km} ;

$t_{km}^{\max}$ é o limitante superior para t_{km} ;

$b_k^{sh\min}$ é o limitante inferior para b_k^{sh} e;

$b_k^{sh\max}$ é o limitante superior para b_k^{sh} .

Apesar deste modelo não representar a realidade das variáveis de controle t_{km} e b_k^{sh} de um FPOR, a comparação entre os modelos (6.1.1) e (6.1.6) nos propicia estimar as variações nas perdas totais de um sistema elétrico.

6.1.3 FPOR clássico

Se além de considerarmos a possibilidade de ajustarmos os *taps* dos transformadores em-fase e as susceptâncias *shunt* de barra de forma contínua, também não considerarmos as restrições de atuação

de dispositivos de controle de tensão, então o modelo mais simples de FPOR é segundo a expressão (6.1.7), a seguir:

$$\begin{aligned}
 \text{Min} \quad & \sum_{k,m \in \mathcal{L}} g_{km} (V_k^2 + V_m^2 - 2V_k V_m \cos(\gamma_k - \gamma_m)) + \\
 & + \sum_{k,m \in \mathcal{T}} g_{km} \left(\frac{1}{(t_{km})^2} V_k^2 + V_m^2 - 2\frac{1}{t_{km}} V_k V_m \cos(\gamma_k - \gamma_m) \right) \\
 \text{s.a :} \quad & P_{G_k} - P_{C_k} - \sum_{m \in \mathcal{V}_k} P_{km} (\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{\text{sh}}) = 0 \quad \forall k \in \mathcal{C} \cup \mathcal{G}' \\
 & Q_{G_k} - Q_{C_k} + b_k^{\text{sh}} V_k^2 - \sum_{m \in \mathcal{V}_k} Q_{km} (\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{\text{sh}}) = 0 \quad \forall k \in \mathcal{C} \quad (6.1.7) \\
 & Q_{Ger_k}^{\min} \leq Q_{C_k} - b_k^{\text{sh}} V_k^2 + \sum_{m \in \mathcal{V}_k} Q_{km} (\mathbf{V}, \gamma, \mathbf{t}, \mathbf{b}^{\text{sh}}) \leq Q_{Ger_k}^{\max} \quad \forall k \in \mathcal{G} \\
 & V_k^{\min} \leq V_k \leq V_k^{\max} \quad \forall k \in \mathcal{B} \\
 & t_{km}^{\min} \leq t_{km} \leq t_{km}^{\max} \quad \forall k, m \in \mathcal{T} \\
 & b_k^{\text{sh}, \min} \leq b_k^{\text{sh}} \leq b_k^{\text{sh}, \max} \quad \forall k \in \mathcal{B}^{\text{sh}}.
 \end{aligned}$$

Apesar de não representar as variáveis de controle t_{km} e b_k^{sh} como discretas, nem as restrições relacionadas à atuação de dispositivos de controle de tensão, o modelo (6.1.7) tem sido tradicionalmente utilizado no planejamento e operação de sistemas de potência. Todavia, para aplicações práticas na operação, este modelo pode fornecer um número elevado de ajustes em dispositivos de controle, o que pode inviabilizar a implementação prática em tempo hábil destes controles para grandes sistemas [Lage 2013, p. 145]. O modelo geral (6.1.1) busca reduzir esse número de ajustes de controles sem, contudo, alterar significativamente as perdas totais do sistema elétrico em relação ao modelo mais simples (6.1.7).

O modelo (6.1.7) tem sido intensamente pesquisado, sendo que várias técnicas de otimização têm sido propostas para resolvê-lo. Algumas abordagens são: método do gradiente reduzido [Dommel e Tinney 1968], método primal-dual de pontos interiores [Wu, Debs e Marsten 1994, Granville 1994], função de complementaridade não-linear [Torres e Quintana 1999], programação semi-definida [Lavari e Low 2012, Madani, Ashraphijuo e Lavari 2016] e método primal-dual barreira logarítmica modificada [Sousa, Baptista e Costa 2012, Pinheiro et al. 2015].

Resultados numéricos

Neste capítulo, apresentamos os resultados numéricos obtidos por meio da aplicação do método PDRNLI para resolvermos os modelos de FPOR apresentados em 6.1.1, página 130; 6.1.6, página 133 e; 6.1.7, página 134. Desenvolvemos a implementação do método PDRNLI no software MATLAB®2012b e realizamos os experimentos em um *notebook* com processador Intel®Core i3-6100 de 2,3GHz, com 4 GB de RAM e sistema operacional de 64bits Windows 10 home.

7.1 Especificações do problema de FPOR

As variáveis de estado do problema de FPOR são os vetores $\mathbf{V} = (V_1, \dots, V_k, \dots, V_{card(\mathcal{B})})$; $\gamma = (\gamma_1, \dots, \gamma_k, \dots, \gamma_{card(\mathcal{B})})$; $\mathbf{t} = (t_1, \dots, t_\kappa, \dots, t_{card(\mathcal{T})})$, onde κ é o respectivo ramo $k - m$ com k e m em $\mathcal{T}$ e; $\mathbf{b}^{sh} = (b_1^{sh}, \dots, b_k^{sh}, \dots, b_{card(\mathcal{B}^{sh})}^{sh})$, onde $card(\cdot)$ significa o cardinal de um conjunto. Por simplicidade, sempre que for conveniente, as variáveis de estado são representadas por meio do vetor $\mathbf{x} = (\mathbf{V}^k, \gamma^k, \mathbf{t}^k, \mathbf{b}^{sh^k})^T$.

Definimos para todos os modelos de FPOR que cada magnitude de tensão V_k está definida no intervalo $[0, 95 ; 1, 05]$.

Quanto aos *taps* dos transformadores em-fase, definimos que:

- Para o modelo de FPOR definido em 6.1.1, cada t_κ está em $\mathbb{T}_{tap} = \mathbb{D}_{tap} \cup \{t_{km_{espec}}\}$, onde os elementos do conjunto $\mathbb{D}_{tap}$ são definidos pela sequência $\{0, 90 + 0, 00625i\}_{i=0}^{32}$.
- Para os modelos de FPOR 6.1.6 e 6.1.7, cada t_κ está no intervalo $[0, 9 ; 1, 1]$.

Quanto aos *shunts* de barra definimos que:

- Para o modelo de FPOR definido em 6.1.1, se $b_{k_{espec}}^{sh} > 0$, então a respectiva susceptância b_k^{sh} está em $\mathbb{B}^{sh} = \{0, b_{k_{espec}}^{sh}, 2b_{k_{espec}}^{sh}\}$. Caso contrário, se $b_{k_{espec}}^{sh} < 0$, então a respectiva b_k^{sh} está em $\mathbb{B}^{sh} = \{b_{k_{espec}}^{sh}, 0\}$.
- Para os modelos de FPOR 6.1.6 e 6.1.7, se $b_{k_{espec}}^{sh} > 0$, então a respectiva susceptância b_k^{sh} está no intervalo $[0, 2b_{k_{espec}}^{sh}]$. Caso contrário, se $b_{k_{espec}}^{sh} < 0$, então a respectiva susceptância b_k^{sh} está no intervalo $[b_{k_{espec}}^{sh}, 0]$.

Quanto às gerações de potência reativa, cada limitante inferior $Q_{Ger_k}^{\min}$ e superior $Q_{Ger_k}^{\max}$, com k em $\mathcal{G}'$, assume o valor especificado pelo banco de dados do sistema convertido para pu . Se k é a barra *slack*, então definimos $Q_{Ger_k}^{\min} = -10$ e $Q_{Ger_k}^{\max} = 10$.

7.2 Os sistemas elétricos investigados e suas características

Neste trabalho, investigamos os modelos de FPOR junto aos seguintes elétricos: IEEE-14, IEEE-30 e IEEE-118 [Power 1999]. Suas características são apresentadas nas próximas subseções.

7.2.1 Características técnicas do sistema IEEE-14 barras

As características técnicas do sistema IEEE-14 barras são apresentadas na Tabela 7.2.1:

Tabela 7.2.1: Características técnicas do sistema IEEE-14 barras.

Conjunto	$card(\cdot)$
$\mathcal{C}$	9
$\mathcal{G}$	5
$\mathcal{L}$	17
$\mathcal{T}$	3
$\mathcal{B}^{sh}$	1

A topologia deste sistema elétrico é apresentada na Figura 7.2.1 a seguir:

Figura 7.2.1: Topologia do sistema IEEE-14 barras [Power 1999].

A topologia deste sistema elétrico é apresentada na Figura 7.2.3 a seguir:

Figura 7.2.3: Topologia do sistema IEEE-118 barras [Power 1999].

7.3 Resultados numéricos

7.3.1 Resultados para o sistema IEEE-14 barras

A Tabela 7.3.1 apresenta as informações sobre os parâmetros iniciais e o ponto inicial. Esclarecemos que μ é o parâmetro de reescala; α_μ é o fator de redução do parâmetro de reescala; ρ é o parâmetro de penalidade; $\alpha_{\rho_{tap}}$ é o fator de aumento associado ao parâmetro penalidade cuja função penalidade trata com o *tap* do transformador em-fase; $\alpha_{\rho_{shunt}}$ é o fator de aumento associado ao parâmetro penalidade cuja função penalidade trata com a susceptância de barra; $\tau \in [0, 1]$ é o parâmetro associado à estratégia de controle de viabilidade [Martínez 2009, seção 2.2]; β é o parâmetro de amortecimento associado à estratégia de convergência global descrita na Seção 5.8; δ_1 é a estimativa do multiplicador de Lagrange vinculada à respectiva restrição de desigualdade; $\gamma(\delta_{1_i}^0, \alpha)$ é o fator de transformação vinculado ao método PDRNLI; $\alpha \in [0, 1]$ é o parâmetro de homotopia; $\psi(t) \in \mathbb{H}$ é a função utilizada pelo método de reescalamento não-linear integrado para o tratamento das restrições de desigualdade; $m(x_i)$ é a função cujos *zeros* são os valores discretos para x_i ; $v(y) \in \mathbb{G}$ é a função vinculada à função penalidade (3.3.4) cuja finalidade é o tratamento das restrições $m(x_i) = 0$ e; ε é a precisão associada à norma-infinito do sistema não-linear primal-dual (5.4.3) e às variáveis discretas.

Tabela 7.3.1: Dados de entrada para resolvermos problema de FPOR para o sistema IEEE-14.

$V_k^0, \forall k = 1, \dots, \text{card}(\mathcal{B}).$	1
$\gamma_k^0, \forall k = 1, \dots, \text{card}(\mathcal{B}).$	0
$t_{km}^0, \forall k, m \in \mathcal{T}.$	1
$t_{kmespec}$	1
$b_k^{sh}, k \in \mathcal{B}^{sh}$	É especificado conforme o banco de dados.
$b_{k espec}^{sh}$	É especificado conforme o banco de dados.
μ_0	0,65
α_μ	0,35
$\rho_{i_0}, \forall i = 1, \dots, \text{card}(\mathcal{T} \cup \mathcal{B}^{sh}).$	10^{-1}
$\alpha_{\rho_{tap}}$	2
$\alpha_{\rho_{shunt}}$	2
τ	0,05
β_0	0,1
$\delta_{1_i}^0, \forall i = 1, \dots, r.$	0,1
$\gamma(\delta_{1_i}^0, \alpha), \forall i = 1, \dots, r.$	$\delta_{1_i}^0(\alpha + (1 - \alpha)\delta_{1_i}^0)^{-1}, \text{ com } \alpha \in [0, 1].$
α	0,4
$\psi(t)$	$\frac{1}{2}(\arg \sinh(t) + t^2 + t\sqrt{t^2 + 1})$
$m_i(t_{km}), \forall i = 1, \dots, \text{card}(\mathcal{T}).$	Função-IHDP, expressão (3.2.5).
$m_i(b_k^{sh}), \forall i = 1, \dots, \text{card}(\mathcal{B}^{sh}).$	Função-IHDP, expressão (3.2.5).
$v(y)$	$y \ln(y + 1)$
ε	10^{-5}

O método PDRNLI obteve sucesso em resolver os problemas FPOR associado ao sistema elétrico IEEE-14 descritos em (6.1.7), (6.1.6) e (6.1.1). Os tempos computacionais foram respectivamente de 0,39; 0,42 e 0,57 segundos. As Tabelas 7.3.2 - 7.3.4 resumem os valores iterativos dos principais resultados obtidos pelo método PDRNLI para resolvermos estes problemas de FPOR.

Tabela 7.3.2: Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-14 por meio do modelo (6.1.7).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	0	-1161,7164529813000	9,42E-01	0,65	0,1	-	-
1	13,8465974770953	-1147,5731103713800	2,23E-01	2,28E-01	0,1	1	1
2	13,7168591741919	-23,7564930481909	2,56E-01	7,96E-02	5,92E-02	0,058786	1
3	13,2390128824006	-7,1007502606241	1,31E-02	2,79E-02	3,50E-02	1	1
4	13,5477364145721	4,5033976757313	7,54E-03	9,75E-03	2,07E-02	1	0,9995
5	13,6173728655397	11,2798336954560	3,17E-03	3,41E-03	1,22E-02	1	1
6	13,6145489274190	13,3521341229405	1,05E-03	1,19E-03	7,24E-03	1	0,97951
7	13,6038289597941	13,5881910060051	6,54E-03	4,18E-04	4,28E-03	0,86651	0,92352
8	13,6047891947207	13,6025913752524	2,84E-04	1,46E-04	2,53E-03	1	0,91674
9	13,6044353129085	13,6042684297864	1,46E-04	5,12E-05	1,50E-03	1	0,97245
10	13,6043269536127	13,6043106906103	7,19E-04	1,79E-05	8,87E-04	0,99051	0,964
11	13,6042857967635	13,6042793723858	4,14E-03	6,28E-06	5,25E-04	1	1
12	13,6042760312642	13,6042699884636	4,01E-03	2,20E-06	3,10E-04	0,093296	0,9953
13	13,6042743087441	13,6042690347609	2,97E-03	7,69E-07	1,84E-04	0,25885	0,99505
14	13,6042689725747	13,6042688507834	7,48E-06	2,69E-07	1,09E-04	1	1

Tabela 7.3.3: Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-14 por meio do modelo (6.1.6).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	0	-1161,9861944929700	9,42E-01	0,65	0,1	-	-
1	13,5393572294608	-1147,5478838746500	2,21E-01	2,28E-01	0,1	0,98983	1
2	13,3901654862251	-23,8110255624072	2,49E-01	7,96E-02	5,92E-02	0,075726	1
3	13,3710977083254	-6,8746844253621	1,78E-01	2,79E-02	3,50E-02	0,11191	1
4	13,4179101268152	4,6319831279881	1,39E-01	9,75E-03	2,07E-02	0,22045	0,99582
5	13,4834565232709	11,3683065185127	8,92E-02	3,41E-03	1,22E-02	0,36244	0,9973
6	13,5004288664462	13,4886950852876	9,23E-02	1,19E-03	7,24E-03	0,14097	0,9555
7	13,5004288664462	13,8918689714261	1,15E-01	4,36E-03	4,28E-03	4,00E-17	1
8	13,5439580958892	13,6713827481984	4,01E-02	1,52E-03	2,53E-03	0,6635	1
9	13,5439580958892	13,6819398970501	4,19E-02	5,34E-04	1,50E-03	8,68E-16	1
10	13,6246935558702	13,6571123229545	1,02E-02	1,87E-04	8,87E-04	0,75333	1
11	13,6378966877305	13,6548447560575	5,03E-03	6,54E-05	5,25E-04	0,49148	1
12	13,6513946765982	13,6512504134140	2,31E-04	2,29E-05	3,10E-04	1	0,99964
13	13,6512517930636	13,6512252888930	3,87E-05	8,01E-06	1,84E-04	1	1
14	13,6512276593115	13,6512241028301	1,03E-05	2,80E-06	1,09E-04	1	1
15	13,6512243057499	13,6512238871895	1,08E-06	9,81E-07	6,42E-05	1	1

Tabela 7.3.4: Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-14 por meio do modelo (6.1.1).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	0	-1377,7420569626700	9,42E-01	0,65	0,1	-	-
1	13,4994650155644	-1363,4215113420500	2,20E-01	2,28E-01	0,1	0,98988	1
2	13,3615192904530	-28,3978698965122	2,77E-01	7,96E-02	6,36E-02	0,061985	1
3	13,3234512109636	-9,4433308457363	1,75E-01	2,79E-02	4,19E-02	0,15155	1
4	13,3783904741893	2,4210747753280	1,44E-01	9,75E-03	2,76E-02	0,2365	0,9971
5	13,4579271680237	10,3193460675109	1,15E-01	3,41E-03	1,69E-02	0,3459	0,99777
6	13,4634510231822	13,3137722003146	1,36E-01	1,19E-03	1,11E-02	0,030886	0,96255
7	13,9000651907265	13,8165300531287	2,20E-02	4,18E-04	1,01E-02	0,85173	0,98235
8	13,7878733695335	13,7725453654911	2,05E-02	1,46E-04	1,23E-02	0,45591	0,94923
9	13,6916989365870	13,6873803776670	7,70E-03	5,12E-05	1,49E-02	0,71209	0,99992
10	13,6524863422266	13,6514906182980	1,14E-03	1,79E-05	8,81E-03	1	1
11	13,6520222676355	13,6515233986829	6,69E-04	6,28E-06	5,21E-03	0,46206	1
12	13,6513482100370	13,6512424545260	5,66E-05	2,20E-06	3,08E-03	1	0,99996
13	13,6512449908503	13,6512254235178	1,50E-05	7,69E-07	1,82E-03	1	1
14	13,6512267747647	13,6512239660796	1,40E-06	2,69E-07	1,08E-03	1	1

A Tabela 7.3.5 mostra que para cada problema de FPOR os geradores de potência reativa estão operando dentro de seus limites.

Tabela 7.3.5: Geração de potência reativa, em MVar, obtida pelo método PDRNLI para o sistema IEEE-14.

Barra k	$Q_{Ger_k}^{\min}$	Modelo (6.1.7)	Modelo (6.1.6)	Modelo (6.1.1)	$Q_{Ger_k}^{\max}$
		Q_{Ger_k}	Q_{Ger_k}	Q_{Ger_k}	
1	-1000	-23,46436	-22,220532	-22,220484	1000
2	-40	35,197129	39,218166	39,218125	50
3	0	27,209242	28,870331	28,87031	40
6	-6	23,999979	23,999974	23,999984	24
8	-6	3,335022	15,778784	15,778801	24

As magnitudes de tensão nas barras de tensão controlada determinadas pelo método PDRNLI aplicado a cada um dos três modelos de FPOR são apresentadas na Tabela 7.3.6.

Tabela 7.3.6: *As magnitudes de tensão nas barras de tensão controlada determinadas pelo método PDRNLI para o sistema elétrico IEEE-14.*

	Modelo (6.1.7)	Modelo (6.1.6)	Modelo (6.1.1)
Barra k	V_k	V_k	V_k
1	1,050000	1,050000	1,050000
2	1,036366	1,036406	1,036406
3	1,006391	1,006076	1,006076
6	1,050000	1,022898	1,022898
8	1,023156	1,050000	1,050000

A Figura 7.3.1 mostra os valores iterativos obtidos pelo método PDRNLI das magnitudes de tensão nas barras do sistema IEEE-14. Conforme as figuras mostram, a trajetória dos valores iterativos é bem-comportada e todas as magnitudes de tensão obedecem a seus limitantes inferiores e superiores no término do processo de otimização.

Figura 7.3.1: Valores iterativos das magnitudes de tensão obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.

A Figura 7.3.2 mostra a configuração final das magnitudes de tensão nas barras do sistema IEEE-14. Conforme as figuras mostram, todas as magnitudes de tensão obedecem a seus limitantes inferi-

ores e superiores.

Figura 7.3.2: Configuração final das magnitudes de tensão nas barras do sistema elétrico IEEE-14 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.

Os *taps* dos transformadores em-fase determinados pelo método PDRNLI para o sistema IEEE-

14 são apresentados na Tabela 7.3.7.

Tabela 7.3.7: Os taps dos transformadores em-fase determinados pelo método PDRNLI para o sistema elétrico IEEE-14.

Modelo (6.1.7)					
Da barra k	Para a barra m	V_k	V_m	t_{km}	$t_{km_{espec}}$
4	7	1,015193	1,017414	1,060404	1,000000
4	9	1,015193	1,046683	0,900001	1,000000
5	6	1,019303	1,050000	0,981051	1,000000
Modelo (6.1.6)					
Da barra k	Para a barra m	V_k	V_m	t_{km}	$t_{km_{espec}}$
4	7	1,011795	1,023529	1,000000	1,000000
4	9	1,011795	1,014563	1,000000	1,000000
5	6	1,016366	1,022898	1,000000	1,000000
Modelo (6.1.1)					
Da barra k	Para a barra m	V_k	V_m	t_{km}	$t_{km_{espec}}$
4	7	1,011795	1,023529	1,000000	1,000000
4	9	1,011795	1,014563	1,000000	1,000000
5	6	1,016366	1,022898	1,000000	1,000000

A Tabela 7.3.7 mostra que no modelo (6.1.7) todos os taps dos transformadores foram ajustados, conforme previsto. Já para os modelos (6.1.6) e (6.1.1), nenhum tap do transformador em-fase foi ajustado, pois as magnitudes de tensão nas barra da impedância operam dentro dos seus limites operacionais. Em virtude da presença das restrições de complementaridade mista, o número de controles ajustados pelos modelos (6.1.6) e (6.1.1) foi reduzido em relação ao modelo (6.1.7) conforme esperado.

Por meio do Teorema 4.1.1, página 75, e por meio do estudo realizado sobre o sistema primal-dual (4.2.14), página 81, deduzimos que os multiplicadores de Lagrange associados à restrição de complementaridade mista $V_m^{\min} \leq V_m \leq V_m^{\max} \perp -(t_{km} - t_{km_{espec}})$ são representados pelos vetores $\pi_1^{\text{tap}} = (\pi_{1_1}^{\text{tap}}, \dots, \pi_{1_i}^{\text{tap}}, \dots, \pi_{1_{\text{card}(\mathcal{T})}}^{\text{tap}})^T \in \mathbb{R}^{\text{card}(\mathcal{T})}$ e $\pi_2^{\text{tap}} = (\pi_{2_1}^{\text{tap}}, \dots, \pi_{2_i}^{\text{tap}}, \dots, \pi_{2_{\text{card}(\mathcal{T})}}^{\text{tap}})^T \in \mathbb{R}^{\text{card}(\mathcal{T})}$. A componente $\pi_{1_i}^{\text{tap}}$ está vinculada ao custo de atendimento da respectiva variável V_m quando esta se torna ativa ao passo que a componente $\pi_{2_i}^{\text{tap}}$ está associada ao custo de atendimento da respectiva expressão $-(t_{km} - t_{km_{espec}})$ caso ela se torne ativa. Para cada iteração k do método PDRNLI, geramos os vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ por meio das expressões apresentadas no sistema primal-dual (4.2.14).

Apresentamos na Tabela 7.3.8 os multiplicadores de Lagrange e definidos pela última iteração do método PDRNLI aplicado aos modelos (6.1.6) e (6.1.1), respectivamente.

Tabela 7.3.8: Os multiplicadores de Lagrange π_1^{tap} e π_2^{tap} vinculados à V_m e à $-(t_{km} - t_{kmespec})$, respectivamente, determinados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14.

$m \in \mathcal{T}_z$	Modelo (6.1.6)		Modelo (6.1.1)	
	π_1^{tap}	π_2^{tap}	π_1^{tap}	π_2^{tap}
7	0,000000	0,000764	0,000000	0,000763
9	0,000000	-0,002106	0,000000	-0,002106
6	0,000000	-0,001076	0,000000	-0,001076

A Figura 7.3.3 mostra os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 por meio do modelo (6.1.6).

Figura 7.3.3: Os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.6).

A Figura 7.3.4 mostra os valores iterativos dos vetores π_1^k e π_2^k gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 por meio do modelo (6.1.1).

Figura 7.3.4: Os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.1).

A Figura 7.3.5 mostra os valores iterativos obtidos pelo método PDRNLI dos *taps* dos transformadores em-fase vinculados aos ramos $k - m$ até a convergência descrita na Tabela 7.3.7. Conforme as figuras mostram, a trajetória dos valores iterativos é bem-comportada, todos os *taps* obedecem a seus limitantes inferiores e superiores no término do processo de otimização e, em particular para problema descrito por meio do modelo (6.1.1), também satisfazem seus valores discretos.

Figura 7.3.5: Valores iterativos dos taps dos transformadores em-fase obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.

As susceptâncias *shunt* de barra determinadas pelo método PDRNLI para o sistema IEEE-14 são apresentados na Tabela 7.3.9.

Tabela 7.3.9: As susceptâncias shunt de barra determinadas pelo método PDRNLI para o sistema elétrico IEEE-14.

Barra k	Modelo (6.1.7)		Modelo (6.1.6)		Modelo (6.1.1)	
	V_k	b_k^{sh}	V_k	b_k^{sh}	V_k	b_k^{sh}
9	1,046683	0,379992	1,014563	0,190000	1,014563	0,190000

A Tabela 7.3.9 mostra que no modelo (6.1.7) a susceptância shunt de barra foi ajustada, conforme previsto. Já para os modelos (6.1.6) e (6.1.1), a susceptância shunt de barra não foi ajustada, pois a respectiva magnitude de tensão na barra controlada não atingiu nenhum de seus limitantes. Em virtude da presença das restrições de complementaridade mista, o número de controles ajustados pelos modelos (6.1.6) e (6.1.1) foi reduzido em relação ao modelo (6.1.7).

Por meio do Teorema 4.1.1, página 75, e por meio do estudo realizado sobre o sistema primal-dual (4.2.10), página 79, deduzimos que os multiplicadores de Lagrange associados à restrição de complementaridade mista $V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{sh} - b_{k_{espec}}^{sh})$ são representados pelos vetores $\pi_1^{sh} = \left(\pi_{1_1}^{sh}, \dots, \pi_{1_i}^{sh}, \dots, \pi_{1_{card(\mathcal{B}^{sh})}}^{sh} \right)^T \in \mathbb{R}^{card(\mathcal{B}^{sh})}$ e $\pi_2^{sh} = \left(\pi_{2_1}^{sh}, \dots, \pi_{2_i}^{sh}, \dots, \pi_{2_{card(\mathcal{B}^{sh})}}^{sh} \right)^T \in \mathbb{R}^{card(\mathcal{B}^{sh})}$. A componente $\pi_{1_i}^{sh}$ está vinculada ao custo de atendimento da respectiva variável V_k quando esta se torna ativa; a componente $\pi_{2_i}^{sh}$ está associada ao custo de atendimento da respectiva expressão $(b_k^{sh} - b_{k_{espec}}^{sh})$ caso ela se torne ativa. Para cada iteração k do método PDRNLI, geramos os vetores $\pi_1^{sh,k}$ e $\pi_2^{sh,k}$ por meio das expressões apresentadas no sistema primal-dual (4.2.10).

Apresentamos na Tabela 7.3.10 os multiplicadores de Lagrange π_1^{sh} e π_2^{sh} definidos pela última iteração do método PDRNLI aplicado aos modelos (6.1.6) e (6.1.1), respectivamente.

Tabela 7.3.10: Os multiplicadores de Lagrange π_1^{sh} e π_2^{sh} e vinculados à V_k e à $(b_k^{sh} - b_{k_{espec}}^{sh})$, respectivamente, determinados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14.

$k \in \mathcal{B}^{sh}$	Modelo (6.1.6)		Modelo (6.1.1)	
	π_1^{sh}	π_2^{sh}	π_1^{sh}	π_2^{sh}
9	0,000000	0,015191	0,000000	0,011395

A Figura 7.3.6 mostra os valores iterativos dos vetores $\pi_1^{sh,k}$ e $\pi_2^{sh,k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 por meio do modelo (6.1.6).

Figura 7.3.6: Os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$, vinculados às restrições $V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{\text{sh}} - b_{k_{\text{espec}}}^{\text{sh}})$, gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.6).

A Figura 7.3.7 mostra os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 por meio do modelo (6.1.1).

Figura 7.3.7: Os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$, vinculados às restrições $V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{\text{sh}} - b_{k_{\text{espec}}}^{\text{sh}})$, gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.1).

A Figura 7.3.8 mostra os valores iterativos obtidos pelo método PDRNLI da susceptância *shunt* de barra até a convergência descrita na Tabela 7.3.9. Conforme as figuras mostram, a trajetória dos valores iterativos é bem-comportada, a susceptância obedece a seus limitante inferior e superior no término do processo de otimização e, em particular para problema descrito por meio do modelo (6.1.1), também satisfazem seus valores discretos.

Figura 7.3.8: Valores iterativos da susceptância capacitiva de barra obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.

A Tabela 7.3.11 mostra o valor final dos parâmetros de penalidades vinculados ao tratamento das variáveis discretas. No caso dos *taps* dos transformadores em-fase, o parâmetro de penalidade está

associado ao respectivo ramo $k - m$, o qual é conforme apresentado na Tabela 7.3.7. No caso da susceptância *shunt* de barra, o parâmetro de penalidade está associado à respectiva barra controlada apresentada na Tabela 7.3.9.

Tabela 7.3.11: Valor final do parâmetro de penalidade definido pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.1).

Variáveis de Controle	ρ^k
Taps dos transformadores em-fase	8,192000
	4,096000
	8,192000
Susceptância <i>shunt</i> de barra	0,102400

Os valores iterativos associado às componentes do vetor de parâmetros de penalidade ρ^k do método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.1) são apresentados na Figura 7.3.9.

Figura 7.3.9: Valores iterativos vinculados às componentes do vetor de parâmetros de penalidade ρ^k obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14 mediante o modelo (6.1.1).

O condicionamento numérico na norma-2 da matriz hessiana θ_k do método PDRNLI, denotado por $\kappa(\theta_k)$, é um dos indicadores utilizados para avaliarmos a dificuldade computacional deste método em determinar um minimizador local para o problema de programação não-linear formulado de acordo com a expressão (5.3.2), página 88. Por definição, quanto maior é $\kappa(\theta_k)$, pior é o condicionamento da matriz θ_k e mais instável é o cálculo de sua inversa. Ao contrário, quanto menor é $\kappa(\theta_k)$, melhor é o condicionamento da matriz θ_k e mais estável é o cálculo de sua inversa.

Com base na Proposição B.1.1, o parâmetro de amortecimento β da estratégia de convergência global proposta neste trabalho, denominada Condição de Positividade, também nos permite aproximar a matriz θ_k , a qual pode ser singular, por uma matriz $\hat{\theta}_k$ não-singular e que seja suficientemente próxima à matriz θ_k . Portanto, esta estratégia de regularização é uma propriedade natural da Condição de Positividade.

Com a finalidade de mensurarmos o condicionamento numérico da matriz θ_k ao longo do processo iterativo, os valores de condicionamento $\kappa(\theta_k)$ foram calculados antes e após a aplicação da estratégia de convergência global. Para isto, consideremos as abreviações ACQ para designar Antes da Condição de Positividade e DCP para designar Depois da Condição de Positividade. De acordo com os resultados apresentados na Tabela (7.3.12) a seguir, o número $\kappa(\theta_k)$ é significativamente melhorado quando os valores $\kappa(\theta_k)$ ACQ e DCQ são comparados entre si e respectivamente para todos os modelos de FPOR estudados para o sistema IEEE-14.

Tabela 7.3.12: Monitoramento do condicionamento numérico da matriz hessiana do método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-14.

Ciclo k	Modelo (6.1.7)		Modelo (6.1.6)		Modelo (6.1.1)	
	$\kappa(\theta_k)$ ACP	$\kappa(\theta_k)$ DCP	$\kappa(\theta_k)$ ACP	$\kappa(\theta_k)$ DCP	$\kappa(\theta_k)$ ACP	$\kappa(\theta_k)$ DCP
0	8,09E+06	8422,4	1,13E+07	5369,6	8,09E+06	8422,4
1	1,76E+07	2,20E+05	2,35E+07	1,15E+05	1,76E+07	2,20E+05
2	3,53E+09	7273,6	6,78E+09	8164,9	3,53E+09	7273,6
3	1,64E+12	9483,9	4,58E+12	9893,9	1,64E+12	9483,9
4	1,81E+14	6130,2	3,87E+14	6505,8	1,81E+14	6130,2
5	3,94E+16	5311,5	1,28E+17	6887,1	3,94E+16	5311,5
6	4,96E+16	8556,1	1,14E+18	17252	4,96E+16	8556,1
7	7,25E+16	19771	6,35E+16	20836	7,25E+16	19771
8	1,93E+17	27650	2,05E+17	2,34E+05	1,93E+17	27650
9	4,00E+17	68590	1,53E+18	3,48E+05	4,00E+17	68590
10	2,80E+17	3,98E+06	4,61E+18	1,65E+06	2,80E+17	3,98E+06
11	1,10E+18	1,24E+07	1,19E+18	7,87E+06	1,10E+18	1,24E+07
12	1,89E+18	1,20E+08	6,52E+17	1,17E+08	1,89E+18	1,20E+08
13	1,33E+18	7,22E+08	8,74E+17	8,26E+08	1,33E+18	7,22E+08
14	2,65E+18	7,22E+08	2,11E+18	4,62E+09	2,65E+18	7,22E+08
15			8,11E+18	4,62E+09		

A Tabela 7.3.13 sintetiza as variações proporcionais das perdas de potência na transmissão mediante os três modelos. Os valores mostrados na tabela foram calculados do seguinte modo: sejam p as perdas obtidas pelo modelo A e q as perdas obtidas pelo modelo B; então os resultados de interesse são aqueles em $q < p$, pois a parte fracionária de $\frac{p}{q}$ representa o aumento proporcional das perdas obtidas, em percentual, pelo modelo A em relação às perdas obtidas pelo modelo B.

Tabela 7.3.13: Comparação entre as proporções das perdas de potência ativa associado ao sistema IEEE-14.

	A = Modelo (6.1.7)	A = Modelo (6.1.6)	A = (6.1.1)
B = Modelo (6.1.7)	1,0000000	1,0034515	1,0034517
B = Modelo (6.1.6)	-	1,0000000	1,0000002
B = Modelo (6.1.1)	-	-	1,0000000

Uma das principais conclusões obtidas com base na Tabela 7.3.13 é que o aumento proporcional das perdas, em megawatts, ao utilizarmos o modelo clássico (6.1.7) e o modelo (6.1.1) é de apenas

0,0034517% com o ajuste de nenhum dispositivo de controle de tensão. Este resultado mostra que o modelo (6.1.1) produz soluções que são operacionalmente mais simples de serem implementadas em tempo hábil pelo operador do sistema elétrico sem comprometer consideravelmente as perdas na transmissão.

Por fim, consideremos a hipótese de resolvermos o FPOR dado pelo modelo (6.1.1) por uma técnica de enumeração explícita associada às variáveis discretas; isto é, para cada possibilidade de escolha para os *taps* dos transformadores em-fase e para as susceptâncias *shunt* de barras, então resolvemos um FPOR com essas variáveis fixadas. Assim, depois de resolvermos todas as possibilidades, então escolhemos a melhor combinação. Verifiquemos quanto tempo será necessário para resolvermos todas as possibilidades ao assumirmos que cada FPOR é resolvido pelo método PDRNLI em 10^{-6} segundos.

O sistema IEEE-14 possui 3 *taps* dos transformadores e 1 susceptâncias *shunt* capacitivas de barra. Cada *tap* do transformador em-fase tem 33 possibilidades e a susceptâncias *shunt* capacitivas de barra tem 3 possibilidades. Portanto, devemos resolver $33^3 \times 3 = 107811$ problemas de FPOR. O tempo total para resolvermos todos esses problemas é 0,1078 segundos. Neste trabalho nós determinamos uma solução viável em apenas 0,39 segundos para uma única tentativa.

7.3.2 Resultados para o sistema IEEE-30 barras

A Tabela 7.3.14 apresenta as informações sobre os parâmetros iniciais e o ponto inicial.

Tabela 7.3.14: Dados de entrada para resolvermos problema de FPOR para o sistema IEEE-30.

$V_k^0, \forall k = 1, \dots, \text{card}(\mathcal{B}).$	1
$\gamma_k^0, \forall k = 1, \dots, \text{card}(\mathcal{B}).$	0
$t_{km}^0, \forall k, m \in \mathcal{T}.$	1.
$t_{km_{espec}}$	1.
$b_k^{sh0}, k \in \mathcal{B}^{sh}$	É especificado conforme o banco de dados.
$b_{k_{espec}}^{sh}$	É especificado conforme o banco de dados.
μ_0	0,05
α_μ	0,35
$\rho_{i_0}, \forall i = 1, \dots, \text{card}(\mathcal{T} \cup \mathcal{B}^{sh}).$	10^{-1}
$\alpha_{\rho_{tap}}$	2
$\alpha_{\rho_{shunt}}$	2
τ	0,05
β_0	0,1
$\delta_{1_i}^0, \forall i = 1, \dots, r.$	1
$\gamma(\delta_{1_i}^0, \alpha), \forall i = 1, \dots, r.$	$(\delta_{1_i}^0)^\alpha, \text{ com } \alpha \in [0, 1].$
α	0,7
$\psi(t)$	$\frac{1}{2}(\arg \sinh(t) + t^2 + t\sqrt{t^2 + 1})$
$m_i(t_{km}), \forall i = 1, \dots, \text{card}(\mathcal{T}).$	Função-IHDP, expressão (3.2.5).
$m_i(b_k^{sh}), \forall i = 1, \dots, \text{card}(\mathcal{B}^{sh}).$	Função-IHDP, expressão (3.2.5).
$v(y)$	$y \ln(y + 1)$
ε	10^{-6}

O método PDRNLI obteve sucesso em resolver o problema FPOR associado ao sistema elétrico

IEEE-30 através dos modelos (6.1.7), (6.1.6) e (6.1.1). Os tempos computacionais foram respectivamente de 0,57; 0,68 e 0,76 segundos. As Tabelas 7.3.15 – 7.3.17 resumem os valores iterativos dos principais resultados obtidos pelo método PDRNLI para resolvermos estes problemas de FPOR.

Tabela 7.3.15: Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-30 por meio do modelo (6.1.7).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	0	-1694,7726757079400	9,49E-01	0,05	0,1	-	-
1	16,8341367532335	-1644,9683999332600	2,46E-01	1,75E-02	0,1	0,97187	1
2	18,0073904942752	-49,0004994840478	1,15E-02	6,13E-03	5,92E-02	1	0,96145
3	18,1105483436420	6,4214643654299	3,23E-03	2,14E-03	3,50E-02	1	1
4	17,8560132594638	16,7418458233632	2,18E-03	7,50E-04	2,07E-02	1	1
5	17,7707469045878	17,6948720372860	2,80E-03	2,63E-04	1,22E-02	1	0,9316
6	17,7591373626640	17,7510648073220	1,51E-03	9,19E-05	7,24E-03	0,96026	0,99719
7	17,7568560776254	17,7564152243821	1,55E-03	3,22E-05	4,28E-03	1	1
8	17,7565835968940	17,7565662310124	1,60E-03	1,13E-05	2,53E-03	1	1
9	17,7565067266341	17,7565163939165	1,27E-03	3,94E-06	1,50E-03	1	1
10	17,7565185921577	17,7565180485803	6,93E-06	1,38E-06	8,87E-04	1	0,82245
11	17,7565135411856	17,7565136282007	1,41E-05	4,83E-07	5,25E-04	1	1
12	17,7565136299779	17,7565136226655	2,11E-07	1,69E-07	3,10E-04	1	1

Tabela 7.3.16: Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-30 por meio do modelo (6.1.6).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	0	-1723,9176613133200	9,49E-01	0,05	0,1	-	-
1	16,7508937075720	-1646,0428980820800	2,45E-01	1,75E-02	0,1	0,96935	1
2	16,8565684966987	-46,4469750748935	2,23E-01	6,13E-03	5,92E-02	0,090279	0,95795
3	16,9820649696221	8,5190065946631	2,00E-01	2,14E-03	3,50E-02	0,11015	0,90774
4	17,6603964248561	16,7086647967118	6,10E-02	7,50E-04	2,07E-02	0,72473	0,98548
5	17,8786231961829	17,7988836436450	8,72E-03	2,63E-04	1,22E-02	0,88008	0,92684
6	17,8845093520840	17,9516364279122	3,34E-02	9,19E-05	7,24E-03	0,24501	1
7	17,8892560145771	17,9719418972821	2,97E-02	3,22E-05	4,28E-03	0,24566	1
8	17,9089465595978	17,9095334155694	8,47E-03	1,13E-05	2,53E-03	0,96163	0,9712
9	17,9068986521021	17,9078199634748	2,01E-03	3,94E-06	1,50E-03	0,77054	1
10	17,9061286233008	17,9111916801559	2,66E-03	5,65E-06	8,87E-04	1	1
11	17,9061910867881	17,9061972830565	4,79E-04	1,98E-06	5,25E-04	0,99623	1
12	17,9062294413614	17,9068392011559	1,43E-04	6,93E-07	3,10E-04	0,58386	1
13	17,9062541137115	17,9062883088547	3,68E-05	2,42E-07	1,84E-04	0,92488	1
14	17,9062556136721	17,9062559983442	3,04E-06	8,48E-08	1,09E-04	0,97027	1
15	17,9062556005237	17,9062556123526	1,46E-06	2,97E-08	6,42E-05	0,49588	1
16	17,9062555863630	17,9062555821843	1,02E-07	1,04E-08	3,80E-05	0,96097	1

Tabela 7.3.17: Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-30 por meio do modelo (6.1.1).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	0	-1723,8871509881700	9,51E-01	0,05	0,1	-	-
1	16,6132242249688	-1646,1289423290900	2,44E-01	1,75E-02	0,1	0,96661	1
2	16,8552436812448	-46,5295408219191	2,56E-01	6,13E-03	1,05E-01	0,16977	0,95719
3	16,9711196022836	8,4930682890812	2,40E-01	2,14E-03	1,27E-01	0,075151	0,91623
4	17,6807616554776	16,7408330881342	7,48E-02	7,50E-04	1,49E-01	0,73989	0,98762
5	17,8824096767499	17,8277887837979	2,07E-02	2,63E-04	2,40E-01	0,85523	0,94097
6	17,8925370257549	17,9690623449915	3,45E-02	9,19E-05	2,91E-01	0,31059	1
7	17,9108392659181	17,9197060645646	1,65E-02	3,22E-05	1,72E-01	0,84872	1
8	17,9103291609827	17,9124445705711	3,82E-03	1,13E-05	1,02E-01	0,75446	1
9	17,9100122739711	17,9101643080460	9,89E-04	3,94E-06	6,03E-02	0,91527	1
10	17,9099410858179	17,9099150914704	1,78E-04	1,38E-06	3,57E-02	1	1
11	17,9099152717655	17,9099089731011	5,59E-05	4,83E-07	2,11E-02	0,97691	1
12	17,9099096166106	17,9099087623624	2,63E-06	1,69E-07	1,25E-02	0,9824	1
13	17,9099088160601	17,9099087549246	2,06E-07	5,91E-08	7,39E-03	1	1

A Tabela 7.3.18 mostra que para cada modelo de FPOR os geradores de potência reativa estão operando dentro de seus limites.

Tabela 7.3.18: Geração de potência reativa, em MVar, obtida pelo método PDRNLI para o sistema IEEE-30.

Barra k	$Q_{Ger_k}^{\min}$	Modelo (6.1.7)	Modelo (6.1.6)	Modelo (6.1.1)	$Q_{Ger_k}^{\max}$
		Q_{Ger_k}	Q_{Ger_k}	Q_{Ger_k}	
1	-1000	-26,932825	-27,596194	-27,608151	1000
2	-40	43,030374	49,825825	49,831214	50
5	-40	32,928504	36,108603	36,110875	40
8	-10	29,815346	39,999999	39,999996	40
11	-6	22,078001	15,839689	15,84684	24
13	-6	15,297342	24	23,999997	24

As magnitudes de tensão nas barras de tensão controlada determinadas pelo método PDRNLI aplicado a cada um dos três modelos de FPOR são apresentadas na Tabela 7.3.19.

Tabela 7.3.19: As magnitudes de tensão nas barras de tensão controlada determinadas pelo método PDRNLI para o sistema elétrico IEEE-30.

Barra k	Modelo (6.1.7)	Modelo (6.1.6)	Modelo (6.1.1)
	V_k	V_k	V_k
1	1,050000	1,050000	1,050000
2	1,035432	1,036418	1,036423
5	1,003016	1,004339	1,004347
8	1,008985	1,008259	1,008265
11	1,050000	1,050000	1,050000
13	1,050000	1,049195	1,049170

A Figura 7.3.10 mostra os valores iterativos obtidos pelo método PDRNLI das magnitudes de tensão nas barras do sistema IEEE-30. Conforme as figuras mostram, a trajetória dos valores iterativos é bem-comportada e todas as magnitudes de tensão obedecem a seus limitantes inferiores e superiores no término do processo de otimização.

Figura 7.3.10: Valores iterativos das magnitudes de tensão obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.

A Figura 7.3.11 mostra a configuração final das magnitudes de tensão nas barras do sistema IEEE-30. Conforme as figuras mostram, todas as magnitudes de tensão obedecem a seus limitantes

inferiores e superiores.

Figura 7.3.11: Configuração final das magnitudes de tensão nas barras do sistema elétrico IEEE-30 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.

Os taps dos transformadores em-fase determinados pelo método PDRNLI para o sistema IEEE-

30 são apresentados na Tabela 7.3.20.

Tabela 7.3.20: Os taps dos transformadores em-fase determinados pelo método PDRNLI para o sistema elétrico IEEE-30.

Modelo (6.1.7)					
Da barra k	Para a barra m	V_k	V_m	t_{km}	$t_{km_{espec}}$
6	9	1,012075	1,006265	1,099999	1,000000
6	10	1,012075	1,029967	0,941663	1,000000
4	12	1,013647	1,029604	0,983610	1,000000
28	27	1,006784	1,030619	0,954478	1,000000
Modelo (6.1.6)					
Da barra k	Para a barra m	V_k	V_m	t_{km}	$t_{km_{espec}}$
6	9	1,007527	1,018622	1,000000	1,000000
6	10	1,007527	1,009262	1,000000	1,000000
4	12	1,011372	1,017171	1,000000	1,000000
28	27	1,004153	0,985963	1,000000	1,000000
Modelo (6.1.1)					
Da barra k	Para a barra m	V_k	V_m	t_{km}	$t_{km_{espec}}$
6	9	1,007534	1,018608	1,000000	1,000000
6	10	1,007534	1,009244	1,000000	1,000000
4	12	1,011378	1,017144	1,000000	1,000000
28	27	1,004159	0,985949	1,000000	1,000000

A Tabela 7.3.20 mostra que no modelo (6.1.7) todos os taps dos transformadores foram ajustados, conforme previsto. Já para os modelos (6.1.6) e (6.1.1), nenhum ajuste sobre os taps dos transformadores em-fase foi necessário. Em virtude da presença das restrições de complementaridade mista, o número de controles ajustados pelos modelos (6.1.6) e (6.1.1) foi reduzido em relação ao modelo (6.1.7) conforme esperado.

Apresentamos na Tabela 7.3.21 os multiplicadores de Lagrange π_1^{tap} e π_2^{tap} definidos pela última iteração do método PDRNLI aplicado aos modelos (6.1.6) e (6.1.1), respectivamente.

Tabela 7.3.21: Os multiplicadores de Lagrange π_1^{tap} e π_2^{tap} vinculados à V_m e à $(t_{km} - t_{km_{espec}})$, respectivamente, determinados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30.

$m \in \mathcal{T}_z$	Modelo (6.1.6)		Modelo (6.1.1)	
	π_1^{tap}	π_2^{tap}	π_1^{tap}	π_2^{tap}
9	0,000000	0,006890	0,000000	0,006845
10	0,000000	0,000503	0,000000	0,000499
12	0,000000	0,001245	0,000000	0,001212
27	0,000000	-0,025711	0,000000	-0,025713

A Figura 7.3.12 mostra os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 por meio do modelo (6.1.6).

Figura 7.3.12: Os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.6).

A Figura 7.3.13 mostra os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 por meio do modelo (6.1.1).

Figura 7.3.13: Os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.1).

A Figura 7.3.14 mostra os valores iterativos obtidos pelo método PDRNLI dos *taps* dos transformadores em-fase vinculados aos ramos $k-m$ até a convergência descrita na Tabela 7.3.20. Conforme as figuras mostram, a trajetória dos valores iterativos é bem-comportada, todos os *taps* obedecem a seus limitantes inferiores e superiores no término do processo de otimização e, em particular para problema descrito por meio do modelo (6.1.1), também satisfazem seus valores discretos.

Figura 7.3.14: Valores iterativos dos taps dos transformadores em-fase obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.

As susceptâncias *shunt* de barra determinadas pelo método PDRNLI para o sistema IEEE-30 são apresentados na Tabela 7.3.22.

Tabela 7.3.22: As susceptâncias shunt de barra determinadas pelo método PDRNLI para o sistema elétrico IEEE-30.

Barra k	Modelo (6.1.7)		Modelo (6.1.6)		Modelo (6.1.1)	
	V_k	b_k^{sh}	V_k	b_k^{sh}	V_k	b_k^{sh}
10	1,029967	0,380000	1,009262	0,190000	1,009244	0,190000
24	1,014573	0,086000	0,983312	0,043000	0,983292	0,043000

A Tabela 7.3.22 mostra que no modelo (6.1.7) a susceptância shunt de barra foi ajustada, conforme previsto. Já para os modelos (6.1.6) e (6.1.1), as susceptâncias shunt de barra não foram ajustadas, pois suas respectivas magnitudes de tensão não atingiriam nenhum de seus limitantes. Em virtude da presença das restrições de complementaridade mista, o número de controles ajustados pelos modelos (6.1.6) e (6.1.1) foi reduzido em relação ao modelo (6.1.7).

Apresentamos na Tabela 7.3.23 os multiplicadores de Lagrange π_1^{sh} e π_2^{sh} definidos pela última iteração do método PDRNLI aplicado aos modelos (6.1.6) e (6.1.1), respectivamente.

Tabela 7.3.23: Os multiplicadores de Lagrange π_1^{sh} e π_2^{sh} e vinculados à V_k e à $(b_k^{sh} - b_{k_{spec}}^{sh})$, respectivamente, determinados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30.

$k \in \mathcal{B}^{sh}$	Modelo (6.1.6)		Modelo (6.1.1)	
	π_1^{sh}	π_2^{sh}	π_1^{sh}	π_2^{sh}
10	0,000000	0,101753	0,000000	0,102036
24	0,000000	0,974690	0,000000	0,170205

A Figura 7.3.15 mostra os valores iterativos dos vetores $\pi_1^{sh,k}$ e $\pi_2^{sh,k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 por meio do modelo (6.1.6).

Figura 7.3.15: Os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$, vinculados às restrições $V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{\text{sh}} - b_{k_{\text{espec}}}^{\text{sh}})$, gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.6).

A Figura 7.3.16 mostra os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 por meio do modelo (6.1.1).

Figura 7.3.16: Os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$, vinculados às restrições $V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{\text{sh}} - b_{k_{\text{espec}}}^{\text{sh}})$, gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.1).

A Figura 7.3.17 mostra os valores iterativos obtidos pelo método PDRNLI das susceptâncias *shunt* de barra até a convergência descrita na Tabela 7.3.22. Conforme as figuras mostram, a trajetória dos valores iterativos é bem-comportada, a susceptância obedece a seus limitante inferior e superior no término do processo de otimização e, em particular para problema descrito por meio do modelo (6.1.1), também satisfazem seus valores discretos.

Figura 7.3.17: Valores iterativos das susceptância shunt de barra obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.

A Tabela 7.3.24 mostra o valor final dos parâmetros de penalidades vinculados ao tratamento das variáveis discretas. No caso dos *taps* dos transformadores em-fase, o parâmetro de penalidade está

associado ao respectivo ramo $k - m$, o qual é conforme apresentado na Tabela 7.3.20. No caso das susceptâncias *shunt* de barra, o parâmetro de penalidade está associado à respectiva barra controlada, a qual é conforme apresentada na Tabela 7.3.22.

Tabela 7.3.24: Valor final do parâmetro de penalidade definido pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.1).

Variáveis de Controle	ρ^k
Taps dos transformadores em-fase	102,4
	102,4
	102,4
	102,4
Susceptância <i>shunt</i> de barra	51,2
	51,2

Os valores iterativos associado às componentes do vetor de parâmetros de penalidade ρ^k do método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.1) são apresentados na Figura 7.3.18.

Figura 7.3.18: Valores iterativos vinculados às componentes do vetor de parâmetros de penalidade ρ^k obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30 mediante o modelo (6.1.1).

Com a finalidade de mensurarmos o condicionamento numérico (na norma-2) da matriz θ_k ao longo do processo iterativo, os valores de condicionamento $\kappa(\theta_k)$ foram calculados antes e após a aplicação da estratégia de convergência global. Para isto, consideremos as abreviações ACQ para designar Antes da Condição de Positividade e DCP para designar Depois da Condição de Positividade. De acordo com os resultados apresentados na Tabela (7.3.25) a seguir, o número $\kappa(\theta_k)$ é significativamente melhorado quando os valores $\kappa(\theta_k)$ ACQ e DCQ são comparados entre si e respectivamente para todos os modelos de FPOR estudados para o sistema IEEE-30.

Tabela 7.3.25: Monitoramento do condicionamento numérico da matriz hessiana do método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-30.

Ciclo k	Modelo (6.1.7)		Modelo (6.1.6)		Modelo (6.1.1)	
	$\kappa(\theta_k)$ ACP	$\kappa(\theta_k)$ DCP	$\kappa(\theta_k)$ ACP	$\kappa(\theta_k)$ DCP	$\kappa(\theta_k)$ ACP	$\kappa(\theta_k)$ DCP
0	4,19E+08	21385	4,19E+08	21385	4,20E+08	21417
1	2,62E+08	22609	2,63E+08	22660	2,64E+08	18373
2	4,99E+09	29163	2,52E+09	15816	2,61E+09	9487,5
3	1,01E+12	14695	1,46E+10	14503	1,59E+10	4061,8
4	2,68E+14	19127	1,34E+12	34190	1,63E+12	4835
5	1,40E+16	1,25E+05	4,50E+13	1,44E+05	6,35E+13	7025,2
6	3,95E+18	6,56E+05	1,29E+17	4,60E+05	1,00E+17	11175
7	9,82E+18	3,23E+06	1,67E+18	1,11E+06	2,77E+18	54502
8	8,40E+19	1,56E+07	8,47E+18	9,19E+06	2,64E+18	2,59E+05
9	1,14E+20	7,53E+07	1,06E+20	4,50E+07	6,02E+18	1,54E+06
10	2,92E+20	3,64E+08	1,13E+19	1,28E+08	1,82E+18	1,28E+07
11	4,91E+20	1,76E+09	6,16E+18	8,72E+08	8,22E+17	8,25E+07
12	5,64E+20	1,76E+09	4,41E+18	3,76E+09	1,31E+18	4,57E+08
13			2,48E+18	2,05E+10	1,57E+18	4,57E+08
14			2,00E+18	1,08E+11		
15			3,11E+19	4,99E+11		
16			2,42E+18	4,99E+11		

A Tabela 7.3.26 sintetiza as variações proporcionais das perdas de potência na transmissão mediante os três modelos. O cálculo é feito da seguinte forma: Sejam p as perdas obtidas pelo modelo A e q as perdas obtidas pelo modelo B; então os resultados de interesse são aqueles em $q < p$, pois a parte fracionária de $\frac{p}{q}$ representa o aumento proporcional das perdas obtidas, em por cento, pelo modelo A em relação às perdas obtidas pelo modelo B.

Tabela 7.3.26: Comparação entre as proporções das perdas de potência ativa associado ao sistema IEEE-30.

	A = Modelo (6.1.7)	A = Modelo (6.1.6)	A = (6.1.1)
B = Modelo (6.1.7)	1,000000	1,008433	1,008639
B = Modelo (6.1.6)	-	1,000000	1,000204
B = Modelo (6.1.1)	-	-	1,000000

Uma das principais conclusões obtidas com base na Tabela 7.3.26 é que o aumento proporcional das perdas, em megawatts, ao utilizarmos o modelo clássico (6.1.7) e o modelo (6.1.1) é de apenas 0,0086397% com o ajuste de nenhum dispositivo de controle de tensão. Este resultado mostra que o modelo (6.1.1) produz soluções que são operacionalmente mais simples de serem implementadas em tempo hábil pelo operador do sistema elétrico sem comprometer consideravelmente as perdas na transmissão.

Por fim, consideremos a hipótese de resolvermos o FPOR dado pelo modelo (6.1.1) por uma técnica de enumeração explícita associada às variáveis discretas; isto é, para cada possibilidade de escolha para os *taps* dos transformadores em-fase e para as susceptâncias *shunt* de barras, então resolvemos um

FPOR com essas variáveis fixadas. Assim, depois de resolvermos todas as possibilidades, então escolhemos a melhor combinação. Verifiquemos quanto tempo será necessário para resolvermos todas as possibilidades ao assumirmos que cada FPOR é resolvido pelo método PDRNLI em 10^{-6} segundos.

O sistema IEEE-30 possui 4 *taps* dos transformadores e possui 2 susceptâncias *shunt* capacitivas de barra. Cada *tap* do transformador em-fase tem 33 possibilidades e cada susceptâncias *shunt* capacitiva de barra tem 3 possibilidades. Portanto, devemos resolver $33^4 \times 3^2 = 10673289$ problemas de FPOR. O tempo total para executarmos todas essas possibilidades é 10,673 segundos. Neste trabalho nós determinamos uma solução viável em apenas 0,76 segundos para uma única tentativa.

7.3.3 Resultados para o sistema IEEE-118 barras

A Tabela 7.3.27 apresenta as informações sobre os parâmetros iniciais e o ponto inicial.

Tabela 7.3.27: Dados de entrada para resolvermos problema de FPOR para o sistema IEEE-118.

$V_k^0, \forall k = 1, \dots, \text{card}(\mathcal{B}).$	1
$\gamma_k^0, \forall k = 1, \dots, \text{card}(\mathcal{B}).$	0
$t_{km}^0, \forall k, m \in \mathcal{T}.$	1
$t_{km_{espec}}$	1.
$b_k^{sh}, k \in \mathcal{B}^{sh}$	É especificado conforme o banco de dados.
$b_{k_{espec}}^{sh}$	É especificado conforme o banco de dados.
μ_0	0,05
α_μ	0,3
$\rho_{i_0}, \forall i = 1, \dots, \text{card}(\mathcal{T}).$	10^{-1}
$\rho_{i_0}, \forall i = 1, \dots, \text{card}(\mathcal{B}^{sh}).$	10^{-4}
$\alpha_{\rho_{tap}}$	2,2
$\alpha_{\rho_{shunt}}$	2
$\tau_i, \forall i = 1, \dots, \text{card}(\mathcal{T}).$	0,45
$\tau_i, \forall i = 1, \dots, \text{card}(\mathcal{B}^{sh}).$	0,05
β_0	0,1
$\delta_{1_i}^0, \forall i = 1, \dots, r.$	0,2
$\gamma(\delta_{1_i}^0, \alpha), \forall i = 1, \dots, r.$	$1 + (\delta_{1_i}^0 - 1) \alpha, \text{ com } \alpha \in [0, 1].$
α	0,3
$\psi(t)$	$\frac{1}{2} (\arg \sinh(t) + t^2 + t\sqrt{t^2 + 1})$
$m_i(t_{km}), \forall i = 1, \dots, \text{card}(\mathcal{T}).$	Função-IHDP, expressão (3.2.5).
$m_i(b_k^{sh}), \forall i = 1, \dots, \text{card}(\mathcal{B}^{sh}).$	Função-IHDP, expressão (3.2.5).
$v(y)$	$y \ln(y + 1)$
ε	10^{-5}

O método PDRNLI obteve sucesso em resolver o problema FPOR associado ao sistema elétrico IEEE-118 através dos modelos (6.1.7), (6.1.6) e (6.1.1). Os tempos computacionais foram respectivamente de 1,89; 2,91 e 3,88 segundos. As Tabelas 7.3.28 - 7.3.30 resumem os valores iterativos dos principais resultados obtidos pelo método PDRNLI para resolvermos estes problemas de FPOR.

Tabela 7.3.28: Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-118 por meio do modelo (6.1.7).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	0	-4179,9770034581500	1,02E+01	0,05	0,1	-	-
1	73,5039831213395	-4039,6919070693400	1,20E+00	1,50E-02	0,1	0,76182	1
2	124,1364896937690	-373,3695070280180	2,56E-01	4,50E-03	5,98E-02	1	1
3	121,5092254996290	-37,8231210706716	6,04E-02	1,35E-03	3,58E-02	0,81347	1
4	118,7453305200290	70,9148142448776	1,03E-02	4,05E-04	2,14E-02	0,96818	1
5	118,4083833696160	103,9860129753210	9,68E-02	1,50E-03	1,28E-02	0,23558	1
6	117,6379639672800	109,9688839158680	5,03E-02	4,50E-04	7,66E-03	0,5911	1
7	117,3236819955400	111,4645675312650	3,40E-02	1,35E-04	4,58E-03	0,5618	1
8	117,1650157627040	113,8528311506100	2,43E-02	4,05E-05	2,74E-03	0,56283	1
9	117,0678395624150	115,7664098074880	6,07E-03	1,22E-05	1,64E-03	0,85691	1
10	117,0617991208370	116,6368410123160	1,37E-02	3,65E-06	9,80E-04	0,21114	1
1	117,0405946562050	116,9130121423090	1,10E-03	1,09E-06	5,86E-04	1	1
12	117,0371401413420	116,9987975124480	4,96E-04	3,28E-07	3,51E-04	0,88741	1
13	117,0365780595830	117,0250364539510	6,21E-04	9,85E-08	2,10E-04	0,72109	1
14	117,0362991093950	117,0328592663740	6,13E-05	2,95E-08	1,25E-04	0,95568	1
15	117,0362504839080	117,0352240051890	3,04E-06	8,86E-09	7,50E-05	1	1

Tabela 7.3.29: Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-118 por meio do modelo (6.1.6).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	0	-4196,3095210413900	1,02E+01	0,05	0,1	-	-
1	51,8934538447794	-4028,1847479827300	1,99E+00	1,50E-02	0,1	0,64289	1
2	103,4834088378770	-365,4908175023130	4,38E-01	4,50E-03	5,98E-02	0,76267	1
3	103,4834088378770	-32,0398675607962	4,38E-01	1,35E-03	3,58E-02	9,46E-16	1
4	117,1787409140240	72,3619945125473	6,23E-02	4,05E-04	2,14E-02	0,84523	1
5	117,2892227426280	106,2012702890300	1,93E-01	1,22E-04	1,28E-02	0,13761	1
6	117,9019563874320	114,6894170495410	3,01E-02	1,39E-03	7,66E-03	0,82598	1
7	117,8472984672440	115,6498741372660	1,46E-01	4,17E-04	4,58E-03	0,18593	1
8	117,8768538793650	115,9367867824530	9,35E-02	1,25E-04	2,74E-03	0,21564	1
9	117,8797113772400	116,5851708314550	1,03E-01	3,75E-05	1,64E-03	0,04339	1
10	117,8797113772400	117,5442123338470	1,24E-01	1,13E-05	9,80E-04	2,94E-17	1
1	117,9442416501920	117,6562860367640	9,44E-03	3,38E-06	5,86E-04	0,84579	1
12	117,9512407636610	117,8743368689590	1,65E-02	5,31E-05	3,51E-04	0,69075	1
13	117,9496740937360	117,8908031477890	1,04E-01	1,59E-05	2,10E-04	0,17796	1
14	117,9485244610500	117,8960417645210	6,19E-02	4,78E-06	1,25E-04	0,32176	1
15	117,9477170347200	117,9028571130980	2,57E-02	1,43E-06	7,50E-05	0,53423	1
16	117,9470948434870	117,9182379662740	8,96E-03	4,30E-07	4,49E-05	0,61977	1
17	117,9464954831800	117,9334851951970	1,31E-03	1,29E-07	2,68E-05	0,81272	1
18	117,9463173778930	117,9421652967180	4,47E-03	3,87E-08	1,61E-05	0,50617	1
19	117,9461469929240	117,9449668999110	5,34E-04	1,63E-06	9,60E-06	0,72806	1
20	117,9460841786360	117,9453167711890	1,07E-03	4,89E-07	5,74E-06	0,65959	1
21	117,9460762414820	117,9453372906660	6,86E-04	1,47E-07	3,44E-06	0,22464	1
22	117,9460568200920	117,9453861150220	1,95E-04	4,40E-08	2,05E-06	0,70931	1
23	117,9460506106170	117,9455169345170	6,98E-05	1,32E-08	1,23E-06	0,62154	1
24	117,9460465533870	117,9457352466420	7,75E-06	3,96E-09	7,35E-07	0,91858	1

Tabela 7.3.30: Valores iterativos dos principais resultados gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema elétrico IEEE-118 por meio do modelo (6.1.1).

Ciclo k	$f(\mathbf{x}^k)$	$L_\alpha(\omega^k)$	$\ \mathbf{a}^k\ _\infty$	μ_k	β_k	$\bar{\alpha}_p$	$\bar{\alpha}_d$
0	0,00000000000000	-4196,3095210413900	1,02E+01	0,05	0,1	-	-
1	53,1412801620950	-4029,6956748075100	1,94E+00	1,50E-02	0,1	0,65018	1
2	105,6823122258870	-363,7468382818520	4,95E-01	4,50E-03	9,17E-02	0,78522	1
3	105,8091339457460	-32,7035251266391	4,95E-01	1,35E-03	8,87E-02	8,19E-16	1
4	116,9049554496820	72,8506506434542	0,09916300000000	4,05E-04	6,93E-02	0,79356	1
5	117,6757806219460	105,1176502446140	9,94E-02	1,22E-04	5,56E-02	0,5988	1
6	117,8586364238940	114,3093816239870	8,34E-02	3,65E-05	1,41E-01	0,58902	1
7	117,9030289646920	117,0142034891760	4,78E-02	1,44E-04	8,43E-02	0,59579	1
8	117,9214138656440	117,3645021175430	5,87E-02	4,33E-05	5,04E-02	0,3205	1
9	117,9367779810040	117,5781323344810	6,35E-02	1,30E-05	3,01E-02	0,17401	1
10	117,9553025045550	117,6471343340690	1,44E-02	3,90E-06	1,80E-02	0,92818	1
11	117,9507313607490	117,8248098773620	1,70E-03	1,17E-06	1,08E-02	0,91364	1
12	117,9495230708630	117,9086689534380	9,74E-03	3,51E-07	6,45E-03	0,48882	1
13	117,9492998977490	117,9371762712560	2,26E-02	1,05E-07	3,86E-03	0,17916	1
14	117,9478063459690	117,9498133483420	9,59E-03	1,44E-06	2,31E-03	0,8993	1
15	117,9475745843220	117,9519682050070	7,85E-04	2,38E-06	1,38E-03	0,91676	1
16	117,9475555281330	117,9463565627660	1,98E-03	7,14E-07	8,26E-04	0,43885	1
17	117,9475363493180	117,9458451842070	6,55E-03	2,14E-07	4,94E-04	0,84124	1
18	117,9475305440220	117,9459302113850	8,43E-03	6,43E-08	2,95E-04	0,85865	1
19	117,9475316850630	117,9464227691380	2,25E-02	1,93E-08	1,77E-04	0,13115	1
20	117,9475286293470	117,9469909701350	2,32E-02	5,78E-09	1,06E-04	0,65987	1
21	117,9475264521030	117,9473340110240	8,25E-03	1,74E-09	6,32E-05	0,96366	1
22	117,9475261884210	117,9474664664040	9,21E-04	5,21E-10	3,78E-05	0,8883	1
23	117,9475261333100	117,9475081118750	9,07E-05	1,56E-10	2,26E-05	0,90164	1
24	117,9475261200710	117,9475207091430	8,90E-06	4,68E-11	1,35E-05	0,90271	1

A Tabela 7.3.31 mostra que para cada modelo de FPOR os geradores de potência reativa estão operando dentro de seus limites.

Tabela 7.3.31: Geração de potência reativa, em MVar, obtida pelo método PDRNLI para o sistema IEEE-118.

Barra k	$Q_{Ger_k}^{\min}$	Modelo (6.1.7)	Modelo (6.1.6)	Modelo (6.1.1)	$Q_{Ger_k}^{\max}$
		Q_{Ger_k}	Q_{Ger_k}	Q_{Ger_k}	
1	-5	14,999956	14,999998	15,000000	15
4	-300	24,406549	97,193662	97,099594	300
6	-13	26,932222	37,302562	37,289062	50
8	-300	-83,407610	-113,142855	-112,760493	300
10	-147	-90,412992	-103,136893	-103,281937	200
12	-35	55,050127	61,629370	61,627569	120
15	-10	29,879413	29,999900	30,000000	30
18	-16	36,071661	37,487444	37,508480	50
19	-8	23,998947	23,999967	24,000000	24
24	-300	2,513590	2,389555	2,080243	300
25	-47	139,998488	140,000000	133,034474	140
26	-1000	-198,645400	-201,438952	-194,904503	1000
27	-300	36,709975	36,191636	35,919386	300
31	-300	27,898922	28,105982	28,119462	300
32	-14	26,964939	26,735279	26,650682	42
34	-8	5,691805	23,995797	23,999958	24
36	-8	23,792615	23,999886	24,000000	24
40	-300	35,342741	36,924774	36,950808	300
42	-300	34,704079	34,115155	34,118401	300
46	-100	-3,712735	-1,864463	-1,861698	100
49	-85	83,342766	83,030768	83,037488	210
54	-300	40,847382	41,208290	41,206865	300
55	-8	22,962621	22,998447	22,999956	23
56	-8	14,997616	14,999927	15,000000	15
59	-60	103,708024	131,669567	131,671953	180
61	-100	0,629774	-26,907041	-26,909778	300
62	-20	13,454679	14,099874	14,100789	20
65	-67	-67,000031	-67,000000	-67,000000	200
66	-67	-67,000025	-27,085501	-27,095912	200
69	-1000	-200,208691	-12,414234	-12,350906	1000
70	-10	26,606504	31,999973	32,000000	32
72	-100	-5,239435	-4,947900	-4,941851	100
73	-100	-2,165860	-0,691503	-0,668327	100
74	-6	8,999898	8,999998	9,000000	9
76	-8	22,999972	22,999998	23,000000	23
77	-20	60,570499	69,999986	70,000000	70
80	-165	-3,270440	58,811949	58,826849	280
85	-8	22,999925	22,999997	23,000000	23
87	-100	3,241053	4,668585	4,668756	1000
89	-210	-39,933119	-33,372286	-33,371550	300
90	-300	49,913468	49,941127	49,941177	300
91	-100	4,871517	5,207695	5,207767	100
92	-3	8,999936	8,999997	9,000000	9
99	-100	-3,099867	-3,103850	-3,103671	100
100	-50	56,586037	60,411420	60,411244	155
103	-15	22,052391	23,187931	23,187505	40
104	-8	22,999567	22,999983	23,000000	23
105	-8	12,269370	21,744017	21,745569	23
107	-200	6,367512	6,376622	6,376169	200
110	-8	22,636958	22,999985	23,000000	23
111	-100	-0,147475	1,953220	1,953164	1000
112	-100	12,746196	14,538337	14,538265	1000
113	-100	-6,077662	-4,699196	-4,676572	200
116	-1000	50,443680	-218,792771	-218,793769	1000

As magnitudes de tensão nas barras de tensão controlada determinadas pelo método PDRNLI aplicado a cada um dos três modelos de FPOR são apresentadas na Tabela 7.3.32.

Tabela 7.3.32: As magnitudes de tensão nas barras de tensão controlada determinadas pelo método PDRNLI para o sistema elétrico IEEE-118.

Barra k	Mod. (6.1.7) V_k	Mod. (6.1.6) V_k	Mod. (6.1.1) V_k	Barra k	Mod. (6.1.7) V_k	Mod. (6.1.6) V_k	Mod. (6.1.1) V_k
1	1,017620	1,014131	1,014133	65	1,050000	1,050000	1,050000
4	1,049052	1,048843	1,048846	66	1,050000	1,050000	1,050000
6	1,041604	1,040452	1,040453	69	1,050000	1,050000	1,050000
8	1,033503	1,037574	1,037620	70	1,027876	1,026568	1,026517
10	1,044654	1,040445	1,040397	72	1,032042	1,030942	1,030787
12	1,038684	1,036500	1,036497	73	1,027436	1,027235	1,027183
15	1,037302	1,030041	1,030011	74	1,012199	1,004417	1,004386
18	1,039680	1,032392	1,032361	76	0,995544	0,991466	0,991450
19	1,035846	1,028560	1,028526	77	1,037590	1,034708	1,034705
24	1,042128	1,040239	1,039973	80	1,050000	1,050000	1,050000
25	1,050000	1,050000	1,050000	85	1,044234	1,039812	1,039812
26	1,050000	1,050000	1,050000	87	1,050000	1,050000	1,050000
27	1,035419	1,031675	1,031424	89	1,050000	1,050000	1,050000
31	1,031039	1,025836	1,025670	90	1,033359	1,033354	1,033354
32	1,033922	1,029695	1,029481	91	1,037969	1,037940	1,037940
34	1,046976	1,037319	1,037319	92	1,039238	1,038761	1,038760
36	1,045107	1,034959	1,034960	99	1,045111	1,045109	1,045109
40	1,025694	1,017310	1,017329	100	1,050000	1,050000	1,050000
42	1,027543	1,021333	1,021349	103	1,041826	1,041887	1,041887
46	1,036584	1,035275	1,035281	104	1,031780	1,031754	1,031755
49	1,050000	1,049018	1,049023	105	1,029454	1,029403	1,029403
54	1,029084	1,027086	1,027089	107	1,016289	1,016242	1,016242
55	1,028018	1,025928	1,025932	110	1,025945	1,024415	1,024416
56	1,028038	1,025990	1,025994	111	1,033946	1,033961	1,033961
59	1,050000	1,047067	1,047069	112	1,010228	1,009805	1,009805
61	1,050000	1,046630	1,046631	113	1,046758	1,039652	1,039597
62	1,045930	1,043292	1,043293	116	1,049723	1,031858	1,031858

A Figura 7.3.19 mostra os valores iterativos obtidos pelo método PDRNLI das magnitudes de tensão nas barras do sistema IEEE-118. Conforme as figuras mostram, a trajetória dos valores iterativos é bem-comportada e todas as magnitudes de tensão obedecem a seus limitantes inferiores e superiores no término do processo de otimização.

Figura 7.3.19: Valores iterativos das magnitudes de tensão obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.

A Figura 7.3.20 mostra a configuração final das magnitudes de tensão nas barras do sistema IEEE-118. Conforme as figuras mostram, todas as magnitudes de tensão obedecem a seus limitantes

inferiores e superiores.

Figura 7.3.20: Configuração final das magnitudes de tensão nas barras do sistema elétrico IEEE-118 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.

Os *taps* dos transformadores em-fase determinados pelo método PDRNLI para o sistema IEEE-

118 são apresentados na Tabela 7.3.33.

Tabela 7.3.33: Os taps dos transformadores em-fase determinados pelo método PDRNLI para o sistema elétrico IEEE-118.

Modelo (6.1.7)					
Da barra k	Para a barra m	V_k	V_m	t_{km}	$t_{km_{espec}}$
8	5	1,033503	1,050000	0,981465	1,000000
26	25	1,050000	1,050000	1,061084	1,000000
30	17	1,044023	1,050000	0,983589	1,000000
38	37	1,043892	1,050000	0,981940	1,000000
63	59	1,040490	1,050000	0,980406	1,000000
64	61	1,048101	1,050000	1,001072	1,000000
65	66	1,050000	1,050000	0,988306	1,000000
68	69	1,048050	1,050000	0,949253	1,000000
81	80	1,050000	1,050000	0,987579	1,000000
Modelo (6.1.6)					
Da barra k	Para a barra m	V_k	V_m	t_{km}	$t_{km_{espec}}$
8	5	1,037574	1,044540	1,000000	1,000000
26	25	1,050000	1,050000	1,059095	1,000000
30	17	1,050000	1,042090	1,000000	1,000000
38	37	1,050000	1,038747	1,000000	1,000000
63	59	1,047771	1,047067	1,000000	1,000000
64	61	1,049739	1,046630	1,000000	1,000000
65	66	1,050000	1,050000	1,000000	1,000000
68	69	1,040729	1,050000	1,000000	1,000000
81	80	1,050000	1,050000	1,000000	1,000000
Modelo (6.1.1)					
Da barra k	Para a barra m	V_k	V_m	t_{km}	$t_{km_{espec}}$
8	5	1,037620	1,044550	1,000000	1,000000
26	25	1,050000	1,050000	1,056250	1,000000
30	17	1,050000	1,042052	1,000000	1,000000
38	37	1,050000	1,038748	1,000000	1,000000
63	59	1,047772	1,047069	1,000000	1,000000
64	61	1,049740	1,046631	1,000000	1,000000
65	66	1,050000	1,050000	1,000000	1,000000
68	69	1,040729	1,050000	1,000000	1,000000
81	80	1,050000	1,050000	1,000000	1,000000

A Tabela 7.3.33 mostra que no modelo (6.1.7) todos os taps dos transformadores foram ajustados, conforme previsto. Já para os modelos (6.1.6) e (6.1.1), apenas o tap do transformador em-fase do ramo 26 – 25 foi ajustado, pois a magnitude de tensão na barra da impedância 25 atingiu seu limitante superior, isto é, 1,05pu. Além disso, esta tabela mostra que os taps dos transformadores em-fase dos ramos 65 – 66, 68 – 69 e 81 – 80 não foram ajustados mesmo que as respectivas magnitudes de tensão nas barras da impedância atingiram seu limitante superior. Em virtude da presença das restrições de complementaridade mista, o número de controles ajustados pelos modelos (6.1.6) e (6.1.1) foi reduzido em relação ao modelo (6.1.7) conforme esperado.

Apresentamos na Tabela 7.3.34 os multiplicadores de Lagrange π_1^{tap} e π_2^{tap} definidos pela última iteração do método PDRNLI aplicado aos modelos (6.1.6) e (6.1.1), respectivamente.

Tabela 7.3.34: Os multiplicadores de Lagrange π_1^{tap} e π_2^{tap} vinculados à V_m e à $-(t_{km} - t_{km\text{espec}})$, respectivamente, determinados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118.

$m \in \mathcal{T}_z$	Modelo (6.1.6)		Modelo (6.1.1)	
	π_1^{tap}	π_2^{tap}	π_1^{tap}	π_2^{tap}
5	0,000000	-0,100559	0,000000	-0,100439
25	-0,122752	0,000000	-0,122447	0,000000
17	0,000000	-0,086603	0,000000	-0,085618
37	0,000000	-0,132146	0,000000	-0,131817
59	0,000000	-0,048249	0,000000	-0,048255
61	0,000000	-0,064748	0,000000	-0,064757
66	0,000000	-0,137318	0,000000	-0,137347
69	0,000000	-0,052469	0,000000	-0,052470
80	0,000000	-0,061572	0,000000	-0,061538

A Figura 7.3.21 mostra os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 por meio do modelo (6.1.6).

Figura 7.3.21: Os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante o modelo (6.1.6).

A Figura 7.3.22 mostra os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 por meio do modelo (6.1.1).

Figura 7.3.22: Os valores iterativos dos vetores $\pi_1^{\text{tap},k}$ e $\pi_2^{\text{tap},k}$ gerados pelo método PDRNLI para o sistema IEEE-118 por meio do modelo de FPOR (6.1.1).

No contexto de MPECs, as magnitudes de tensão nas barras da impedância e os *taps* dos transformadores em-fase dos ramos 65 – 66, 68 – 69 e 81 – 80 são elementos do conjunto $\mathbb{B}_4$ das restrições bi-ativas, vide Definição 4.2.2, página 80. Sabemos que se um problema de programação não-linear possui restrições de complementaridade formuladas conforme (4.1.7), página 72, cuja solução ótima é tal que o conjunto bi-ativo seja não-vazio, então as condições de KKT deste problema estão garantidas somente sob a hipótese de condição de qualificação de Guignard, a qual é a mais fraca dentre as condição de qualificação conhecidas na literatura, vide um exemplo em [Flegel 2005, p. 22]. Como os resultados mostram que os minimizadores destes problemas possuem restrições bi-ativas para o sistema IEEE-118, concluímos que a solução primal-dual determinadas pelo método PDRNLI só pode ter existido sob a hipótese de condição de qualificação de Guignard. Conforme descrito em [Flegel e Kanzow 2005], situações práticas em que ocorrem condições de qualificação de Guignard são muito raras.

Em [Liu e Sun 2004, Leyffer, López-Calva e Nocedal 2006, Fletcher et al. 2006, Kadrani, Dussault e Benchakroun 2015], os autores mostram respectivamente que os métodos de pontos interiores, de

programação quadrática sequencial e métodos tipo penalidade garantem solução primal-dual para MPECs mediante a condição de qualificação MPEC-LICQ. Já em [Flegel e Kanzow 2005, teor. 4.5], os autores demonstram que se a condição de qualificação de MPEC-LICQ é verificada, então a condição de qualificação de Guignard também se verifica, todavia a recíproca não é verdadeira.

Como o método PDRNLI proposto foi capaz de encontrar uma solução primal-dual mesmo para problemas que apresentam condições de qualificação estritamente mais fracas que a MPEC-LICQ, então este método se torna atrativo para resolver uma maior classe de problemas MPECs e de problemas de programação não-linear. Entretanto, são necessário investigações futuras no sentido de comparar o desempenho do método proposto com os métodos descritos em [Liu e Sun 2004, Leyffer, López-Calva e Nocedal 2006, Fletcher et al. 2006, Kadrani, Dussault e Benchakroun 2015].

A Figura 7.3.23 mostra os valores iterativos obtidos pelo método PDRNLI dos *taps* dos transformadores em-fase vinculados aos ramos $k-m$ até a convergência descrita na Tabela 7.3.33. Conforme as figuras mostram, a trajetória dos valores iterativos é bem-comportada, todos os *taps* obedecem a seus limitantes inferiores e superiores no término do processo de otimização e, em particular para problema descrito por meio do modelo (6.1.1), também satisfazem seus valores discretos.

Figura 7.3.23: Valores iterativos dos taps dos transformadores em-fase obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.

As susceptâncias *shunt* de barra determinadas pelo método PDRNLI para o sistema IEEE-118 são apresentados na Tabela 7.3.35.

Tabela 7.3.35: As susceptâncias shunt de barra determinadas pelo método PDRNLI para o sistema elétrico IEEE-118.

Barra k	Modelo (6.1.7)		Modelo (6.1.6)		Modelo (6.1.1)	
	V_k	b_k^{sh}	V_k	b_k^{sh}	V_k	b_k^{sh}
5	1,050000	-0,000053	1,044540	-0,400000	1,044550	-0,400000
34	1,046976	0,147779	1,037319	0,140000	1,037319	0,140000
37	1,050000	-0,071063	1,038747	-0,250000	1,038748	-0,250000
44	1,025584	0,044725	1,022639	0,100000	1,022643	0,100000
45	1,026383	0,199970	1,019939	0,100000	1,019944	0,100000
46	1,036584	0,099607	1,035275	0,100000	1,035281	0,100000
48	1,044419	0,079139	1,046338	0,150000	1,046343	0,150000
74	1,012199	0,239999	1,004417	0,120000	1,004386	0,120000
79	1,038923	0,399975	1,031911	0,200000	1,031909	0,200000
82	1,034670	0,347024	1,025236	0,200000	1,025235	0,200000
83	1,037069	0,174700	1,027104	0,100000	1,027103	0,100000
105	1,029454	0,282744	1,029403	0,200000	1,029403	0,200000
107	1,016289	0,060066	1,016242	0,060000	1,016242	0,060000
110	1,025945	0,116603	1,024415	0,060000	1,024416	0,060000

A Tabela 7.3.35 mostra que no modelo (6.1.7) as susceptâncias shunt de barra foram ajustadas, conforme previsto. Já para os modelos (6.1.6) e (6.1.1), as susceptâncias shunt de barra não foram ajustadas, pois suas respectivas magnitudes de tensão não atingiriam nenhum de seus limitantes. Em virtude da presença das restrições de complementaridade mista, o número de controles ajustados pelos modelos (6.1.6) e (6.1.1) foi reduzido em relação ao modelo (6.1.7).

Apresentamos na Tabela 7.3.36 os multiplicadores de Lagrange π_1^{sh} e π_2^{sh} definidos pela última iteração do método PDRNLI aplicado aos modelos (6.1.6) e (6.1.1), respectivamente.

Tabela 7.3.36: Os multiplicadores de Lagrange π_1^{sh} e π_2^{sh} e vinculados à V_k e à $(b_k^{\text{sh}} - b_{k_{\text{espec}}}^{\text{sh}})$, respectivamente, determinados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118.

$k \in \mathcal{B}^{\text{sh}}$	Modelo (6.1.6)		Modelo (6.1.1)	
	π_1^{sh}	π_2^{sh}	π_1^{sh}	π_2^{sh}
5	-0,000003	2,498415	0,000000	0,929270
34	0,000000	0,030128	0,000000	0,045994
37	0,000000	0,228277	0,000000	1,909983
44	0,000000	0,153961	0,000000	0,121699
45	0,000000	0,184554	0,000000	0,134805
46	0,000000	0,042096	0,000000	0,060435
48	-0,000001	0,004359	0,000000	0,012678
74	0,000000	0,209002	0,000000	0,146539
79	0,000000	0,080097	0,000000	0,081509
82	0,000000	0,139692	0,000000	0,114733
83	0,000000	0,145161	0,000000	0,105855
105	0,000000	0,062383	0,000000	0,096144
107	0,000000	0,205594	0,000000	0,145078
110	0,000000	0,128987	0,000000	0,122403

A Figura 7.3.24 mostra os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 por meio do modelo (6.1.6).

Figura 7.3.24: Os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$, vinculados às restrições $V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{\text{sh}} - b_{k_{\text{espec}}}^{\text{sh}})$, gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante o modelo (6.1.6).

A Figura 7.3.25 mostra os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$ gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 por meio do modelo (6.1.1).

Figura 7.3.25: Os valores iterativos dos vetores $\pi_1^{\text{sh},k}$ e $\pi_2^{\text{sh},k}$, vinculados às restrições $V_k^{\min} \leq V_k \leq V_k^{\max} \perp (b_k^{\text{sh}} - b_{k_{\text{espec}}}^{\text{sh}})$, gerados pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante o modelo (6.1.1).

A Figura 7.3.26 mostra os valores iterativos obtidos pelo método PDRNLI das susceptâncias *shunt* de barra até a convergência descrita na Tabela 7.3.35. Conforme as figuras mostram, a trajetória dos valores iterativos é bem-comportada, a susceptância obedece a seus limitante inferior e superior no término do processo de otimização e, em particular para problema descrito por meio do modelo (6.1.1), também satisfazem seus valores discretos.

Figura 7.3.26: Valores iterativos das susceptância shunt de barra obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante os modelos (6.1.7), (6.1.6) e (6.1.1), respectivamente.

A Tabela 7.3.37 mostra o valor final dos parâmetros de penalidades vinculados ao tratamento das variáveis discretas. No caso dos *taps* dos transformadores em-fase, o parâmetro de penalidade está

associado ao respectivo ramo $k - m$, o qual é conforme apresentado na Tabela 7.3.33. No caso das susceptâncias *shunt* de barra, o parâmetro de penalidade está associado à respectiva barra controlada, a qual é conforme apresentada na Tabela 7.3.35.

Tabela 7.3.37: Valores iterativos vinculados às componentes do vetor de parâmetros de penalidade ρ^k obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante o modelo (6.1.1).

Variáveis de Controle	ρ^k
Taps dos transformadores em-fase	54,876
	30114
	120,73
	120,73
	120,73
	120,73
	120,73
	265,6
	54,876
Susceptância <i>shunt</i> de barra	584,32
	0,8192
	0,8192
	6,5536
	0,4096
	0,4096
	0,8192
	26,214
	0,0256
	0,4096
	0,4096
	0,4096
	0,2048
	0,0512
	0,4096

Os valores iterativos associado às componentes do vetor de parâmetros de penalidade ρ^k do método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante o modelo (6.1.1) são apresentados Figura 7.3.27.

Figura 7.3.27: Valores iterativos vinculados às componentes do vetor de parâmetros de penalidade ρ^k obtidos pelo método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118 mediante o modelo (6.1.1).

O valor 30114 é muito grande em relação aos outros parâmetros. Então seu valor iterativo é apresentado de modo truncado na Figura 7.3.27.

Com a finalidade de mensurarmos o condicionamento numérico (na norma-2) da matriz θ_k ao longo do processo iterativo, os valores de condicionamento $\kappa(\theta_k)$ foram calculados antes e após a aplicação da estratégia de convergência global. Para isto, consideremos as abreviações ACQ para designar Antes da Condição de Positividade e DCP para designar Depois da Condição de Positividade. De acordo com os resultados apresentados na Tabela (7.3.38) a seguir, o número $\kappa(\theta_k)$ é significativamente melhorado quando os valores $\kappa(\theta_k)$ ACQ e DCQ são comparados entre si e respectivamente para todos os modelos de FPOR estudados para o sistema IEEE-118.

Tabela 7.3.38: Monitoramento do condicionamento numérico da matriz hessiana do método PDRNLI aplicado ao problema de FPOR junto ao sistema IEEE-118.

Ciclo k	Modelo (6.1.7)		Modelo (6.1.6)		Modelo (6.1.1)	
	$\kappa(\theta_k)$ ACP	$\kappa(\theta_k)$ DCP	$\kappa(\theta_k)$ ACP	$\kappa(\theta_k)$ DCP	$\kappa(\theta_k)$ ACP	$\kappa(\theta_k)$ DCP
0	1,29E+10	5,01E+05	1,29E+10	5,01E+05	1,29E+14	5,01E+08
1	2,32E+10	1,31E+06	2,33E+10	1,32E+06	2,33E+14	1,10E+10
2	3,85E+10	1,78E+06	3,84E+10	1,62E+06	1,01E+09	3,73E+09
3	4,14E+10	4,76E+05	4,09E+10	4,69E+05	2,67E+09	1,11E+08
4	4,55E+10	2,61E+05	4,53E+10	7,09E+05	4,51E+14	97670
5	1,10E+11	46424	6,44E+10	3,82E+05	8,34E+14	65176
6	1,15E+11	2,00E+05	1,39E+12	95069	4,25E+15	51332
7	2,60E+11	1,00E+06	1,40E+12	4,91E+05	1,33E+17	1,31E+09
8	1,38E+12	5,07E+06	1,48E+12	2,11E+06	1,16E+17	1,79E+10
9	1,41E+13	3,07E+07	4,53E+12	1,40E+07	1,16E+17	9,89E+09
10	1,40E+14	1,57E+08	3,18E+13	9,17E+07	1,48E+18	7,82E+09
11	1,71E+15	9,67E+08	3,48E+14	5,86E+08	2,04E+19	6,44E+11
12	1,93E+16	6,37E+10	2,85E+15	8,56E+07	1,30E+20	2,13E+12
13	2,11E+17	2,98E+10	2,98E+15	4,69E+08	5,44E+20	4,50E+12
14	3,01E+18	1,72E+11	3,56E+15	2,50E+09	3,59E+22	5,20E+12
15	3,76E+18	1,72E+11	7,01E+15	1,33E+10	6,36E+22	5,29E+12
16			3,34E+16	7,25E+10	3,07E+18	2,95E+12
17			2,99E+17	4,17E+11	6,04E+22	1,64E+14
18			3,23E+18	2,22E+12	1,51E+23	9,15E+14
19			4,82E+19	1,11E+11	8,78E+22	5,10E+15
20			4,27E+19	6,18E+11	1,31E+23	2,84E+16
21			1,99E+19	3,44E+12	5,98E+23	3,72E+17
22			5,67E+18	1,92E+13	8,33E+23	2,65E+18
23			1,92E+19	1,07E+14	3,54E+23	1,55E+19
24			1,60E+19	1,07E+14	3,64E+25	1,55E+19

A Tabela 7.3.39 sintetiza as variações proporcionais das perdas de potência na transmissão através dos três modelos. O cálculo é feito da seguinte forma: Sejam p as perdas obtidas pelo modelo A e q as perdas obtidas pelo modelo B; então os resultados de interesse são aqueles em $q < p$, pois a parte fracionária de $\frac{p}{q}$ representa o aumento proporcional das perdas obtidas, em por cento, pelo modelo A em relação às perdas obtidas pelo modelo B.

Tabela 7.3.39: Comparação entre as proporções das perdas de potência ativa associado ao sistema IEEE-118.

	A = Modelo (6.1.7)	A = Modelo (6.1.6)	A = (6.1.1)
B = Modelo (6.1.7)	1,000000	1,007774	1,007786
B = Modelo (6.1.6)	-	1,000000	1,000013
B = Modelo (6.1.1)	-	-	1,000000

Uma das principais conclusões obtidas com base na Tabela 7.3.39 é que o aumento proporcional das perdas, em megawatts, ao utilizarmos o modelo clássico (6.1.7) e o modelo (6.1.1) é de apenas

0,007786% com o ajuste de apenas 1 dispositivo de controle de tensão dentre os 23 possíveis. Este resultado mostra que o modelo (6.1.1) produz soluções que são operacionalmente mais simples de serem implementadas em tempo hábil pelo operador do sistema elétrico sem comprometer consideravelmente as perdas na transmissão.

Por fim, consideremos a hipótese de resolvermos o FPOR dado pelo modelo (6.1.1) por uma técnica de enumeração explícita associada às variáveis discretas; isto é, para cada possibilidade de escolha para os *taps* dos transformadores em-fase e para as susceptâncias *shunt* de barras, então resolvemos um FPOR com essas variáveis fixadas. Assim, depois de resolvermos todas as possibilidades, então escolhemos a melhor combinação. Verifiquemos quanto tempo será necessário para resolvermos todas as possibilidades ao assumirmos que cada FPOR é resolvido pelo método PDRNLI em 10^{-6} segundos.

O sistema IEEE-118 possui 9 *taps* dos transformadores, 12 susceptâncias *shunt* capacitivas de barra e 2 susceptâncias *shunt* indutivas de barra. Cada *tap* do transformador em-fase tem 33 possibilidades, cada susceptâncias *shunt* capacitivas de barra tem 3 possibilidades e cada susceptâncias *shunt* indutiva de barra tem duas possibilidades. Portanto, devemos resolver $33^9 \times 3^{12} \times 2^2 = 9,866 \times 10^{19}$ problemas de FPOR. O tempo total para executarmos todas essas possibilidades é $9,866 \times 10^{13}$ segundos ou 31284,84 séculos. Neste trabalho nós determinamos uma solução viável em apenas 3,88 segundos para uma única tentativa.

Conclusões

Neste trabalho, propomos uma abordagem de solução para resolver problemas de programação matemática discreta-mista com restrições de equilíbrio (MPEC). De uma forma geral, esta abordagem seguiu os seguintes procedimentos: i) inicialmente, transformamos o MPEC discreto-misto em um MPEC contínuo. ii) o MPEC contínuo foi transformado em um problema de programação não-linear (PNL) padrão equivalente e; iii) O PNL resultante foi abordado por meio do método primal-dual de reescalamiento não-linear integrado (PDRNLI) proposto.

No método PDRNLI, as direções de busca primal e dual são obtidas por meio do método de Newton com uma abordagem previsor-corretor e a atualização de todos parâmetros (penalidade, reescala e estimativa da variável dual das restrições de desigualdade) foi feita de forma dinâmica, isto é, após cada cálculo de direções de busca primal-dual ao fim do passo corretor.

Foi incorporado ao método PDRNLI a estratégia de convergência global denominada de condição de positividade, a qual não necessita que a matriz hessiana da função lagrangiana aumentada integrada associada à variável primal seja definida positiva em todo o $\mathbb{R}^n$, mas apenas na vizinhança do ponto corrente. Em outras palavras, esta estratégia de convergência global não exige que a matriz hessiana seja decomponível no sentido de Cholesky. Esta estratégia proposta é inspirada nas condições suficientes de KKT de segunda ordem e consiste em adicionar um número positivo adequado nos elementos diagonais da matriz hessiana avaliada no ponto corrente com o intuito de que forma quadrática envolvendo a matriz hessiana resultante desta operação e a direção primal seja um número positivo.

Um dos aspectos mais importantes do método PDRNLI é o tratamento das restrições de desigualdade. A concepção da função lagrangiana aumentada integrada surgiu da investigação de métodos de ponto proximal aplicados ao problema dual associado ao problema primal de programação não-linear convexo envolvendo somente restrições de desigualdade. Neste trabalho, investigamos dois métodos duais de ponto proximal, a saber: os métodos descritos em [Polyak e Teboulle 1997, eq. 3.8] e [Matioli e Gonzaga 2008, alg. 2.2], os quais respectivamente utilizam as distâncias entrópicas φ -divergente de Csiszar e de Bregman junto ao método dual de ponto proximal aplicado ao problema dual. Desta investigação, foi apresentado o conceito de uma medida de distância integrada. E este conceito de medida de distância foi utilizado como uma abordagem para o método dual de ponto proximal. Demonstramos que o método dual de ponto proximal com a medida de distância integrada é equivalente a um nova classe de métodos de lagrangiano aumentado, na qual o conceito de função

lagrangiana aumentada integrada desempenha papel fundamental. Destacamos que as abordagens apresentadas em [Polyak e Teboulle 1997] e [Matioli e Gonzaga 2008] são casos particulares de nossa proposta.

Uma das deficiências da função barreira logarítmica modificada é que esta função não é definida para todos os números reais. Algumas técnicas de extrapolação foram propostas para contornar essa deficiência, tais como extrapolação quadrática, exponencial e cúbica [Polyak e Teboulle 1997, Pınheiro et al. 2015]. Em [Auslender, Teboulle e Ben-Tiba 1999], os autores propõem uma classe de funções denominadas barreira logarítmica-quadrática e a utilizaram para induzir uma nova medida de distância. A função logarítmica-quadrática é resultado de uma regularização sobre a função conjugada da função logarítmica modificada. Um dos principais efeitos desta regularização é que o conjugado convexo da função logarítmica-quadrática permite obter funções que não carecem de técnica de extrapolação. Neste trabalho, utilizamos a função logarítmica-quadrática para induzir uma nova e particular medida de distância integrada substancialmente diferente da medida de Auslender. A função conjugada da medida de distância integrada induzida pela função logarítmica-quadrática nos conduziu naturalmente a uma abordagem particular para o método PDRNLI.

Outro aspecto relevante do método PDRNLI proposto é o tratamento das variáveis discretas, que compreendeu os seguintes procedimentos: i) dado um conjunto discreto associado a uma variável, construímos uma classe conveniente de funções cujos *zeros* coincidem com os valores discretos estabelecidos por este conjunto. Dentre tais funções, destacamos o conceito da função interpolante hiperbólica definida por partes (função-IHDP), a qual limita o fenômeno de Runge no valor $\frac{1}{2}$ para qualquer conjunto discreto; ii) criamos um problema de programação não-linear em que funções da classe descrita em (i) são incorporadas como restrições de igualdade artificiais ao passo que as variáveis discretas são tratadas como contínuas e restritas a um intervalo contínuo (restrições canalizadas artificiais); iii) tratamos as restrições de igualdade artificiais descritas em (ii) pelo método de funções de penalidade e as restrições canalizadas pelo método de lagrangiano aumentado integrado; iv) propusemos novas funções penalidades de natureza não-quadrática que possuem propriedades específicas para lidarmos com as restrições de igualdade artificiais descritas em (ii). Dentre essas funções, destacamos as funções- q_{inf} e y -logarítmica. Mostramos que as funções penalidades senoidal descritas em [Soler, Asada e Costa 2013] e [Lage 2013] são casos particulares das funções penalidades propostas.

Aplicamos o método PDRNLI para resolvermos três diferentes modelos vinculados ao problema de Fluxo de Potência Ótimo Reativo (FPOR): i) FPOR clássico cujas variáveis de controle são consideradas contínuas; ii) FPOR com restrições de atuação de dispositivos de controle de tensão cujas variáveis de controle são consideradas contínuas e; iii) FPOR com restrições de atuação de dispositivos de controle de tensão cujas variáveis de controle associadas aos *tap* dos transformadores em-fase e às susceptâncias *shunt* de barra são consideradas discretas. A abordagem desses três problemas de FPOR junto aos sistemas elétricos IEEE-14, IEEE-30 e IEEE-118 com o método PDRNLI produziu resultados numéricos satisfatórios e robustos. Os resultados nos encorajam a comparar o desempenho do método PDRNLI com outros métodos desenvolvidos para resolver problemas de FPOR e outros MPECs apresentados na literatura.

Uma questão importante sobre a solução de problemas FPOR com restrição de atuação de dispositivos de controle de tensão diz respeito a otimalidade destas. Conforme apresentado no Capítulo 4,

sempre que uma solução estacionária primal e dual de um MPEC é determinada mediante a solução do PNL padrão equivalente utilizado neste trabalho, então, tal ponto de KKT de primeira ordem é equivalente a condição S-estacionariedade associada ao MPEC. Entretanto, esta equivalência não é assegurada quando se investigam as condições de segunda ordem [Guo, Lin e Ye 2012]. A saber: a condição suficiente e necessária de KKT do PNL equivalente nos assegura que tal ponto é pelo menos um minimizador não estrito para o MPEC contínuo. Com a inserção da estratégia de condição de positividade junto ao método PDRNLI, então os pontos primais e estacionários são pelo menos minimizadores locais não-estritos do problema FPOR com restrições de atuação de dispositivos de controle de tensão. Para as soluções obtidas para o MPEC discreto-misto podemos afirmar somente que estas soluções primais são minimizadores para a sequência dos problemas de MPEC contínuos. Porém, conjecturamos que tais pontos estacionários são, de fato, minimizadores do MPEC discreto-misto original.

Referências

- AUSLENDER, A.; TEBOULLE, M.; BEN-TIBA, S. Interior Proximal and Multiplier Methods Based on Second Order Homogeneous Kernels. **Mathematics of Operations Research**, 1999. v. 24, n. 3, p. 645–668, ago. 1999. ISSN 0364-765X. Disponível em: <<http://pubsonline.informs.org/doi/abs/10.1287/moor.24.3.645>>. Citado 3 vezes nas páginas 34, 39 e 51.
- _____. A Logarithmic-Quadratic Proximal Method for Variational Inequalities. In: PANG, J.-S. (Ed.). **Computational Optimization**. Springer US, 1999. p. 31–40. ISBN 978-1-4613-7367-4 978-1-4615-5197-3. DOI: 10.1007/978-1-4615-5197-3_3. Disponível em: <http://link.springer.com/chapter/10.1007/978-1-4615-5197-3_3>. Citado 4 vezes nas páginas 35, 39, 51 e 192.
- BAZARAA, M. S.; SHERALI, H. D.; SHETTY, C. M. **Nonlinear Programming: Theory and Algorithms**. 3 edition. ed. Hoboken, N.J.: Wiley-Interscience, 2006. ISBN 978-0-471-48600-8. Citado 7 vezes nas páginas 30, 41, 44, 62, 63, 68 e 99.
- BENSON, H. Y.; SHANNO, D. F.; VANDERBEI, R. J. **Interior-Point methods for nonconvex nonlinear programming: Complementary Constraints**. [S.l.], ago. 2002. Citado 4 vezes nas páginas 30, 71, 77 e 78.
- BERTSEKAS, D. P. **Constrained Optimization and Lagrange Multiplier Methods**. 1 edition. ed. [S.l.]: Athena Scientific, 1996. ISBN 978-1-886529-04-5. Citado 2 vezes nas páginas 55 e 63.
- _____. **Nonlinear Programming**. 2nd edition. ed. Belmont, Mass: Athena Scientific, 1999. ISBN 978-1-886529-00-7. Citado 3 vezes nas páginas 28, 41 e 44.
- BIRGIN, E. G.; CASTILLO, R. A.; MARTÍNEZ, J. M. Numerical Comparison of Augmented Lagrangian Algorithms for Nonconvex Problems. **Computational Optimization and Applications**, 2005. v. 31, n. 1, p. 31–55, maio 2005. ISSN 0926-6003, 1573-2894. Disponível em: <<http://link.springer.com/article/10.1007/s10589-005-1066-7>>. Citado na página 55.
- BREGMAN, L. M. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. **USSR Computational Mathematics and Mathematical Physics**, 1967. v. 7, n. 3, p. 200–217, 1967. ISSN 0041-5553. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0041555367900407>>. Citado 2 vezes nas páginas 31 e 35.
- CARMO, M. P. d. **Geometria Riemanniana**. Rio de Janeiro: LTC, 2008. ISBN 978-85-244-0036-0. Citado na página 215.

- CENSOR, Y.; ZENIOS, S. A. Proximal minimization algorithm with D-functions. **Journal of Optimization Theory and Applications**, 1992. v. 73, n. 3, p. 451–464, 1992. ISSN 0022-3239, 1573-2878. Disponível em: <<http://link.springer.com/article/10.1007/BF00940051>>. Citado na página 35.
- CHEN, B.; HARKER, P. A Non-Interior-Point Continuation Method for Linear Complementarity Problems. **SIAM Journal on Matrix Analysis and Applications**, 1993. v. 14, n. 4, p. 1168–1190, out. 1993. ISSN 0895-4798. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/0614081>>. Citado na página 71.
- _____. Smooth Approximations to Nonlinear Complementarity Problems. **SIAM Journal on Optimization**, 1997. v. 7, n. 2, p. 403–420, maio 1997. ISSN 1052-6234. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/S1052623495280615>>. Citado na página 71.
- CHEN, C.; MANGASARIAN, O. L. A class of smoothing functions for nonlinear and mixed complementarity problems. **Computational Optimization and Applications**, 1996. v. 5, n. 2, p. 97–138, mar. 1996. ISSN 0926-6003, 1573-2894. Disponível em: <<http://link.springer.com/article/10.1007/BF00249052>>. Citado 2 vezes nas páginas 71 e 78.
- CHEN, Y.; FLORIAN, M. The nonlinear bilevel programming problem: formulations, regularity and optimality conditions. **Optimization**, 1995. v. 32, n. 3, p. 193–209, jan. 1995. ISSN 0233-1934. Disponível em: <<http://dx.doi.org/10.1080/02331939508844048>>. Citado na página 72.
- CHUQUIPOMA, J. A. D. Inequações Variacionais: Teoria e Aplicações. In: . **ERMAC 2010: I Encontro Regional de Matemática Aplicada e Computacional**: [s.n.], 2010. p. 36 – 64. Citado na página 76.
- CSISZÁR, I. Information-type measures of difference of probability distributions and indirect observations. **Studia Scientiarum Mathematicarum Hungarica**, 1967. v. 2, p. 299–318, 1967. Citado 3 vezes nas páginas 31, 35 e 37.
- DIRKSE, S. P.; FERRIS, M. C. The path solver: a nonmonotone stabilization scheme for mixed complementarity problems. **Optimization Methods and Software**, 1995. v. 5, n. 2, p. 123–156, jan. 1995. ISSN 1055-6788. Disponível em: <<http://dx.doi.org/10.1080/10556789508805606>>. Citado na página 78.
- DOMINGUES, H. H. **Álgebra Linear e Aplicações**. Edição: 6. São Paulo: Atual Paradidático, 2009. ISBN 978-85-7056-297-5. Citado 2 vezes nas páginas 222 e 225.
- DOMMEL, H.; TINNEY, W. Optimal Power Flow Solutions. **IEEE Transactions on Power Apparatus and Systems**, 1968. PAS-87, n. 10, p. 1866–1876, out. 1968. ISSN 0018-9510. Citado 2 vezes nas páginas 129 e 134.
- FERRIS, M.; PANG, J. Engineering and Economic Applications of Complementarity Problems. **SIAM Review**, 1997. v. 39, n. 4, p. 669–713, jan. 1997. ISSN 0036-1445. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/S0036144595285963>>. Citado 3 vezes nas páginas 69, 71 e 72.
- FISCHER, A. A special newton-type optimization method. **Optimization**, 1992. v. 24, n. 3-4, p. 269–284, jan. 1992. ISSN 0233-1934. Disponível em: <<http://dx.doi.org/10.1080/02331939208843795>>. Citado na página 30.
- FLEGEL, M. L. Constraint qualifications and stationarity concepts for mathematical programs with equilibrium constraints, Regularitätsbedingungen und Stationaritätskonzepte für Mathematische Programme mit Gleichgewichtsnebenbedingungen. 2005. 2005. Disponível em: <<https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/1068>>. Citado 2 vezes nas páginas 73 e 179.

FLEGEL, M. L.; KANZOW, C. Abadie-Type Constraint Qualification for Mathematical Programs with Equilibrium Constraints. **Journal of Optimization Theory and Applications**, 2005. v. 124, n. 3, p. 595–614, mar. 2005. ISSN 0022-3239, 1573-2878. Disponível em: <<http://link.springer.com/article/10.1007/s10957-004-1176-x>>. Citado 3 vezes nas páginas 29, 73 e 75.

_____. On M-stationary points for mathematical programs with equilibrium constraints. **Journal of Mathematical Analysis and Applications**, 2005. v. 310, n. 1, p. 286–302, out. 2005. ISSN 0022-247X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0022247X05001101>>. Citado 2 vezes nas páginas 73 e 76.

_____. On the Guignard constraint qualification for mathematical programs with equilibrium constraints. **Optimization**, 2005. v. 54, n. 6, p. 517–534, dez. 2005. ISSN 0233-1934. Disponível em: <<http://dx.doi.org/10.1080/02331930500342591>>. Citado 5 vezes nas páginas 70, 72, 73, 179 e 180.

FLETCHER, R. et al. Local Convergence of SQP Methods for Mathematical Programs with Equilibrium Constraints. **SIAM Journal on Optimization**, 2006. v. 17, n. 1, p. 259–286, jan. 2006. ISSN 1052-6234. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/S1052623402407382>>. Citado 3 vezes nas páginas 72, 179 e 180.

FRANCO, N. M. B. **Cálculo Numérico**. São Paulo: Pearson, 2007. ISBN 978-85-7605-087-2. Citado 2 vezes nas páginas 57 e 98.

GAUVIN, J. A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming. **Mathematical Programming**, 1977. v. 12, n. 1, p. 136–138, dez. 1977. ISSN 0025-5610, 1436-4646. Disponível em: <<http://link.springer.com/article/10.1007/BF01593777>>. Citado na página 41.

GOULD, F. J.; TOLLE, J. W. A Necessary and Sufficient Qualification for Constrained Optimization. **SIAM Journal on Applied Mathematics**, 1971. v. 20, n. 2, p. 164–172, mar. 1971. ISSN 0036-1399. Disponível em: <<http://www.jstor.org/stable/2099916>>. Citado na página 72.

GRANVILLE, S. Optimal reactive dispatch through interior point methods. **IEEE Transactions on Power Systems**, 1994. v. 9, n. 1, p. 136–146, fev. 1994. ISSN 0885-8950. Citado 2 vezes nas páginas 103 e 134.

GREGÓRIO, R.; OLIVEIRA, P. R. A logarithmic-quadratic proximal point scalarization method for multiobjective programming. **Journal of Global Optimization**, 2010. v. 49, n. 2, p. 281–291, abr. 2010. ISSN 0925-5001, 1573-2916. Disponível em: <<http://link.springer.com/article/10.1007/s10898-010-9544-6>>. Citado na página 51.

GRIVA, I.; NASH, S. G.; SOFER, A. **Linear and Nonlinear Optimization, Second Edition**. 2 edition. ed. Philadelphia: Society for Industrial Mathematics, 2008. ISBN 978-0-89871-661-0. Citado na página 62.

GRIVA, I.; POLYAK, R. A. Primal-dual nonlinear rescaling method with dynamic scaling parameter update. **Mathematical Programming**, 2006. v. 106, n. 2, p. 237–259, 2006. ISSN 0025-5610, 1436-4646. Disponível em: <<http://link.springer.com/article/10.1007/s10107-005-0603-6>>. Citado na página 50.

GUIGNARD, M. Generalized Kuhn-Tucker Conditions for Mathematical Programming Problems in a Banach Space. **SIAM Journal on Control**, 1969. v. 7, n. 2, p. 232–241, maio 1969. ISSN 0036-1402. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/0307016>>. Citado na página 72.

GUO, L. et al. Sensitivity Analysis of the Value Function for Parametric Mathematical Programs with Equilibrium Constraints. **SIAM Journal on Optimization**, 2014. v. 24, n. 3, p. 1206–1237, jan. 2014. ISSN 1052-6234. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/130929783>>. Citado na página 72.

GUO, L.; LIN, G.-H. Notes on Some Constraint Qualifications for Mathematical Programs with Equilibrium Constraints. **Journal of Optimization Theory and Applications**, 2012. v. 156, n. 3, p. 600–616, maio 2012. ISSN 0022-3239, 1573-2878. Disponível em: <<http://link.springer.com/article/10.1007/s10957-012-0084-8>>. Citado na página 73.

GUO, L.; LIN, G.-H.; YE, J. J. Second-Order Optimality Conditions for Mathematical Programs with Equilibrium Constraints. **Journal of Optimization Theory and Applications**, 2012. v. 158, n. 1, p. 33–64, dez. 2012. ISSN 0022-3239, 1573-2878. Disponível em: <<http://link.springer.com/article/10.1007/s10957-012-0228-x>>. Citado 4 vezes nas páginas 72, 75, 98 e 193.

_____. Solving Mathematical Programs with Equilibrium Constraints. **Journal of Optimization Theory and Applications**, 2015. v. 166, n. 1, p. 234–256, jan. 2015. ISSN 0022-3239, 1573-2878. Disponível em: <<http://link.springer.com/article/10.1007/s10957-014-0699-z>>. Citado 4 vezes nas páginas 31, 72, 75 e 76.

HATCHER, A. **Algebraic Topology**. 1 edition. ed. Cambridge ; New York: Cambridge University Press, 2001. ISBN 978-0-521-79540-1. Citado na página 39.

HU, X. M.; RALPH, D. Convergence of a Penalty Method for Mathematical Programming with Complementarity Constraints. **Journal of Optimization Theory and Applications**, 2004. v. 123, n. 2, p. 365–390, nov. 2004. ISSN 0022-3239, 1573-2878. Disponível em: <<http://link.springer.com/article/10.1007/s10957-004-5154-0>>. Citado na página 72.

IUSEM, A. N.; SVAITER, B. F.; TEBOULLE, M. Entropy-Like Proximal Methods in Convex Programming. **Mathematics of Operations Research**, 1994. v. 19, n. 4, p. 790–814, nov. 1994. ISSN 0364-765X. Disponível em: <<http://pubsonline.informs.org/doi/abs/10.1287/moor.19.4.790>>. Citado 5 vezes nas páginas 34, 35, 36, 37 e 42.

JIANG, H.; RALPH, D. Smooth SQP Methods for Mathematical Programs with Nonlinear Complementarity Constraints. **SIAM Journal on Optimization**, 2000. v. 10, n. 3, p. 779–808, jan. 2000. ISSN 1052-6234. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/S1052623497332329>>. Citado na página 72.

JIANG, Y. et al. An augmented Lagrangian multiplier method based on a CHKS smoothing function for solving nonlinear bilevel programming problems. **Knowledge-Based Systems**, 2014. v. 55, p. 9–14, jan. 2014. ISSN 0950-7051. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0950705113002530>>. Citado na página 72.

KADRANI, A.; DUSSAULT, J. P.; BENCHAKROUN, A. A globally convergent algorithm for MPCC. **EURO Journal on Computational Optimization**, 2015. v. 3, n. 3, p. 263–296, set. 2015. ISSN 2192-4406, 2192-4414. Disponível em: <<http://link.springer.com/article/10.1007/s13675-015-0044-9>>. Citado 2 vezes nas páginas 179 e 180.

KANNIAPPAN, P. Fenchel-Rockafellar type duality for a non-convex non-differential optimization problem. **Journal of Mathematical Analysis and Applications**, 1983. v. 97, n. 1, p. 266–276, nov. 1983. ISSN 0022-247X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0022247X83902500>>. Citado na página 44.

- KOJIMA, M. et al. **A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems**. Berlin: Springer-Verlag, 1991. ISBN 978-0-387-54509-7. Citado na página 71.
- KORT, B.; BERTSEKAS, D. Multiplier methods for convex programming. In: **1973 IEEE Conference on Decision and Control including the 12th Symposium on Adaptive Processes**. [S.l.: s.n.], 1973. p. 428–432. Citado na página 46.
- LAGE, G. G. **O fluxo de potência ótimo reativo com variáveis de controle discretas e restrições de atuação de dispositivos de controle de tensão**. Tese (text) — Universidade de São Paulo, mar. 2013. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/18/18154/tde-29042013-114259/>>. Citado 12 vezes nas páginas 27, 28, 30, 55, 57, 83, 85, 88, 106, 130, 134 e 192.
- LAM, T. Y. **Introduction To Quadratic Forms Over Fields**. Providence, R.I: Amer Mathematical Society, 2004. ISBN 978-0-8218-1095-8. Citado 2 vezes nas páginas 96 e 215.
- LAVAEI, J.; LOW, S. H. Zero Duality Gap in Optimal Power Flow Problem. **IEEE Transactions on Power Systems**, 2012. v. 27, n. 1, p. 92–107, fev. 2012. ISSN 0885-8950. Citado na página 134.
- LEVENBERG, K. A method for the solution of certain problems in least squares. **Quarterly of Applied Mathematics**, 1944. v. 2, p. 164–168, 1944. Citado na página 68.
- LEYFFER, S.; LÓPEZ-CALVA, G.; NOCEDAL, J. Interior Methods for Mathematical Programs with Complementarity Constraints. **SIAM Journal on Optimization**, 2006. v. 17, n. 1, p. 52–77, jan. 2006. ISSN 1052-6234. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/040621065>>. Citado 2 vezes nas páginas 179 e 180.
- LIU, X.; SUN, J. Generalized stationary points and an interior-point method for mathematical programs with equilibrium constraints. **Mathematical Programming**, 2004. v. 101, n. 1, p. 231–261, jul. 2004. ISSN 0025-5610, 1436-4646. Disponível em: <<http://link.springer.com/article/10.1007/s10107-004-0543-6>>. Citado 5 vezes nas páginas 30, 72, 75, 179 e 180.
- LUCIDI, S.; RINALDI, F. Exact Penalty Functions for Nonlinear Integer Programming Problems. **Journal of Optimization Theory and Applications**, 2010. v. 145, n. 3, p. 479–488, abr. 2010. ISSN 0022-3239, 1573-2878. Disponível em: <<http://link.springer.com/article/10.1007/s10957-010-9700-7>>. Citado na página 55.
- LUO, Z.-Q.; PANG, J.-S.; RALPH, D. **Mathematical Programs with Equilibrium Constraints**. 1 edition. ed. Cambridge ; New York: Cambridge University Press, 1996. ISBN 978-0-521-57290-3. Citado 3 vezes nas páginas 27, 69 e 70.
- LUSTIG, I.; MARSTEN, R.; SHANNO, D. On Implementing Mehrotra's Predictor-Corrector Interior-Point Method for Linear Programming. **SIAM Journal on Optimization**, 1992. v. 2, n. 3, p. 435–449, ago. 1992. ISSN 1052-6234. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/0802022>>. Citado na página 95.
- LUSTIG, I. J.; MARSTEN, R. E.; SHANNO, D. F. Computational experience with a globally convergent primal-dual predictor-corrector algorithm for linear programming. **Mathematical Programming**, 1994. v. 66, n. 1-3, p. 123–135, ago. 1994. ISSN 0025-5610, 1436-4646. Disponível em: <<http://link.springer.com/article/10.1007/BF01581140>>. Citado na página 95.
- MADANI, R.; ASHRAPHIJUO, M.; LAVAEI, J. Promises of Conic Relaxation for Contingency-Constrained Optimal Power Flow Problem. **IEEE Transactions on Power Systems**, 2016. v. 31, n. 2, p. 1297–1307, mar. 2016. ISSN 0885-8950. Citado na página 134.

MARQUARDT, D. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. **Journal of the Society for Industrial and Applied Mathematics**, 1963. v. 11, n. 2, p. 431–441, jun. 1963. ISSN 0368-4245. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/0111030>>. Citado na página 68.

MARTINET, B. Perturbation des méthodes d'optimisation. Applications. **ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique**, 1978. v. 12, n. 2, p. 153–171, 1978. ISSN 0764-583X. Disponível em: <<https://eudml.org/doc/193317>>. Citado na página 34.

MARTÍNEZ, J. M. **Otimização Prática Usando o Lagrangiano Aumentado**. [S.l.]: Departamento de Matemática Aplicada, IMECC-UNICAMP, 2009. Citado 4 vezes nas páginas 62, 67, 103 e 138.

MARTÍNEZ, J. M.; SANTOS, S. A. **Métodos Computacionais de Otimização**. [S.l.]: Departamento de Matemática Aplicada, IMECC-UNICAMP, 1998. Citado na página 62.

MATIOLI, L. C.; GONZAGA, C. C. A new family of penalties for augmented Lagrangian methods. **Numerical Linear Algebra with Applications**, 2008. v. 15, n. 10, p. 925–944, dez. 2008. ISSN 1099-1506. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1002/nla.596/abstract>>. Citado 17 vezes nas páginas 30, 31, 33, 34, 35, 36, 38, 40, 41, 42, 43, 45, 46, 48, 50, 191 e 192.

MEHROTRA, S. On the Implementation of a Primal-Dual Interior Point Method. **SIAM Journal on Optimization**, 1992. v. 2, n. 4, p. 575–601, nov. 1992. ISSN 1052-6234. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/0802028>>. Citado na página 95.

MEYER, C. D. **Matrix Analysis and Applied Linear Algebra Book and Solutions Manual**. Har/cdr edition. Philadelphia: SIAM: Society for Industrial and Applied Mathematics, 2001. ISBN 978-0-89871-454-8. Citado 2 vezes nas páginas 223 e 224.

MOORE, E. On the reciprocal of the general algebraic matrix. **Bulletin of the American Mathematical Society**, 1920. v. 26, p. 394–395, 1920. Citado na página 219.

NOWAK, I. **Relaxation and Decomposition Methods for Mixed Integer Nonlinear Programming**. 2005 edition. ed. Basel: Birkhäuser, 2005. ISBN 978-3-7643-7238-5. Citado na página 29.

O'NEILL, B. **Semi-Riemannian Geometry With Applications to Relativity, 103, Volume 103**. 1 edition. ed. New York: Academic Press, 1983. ISBN 978-0-12-526740-3. Citado na página 215.

PENROSE, R. A generalized inverse for matrices. **Mathematical Proceedings of the Cambridge Philosophical Society**, 1955. v. 51, n. 03, p. 406–413, jul. 1955. ISSN 1469-8064. Disponível em: <http://journals.cambridge.org/article_S0305004100030401>. Citado na página 219.

PIERRO, A. R. D.; IUSEM, A. N. A relaxed version of Bregman's method for convex programming. **Journal of Optimization Theory and Applications**, 1986. v. 51, n. 3, p. 421–440, dez. 1986. ISSN 0022-3239, 1573-2878. Disponível em: <<http://link.springer.com/article/10.1007/BF00940283>>. Citado na página 35.

PINHEIRO, R. B. N. et al. Interior - exterior point method with global convergence strategy for solving the reactive optimal power flow problem. **International Journal of Electrical Power & Energy Systems**, 2015. v. 66, p. 235–246, mar. 2015. ISSN 0142-0615. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0142061514006061>>. Citado 8 vezes nas páginas 51, 90, 95, 98, 102, 103, 134 e 192.

PLATBROOD, L. et al. A Generic Approach for Solving Nonlinear-Discrete Security-Constrained Optimal Power Flow Problems in Large-Scale Systems. **IEEE Transactions on Power Systems**, 2014. v. 29, n. 3, p. 1194–1203, maio 2014. ISSN 0885-8950. Citado 2 vezes nas páginas 27 e 129.

POLYAK, R. Modified barrier functions (theory and methods). **Mathematical Programming**, 1992. v. 54, n. 1-3, p. 177–222, fev. 1992. ISSN 0025-5610, 1436-4646. Disponível em: <<http://link.springer.com/article/10.1007/BF01586050>>. Citado 2 vezes nas páginas 40 e 46.

_____. Lagrangian Transformation and Interior Ellipsoid Methods in Convex Optimization. **Journal of Optimization Theory and Applications**, 2014. v. 164, n. 3, p. 966–992, jan. 2014. ISSN 0022-3239, 1573-2878. Disponível em: <<http://link.springer.com/article/10.1007/s10957-014-0527-5>>. Citado 3 vezes nas páginas 33, 38 e 50.

POLYAK, R.; TEBOULLE, M. Nonlinear rescaling and proximal-like methods in convex optimization. **Mathematical Programming**, 1997. v. 76, n. 2, p. 265–284, fev. 1997. ISSN 0025-5610, 1436-4646. Disponível em: <<http://link.springer.com/article/10.1007/BF02614440>>. Citado 13 vezes nas páginas 30, 31, 33, 34, 40, 42, 43, 45, 50, 51, 53, 191 e 192.

POLYAK, R. A. Nonlinear Rescaling as Interior Quadratic Prox Method in Convex Optimization. **Computational Optimization and Applications**, 2006. v. 35, n. 3, p. 347–373, ago. 2006. ISSN 0926-6003, 1573-2894. Disponível em: <<http://link.springer.com/article/10.1007/s10589-006-9759-0>>. Citado na página 51.

POOLE, D. **Linear Algebra: A Modern Introduction**. 3 edition. ed. Boston, MA: Cengage Learning, 2010. ISBN 978-0-538-73545-2. Citado 5 vezes nas páginas 205, 206, 209, 210 e 223.

Power System Case: <https://www.ee.washington.edu/research/pstca/>. Citado 4 vezes nas páginas 16, 136, 137 e 138.

RAO, S. S. **Engineering Optimization: Theory and Practice**. 4 edition. ed. Hoboken, NJ: Wiley, 2009. ISBN 978-0-470-18352-6. Citado 2 vezes nas páginas 29 e 55.

RIBEIRO, A. A.; KARAS, E. **Otimização Continua. Aspectos Teóricos E Computacionais**. [S.l.]: Cengage, 2013. ISBN 978-85-221-1501-3. Citado 2 vezes nas páginas 28 e 73.

RINALDI, F. New results on the equivalence between zero-one programming and continuous concave programming. **Optimization Letters**, 2009. v. 3, n. 3, p. 377–386, mar. 2009. ISSN 1862-4472, 1862-4480. Disponível em: <<http://link.springer.com/article/10.1007/s11590-009-0117-x>>. Citado na página 55.

ROCHA, R. A. et al. Logarithmic quasi-distance proximal point scalarization method for multi-objective programming. **Applied Mathematics and Computation**, 2016. v. 273, p. 856–867, jan. 2016. ISSN 0096-3003. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0096300315014162>>. Citado na página 51.

ROCKAFELLAR, R. Monotone Operators and the Proximal Point Algorithm. **SIAM Journal on Control and Optimization**, 1976. v. 14, n. 5, p. 877–898, ago. 1976. ISSN 0363-0129. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/0314056>>. Citado 3 vezes nas páginas 33, 34 e 41.

ROCKAFELLAR, R. T. A dual approach to solving nonlinear programming problems by unconstrained optimization. **Mathematical Programming**, 1973. v. 5, n. 1, p. 354–373, dez. 1973. ISSN 0025-5610, 1436-4646. Disponível em: <<http://link.springer.com/article/10.1007/BF01580138>>. Citado 4 vezes nas páginas 33, 34, 41 e 43.

_____. Augmented Lagrangians and Applications of the Proximal Point Algorithm in Convex Programming. **Mathematics of Operations Research**, 1976. v. 1, n. 2, p. 97–116, maio 1976. ISSN 0364-765X. Disponível em: <<http://pubsonline.informs.org.ez87.periodicos.capes.gov.br/doi/abs/10.1287/moor.1.2.97>>. Citado na página 41.

ROSEHART, W.; ROMAN, C.; SCHELLENBERG, A. Optimal power flow with complementarity constraints. **IEEE Transactions on Power Systems**, 2005. v. 20, n. 2, p. 813–822, maio 2005. ISSN 0885-8950. Citado 4 vezes nas páginas 27, 70, 78 e 129.

RUGGIERO, M. A. G.; LOPES, V. L. d. R. **Cálculo Numérico. Aspectos Teóricos e Computacionais**. Ciencias exatas edition. São Paulo: PEARSON - UNIVERSITARIOS/KOTLER, 1996. ISBN 978-85-346-0204-4. Citado na página 57.

SANTIS, M. D.; RINALDI, F. Continuous Reformulations for Zero-One Programming Problems. **Journal of Optimization Theory and Applications**, 2011. v. 153, n. 1, p. 75–84, out. 2011. ISSN 0022-3239, 1573-2878. Disponível em: <<http://link.springer.com/article/10.1007/s10957-011-9935-y>>. Citado na página 55.

SCHEEL, H.; SCHOLTES, S. Mathematical Programs with Complementarity Constraints: Stationarity, Optimality, and Sensitivity. **Mathematics of Operations Research**, 2000. v. 25, n. 1, p. 1–22, fev. 2000. ISSN 0364-765X. Disponível em: <<http://pubsonline.informs.org/doi/abs/10.1287/moor.25.1.1.15213>>. Citado na página 72.

SHAFAREVICH, I. R.; REMIZOV, A. **Linear Algebra and Geometry**. 2013 edition. ed. Berlin ; New York: Springer, 2012. ISBN 978-3-642-30993-9. Citado 2 vezes nas páginas 227 e 228.

SHIN, D. K.; GURDAL, Z.; JR, O. H. G. A Penalty Approach for Nonlinear Optimization with Discrete Design Variables. In: ESCHENAUER, H. A.; THIERAUF, G. (Ed.). **Discretization Methods and Structural Optimization à Procedures and Applications**. Springer Berlin Heidelberg, 1989, (Lecture Notes in Engineering, 42). p. 326–334. ISBN 978-3-540-50784-0 978-3-642-83707-4. Disponível em: <http://link.springer.com/chapter/10.1007/978-3-642-83707-4_2>. Citado na página 29.

SOLER, E.; ASADA, E.; COSTA, G. da. Penalty-Based Nonlinear Solver for Optimal Reactive Power Dispatch With Discrete Controls. **IEEE Transactions on Power Systems**, 2013. v. 28, n. 3, p. 2174–2182, ago. 2013. ISSN 0885-8950. Citado 5 vezes nas páginas 28, 29, 55, 65 e 192.

SOUSA, V. A.; BAPTISTA, E. C.; COSTA, G. R. M. da. Optimal reactive power flow via the modified barrier Lagrangian function approach. **Electric Power Systems Research**, 2012. v. 84, n. 1, p. 159–164, mar. 2012. ISSN 0378-7796. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378779611002677>>. Citado na página 134.

SOUSA, V. A. de; BAPTISTA, E. C.; COSTA, G. R. M. da. Loss minimization by the predictor-corrector modified barrier approach. **Electric Power Systems Research**, 2009. v. 79, n. 5, p. 803–808, maio 2009. ISSN 0378-7796. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378779608002915>>. Citado na página 95.

TEBOULLE, M. Entropic Proximal Mappings with Applications to Nonlinear Programming. **Mathematics of Operations Research**, 1992. v. 17, n. 3, p. 670–690, ago. 1992. ISSN 0364-765X. Disponível em: <<http://pubsonline.informs.org/doi/abs/10.1287/moor.17.3.670>>. Citado 7 vezes nas páginas 33, 34, 35, 37, 38, 41 e 42.

- _____. Convergence of Proximal-Like Algorithms. **SIAM Journal on Optimization**, 1997. v. 7, n. 4, p. 1069–1083, nov. 1997. ISSN 1052-6234. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/S1052623495292130>>. Citado na página 37.
- TORRES, G.; QUINTANA, V. Optimal power flow by a nonlinear complementarity method. In: **Power Industry Computer Applications, 1999. PICA '99. Proceedings of the 21st 1999 IEEE International Conference**. [S.l.: s.n.], 1999. p. 211–216. Citado na página 134.
- WU, Y.-C.; DEBS, A. S.; MARSTEN, R. E. A Nonlinear Programming Approach Based on an Interior Point Method for Optimal Power Flows. In: **Athens Power Tech, 1993. APT 93. Proceedings. Joint International Power Conference**. [S.l.: s.n.], 1993. v. 1, p. 196–200. Citado na página 95.
- _____. A direct nonlinear predictor-corrector primal-dual interior point algorithm for optimal power flows. **IEEE Transactions on Power Systems**, 1994. v. 9, n. 2, p. 876–883, maio 1994. ISSN 0885-8950. Citado na página 134.
- YANG, Y.-F.; QI, L. Smoothing Trust Region Methods for Nonlinear Complementarity Problems with P 0-Functions. **Annals of Operations Research**, 2005. v. 133, n. 1–4, p. 99–117, jan. 2005. ISSN 0254-5330, 1572-9338. Disponível em: <<http://link.springer.com/article/10.1007/s10479-004-5026-x>>. Citado na página 72.
- YE, J.; ZHU, D.; ZHU, Q. Exact Penalization and Necessary Optimality Conditions for Generalized Bilevel Programming Problems. **SIAM Journal on Optimization**, 1997. v. 7, n. 2, p. 481–507, maio 1997. ISSN 1052-6234. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/S1052623493257344>>. Citado na página 72.
- YE, J. J. Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. **Journal of Mathematical Analysis and Applications**, 2005. v. 307, n. 1, p. 350–369, jul. 2005. ISSN 0022-247X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0022247X04008741>>. Citado 4 vezes nas páginas 27, 69, 72 e 73.

Uma introdução ao estudo das matrizes de projeção em subespaços

A.1 Conceitos Iniciais

Nesta seção, apresentamos e construímos o método dos quadrados mínimos a partir das definições e proposições de complementos ortogonais, que é uma extensão da noção de vetor normal a um hiperplano; e de projeção ortogonal sobre um subespaço, que é uma extensão de projeção de um vetor sobre o outro, em espaços vetoriais com produto interno usual. A partir do cálculo diferencial, generalizamos o conceito de projeção ortogonal e apresentamos as expressões que definem projeções oblíquas de vetores em subespaços do $\mathbb{R}^n$.

Proposição A.1.1. *Sejam $\mathbb{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$, $1 \leq k \leq n$, $\mathbf{u}_k \in \mathbb{R}^n$, uma base ortogonal de um subespaço $\mathbb{W} \subset \mathbb{R}^n$ e $\mathbf{w} \in \mathbb{R}^n$, então os únicos escalares $\alpha_1, \dots, \alpha_k$ tais que $\mathbf{w} = \sum_{t=1}^k \alpha_t \mathbf{u}_t$ são dados por $\alpha_t = \frac{\langle \mathbf{u}_t, \mathbf{w} \rangle}{\langle \mathbf{u}_t, \mathbf{u}_t \rangle}$, $t = 1, \dots, k$.* —

Demonstração. Como $\mathbb{B}$ é uma base cujos vetores são ortogonais dois a dois, então há escalares $\alpha_1, \dots, \alpha_k$ unicamente determinados de modo que $\mathbf{w} = \sum_{t=1}^k \alpha_t \mathbf{u}_t$. Assim, temos:

$$\begin{aligned} \mathbf{w} &= \sum_{t=1}^k \alpha_t \mathbf{u}_t = \alpha_1 \mathbf{u}_1 + \dots + \alpha_t \mathbf{u}_t + \dots + \alpha_k \mathbf{u}_k \Rightarrow \\ \mathbf{u}_t^T \mathbf{w} &= \alpha_1 \mathbf{u}_t^T \mathbf{u}_1 + \dots + \alpha_t \mathbf{u}_t^T \mathbf{u}_t + \dots + \alpha_k \mathbf{u}_t^T \mathbf{u}_k \Rightarrow \\ \alpha_t &= \frac{\mathbf{u}_t^T \mathbf{w}}{\mathbf{u}_t^T \mathbf{u}_t} = \frac{\langle \mathbf{u}_t, \mathbf{w} \rangle}{\langle \mathbf{u}_t, \mathbf{u}_t \rangle}. \quad \blacksquare \end{aligned}$$

Definição A.1.1 ([Poole 2010]). O conjunto $\mathbb{W}^\perp = \{\mathbf{w}^\perp \in \mathbb{R}^n; \langle \mathbf{w}^\perp, \mathbf{w} \rangle = 0, \forall \mathbf{w} \in \mathbb{W}\}$ de todos os vetores ortogonais a $\mathbb{W}$ é chamado “complemento ortogonal de $\mathbb{W}$ ”. □

A Figura A.1.1 ilustra a Definição A.1.1:

Figura A.1.1: O vetor $\mathbf{w}^\perp \in \mathbb{W}^\perp$.

Agora, suponhamos que o subespaço $\mathbb{W}$ fosse gerado por uma base ortogonal $\mathbb{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$, $1 \leq k \leq n$ e, também, consideremos um vetor $\mathbf{v} = (v_1, \dots, v_j, \dots, v_n)^T \in \mathbb{R}^n$ não-nulo. Assim, para um vetor $\mathbf{u}_k \in \mathbb{W}$ e um ângulo γ definido pelos vetores $\mathbf{v}$ e $\mathbf{u}_k$, temos o vetor $\mathbf{w}_k = \text{proj}_{\mathbf{u}_k}(\mathbf{v})$ definido pela projeção ortogonal de $\mathbf{v}$ sobre $\mathbf{u}_k$. Assim, o vetor $\mathbf{w}_k^\perp \in \mathbb{W}^\perp$ é definido por $\mathbf{w}_k^\perp = \text{perp}_{\mathbf{u}_k}(\mathbf{v}) = \mathbf{v} - \text{proj}_{\mathbf{u}_k}(\mathbf{v})$. Este resultado é mostrado na Figura A.1.2.

Figura A.1.2: O vetor $\mathbf{v}$ é resultante da soma de dois vetores: um $\mathbf{w}_k \in \mathbb{W}$ e outro $\mathbf{w}_k^\perp \in \mathbb{W}^\perp$.

Note que $\mathbf{w}_k = \frac{\|\mathbf{w}_k\| \mathbf{u}_k}{\|\mathbf{u}_k\|}$. Conforme [Poole 2010], temos ainda que $\|\mathbf{w}_k\| = \|\mathbf{v}\| \cos(\gamma)$, isto é, $\|\mathbf{w}_k\| = \frac{\|\mathbf{v}\| \langle \mathbf{u}_k, \mathbf{v} \rangle}{\|\mathbf{u}_k\|}$, portanto, temos:

$$\mathbf{w}_k = \text{proj}_{\mathbf{u}_k}(\mathbf{v}) = \frac{\langle \mathbf{u}_k, \mathbf{v} \rangle}{\langle \mathbf{u}_k, \mathbf{u}_k \rangle} \mathbf{u}_k. \quad (\text{A.1.1})$$

Por conveniência, suponha que $\|\mathbf{u}_k\|^2 = \langle \mathbf{u}_k, \mathbf{u}_k \rangle = 1$, então (A.1.1) é expressa em (A.1.2):

$$\mathbf{w}_k = \text{proj}_{\mathbf{u}_k}(\mathbf{v}) = \langle \mathbf{u}_k, \mathbf{v} \rangle \mathbf{u}_k. \quad (\text{A.1.2})$$

O conceito de projeção de vetor sobre outro vetor agora é estendido para projeção de vetor sobre subespaço.

Definição A.1.2. Sejam $\mathbb{W}$ um subespaço de $\mathbb{R}^n$ e $\mathbb{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$, $1 \leq k \leq n$, uma base ortonormal para $\mathbb{W}$. Dado um vetor $\mathbf{v} \in \mathbb{R}^n$, a projeção ortogonal de $\mathbf{v}$ em $\mathbb{W}$ é definida por:

$$\mathbf{w} = \text{proj}_{\mathbb{W}}(\mathbf{v}) = \langle \mathbf{u}_1, \mathbf{v} \rangle \mathbf{u}_1 + \dots + \langle \mathbf{u}_k, \mathbf{v} \rangle \mathbf{u}_k \quad (\text{A.1.3})$$

e a componente ortogonal de $\mathbf{v}$ em relação a $\mathbb{W}$ é o vetor:

$$\mathbf{w}^\perp = \mathbf{v} - \text{proj}_{\mathbb{W}}(\mathbf{v}). \quad (\text{A.1.4})$$

□

A expressão (A.1.3) é verdadeira devido à Proposição A.1.1. Seja a matriz $U = [\mathbf{u}_1 \dots \mathbf{u}_i \dots \mathbf{u}_k] \in \mathbb{R}^{n \times k}$ cujas colunas são elementos de $\mathbb{B}$. Ao utilizarmos a definição de produto de matrizes, temos que (A.1.3) é reformulada para sua forma equivalente em (A.1.5).

$$\mathbf{w} = \text{proj}_{\mathbb{W}}(\mathbf{v}) = [\mathbf{u}_1 \dots \mathbf{u}_i \dots \mathbf{u}_k] \begin{bmatrix} \langle \mathbf{u}_1, \mathbf{v} \rangle \\ \vdots \\ \langle \mathbf{u}_i, \mathbf{v} \rangle \\ \vdots \\ \langle \mathbf{u}_k, \mathbf{v} \rangle \end{bmatrix} = U \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_i^T \\ \vdots \\ \mathbf{u}_k^T \end{bmatrix} \mathbf{v} \Rightarrow \mathbf{w} = \text{proj}_{\mathbb{W}}(\mathbf{v}) = UU^T \mathbf{v}, \quad (\text{A.1.5})$$

logo, a componente ortogonal de $\mathbf{v}$ em relação a $\mathbb{W}$ é apresentada em (A.1.6):

$$\mathbf{w}^\perp = \mathbf{v} - \text{proj}_{\mathbb{W}}(\mathbf{v}) = \mathbf{w}^\perp = (I_n - UU^T) \mathbf{v}, \quad (\text{A.1.6})$$

em que I_n é a matriz identidade de ordem n . Isso nos conduz à seguinte definição.

Definição A.1.3. Sejam $\mathbb{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$, $1 \leq k \leq n$, uma base ortonormal para $\mathbb{W}$ em $\mathbb{R}^n$, a matriz $U = [\mathbf{u}_1 \dots \mathbf{u}_k] \in \mathbb{R}^{n \times k}$ e $\mathbf{v} \in \mathbb{R}^n$. A transformação linear $P_{\mathbb{W}} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ que projeta ortogonalmente $\mathbf{v}$ em $\mathbb{W}$ tem UU^T como sua matriz canônica e a transformação linear $P_{\mathbb{W}^\perp} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ que projeta ortogonalmente $\mathbf{v}$ em $\mathbb{W}^\perp$ tem $I_n - UU^T$ como sua matriz canônica. □

A Figura A.1.3 ilustra a Definição A.1.3 para $\mathbb{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$ em $\mathbb{R}^2$.

Figura A.1.3: Ilustração da Definição A.1.3.

Proposição A.1.2. As matrizes de projeção P_W e $P_{W^\perp}$ possuem a propriedade de ser idempotente, isto é, $(P_W)^2 = P_W$ e $(P_{W^\perp})^2 = P_{W^\perp}$. —

Demonstração. De fato, desde que $P_W = UU^T$ e $P_{W^\perp} = I_n - UU^T$ e que as colunas de U são mutuamente ortogonais, então:

$$(P_W)^2 = UU^T (UU^T) = U (U^T U) U^T = UU^T = P_W$$

$$(P_{W^\perp})^2 = (I_n - UU^T) (I_n - UU^T) = I_n - 2UU^T + U (U^T U) U^T = I_n - UU^T = P_{W^\perp}. \quad \blacksquare$$

Proposição A.1.3. Seja P uma matriz idempotente, então seus autovalores só podem assumir os valores 1 ou 0. —

Demonstração. De fato, seja v um autovalor de P e $\mathbf{v} \neq \mathbf{0}$ um autovetor associado ao autovalor v , então:

1. $P\mathbf{v} = v\mathbf{v}$;
2. $P^2\mathbf{v} = P(P\mathbf{v}) = P(v\mathbf{v}) = v(P\mathbf{v}) = v^2\mathbf{v}$.

Ao subtrairmos (2) de (1) vem:

$$(P^2 - P)\mathbf{v} = (v^2 - v)\mathbf{v} \Rightarrow (v^2 - v) = 0 \Rightarrow v = 1 \text{ ou } v = 0. \quad \blacksquare$$

Proposição A.1.4. As matrizes de projeção ortogonal P_W e $P_{W^\perp}$ são matrizes simétricas e semidefinidas positiva. —

Demonstração. Como $P_W = UU^T$ e $P_{W^\perp} = I_n - UU^T$, então:

$$(P_W)^T = (UU^T)^T = (U^T)^T U^T = UU^T = P_W$$

$$(P_{W^\perp})^T = (I_n - UU^T)^T = I_n - (U^T)^T U^T = I_n - UU^T = P_{W^\perp}.$$

Pela Proposição A.1.2, $P_{\mathbb{W}}$ e $P_{\mathbb{W}^\perp}$ são matrizes idempotentes. Seja $\mathbf{0} \neq \mathbf{v} \in \mathbb{R}^n$, então:

$$\begin{aligned}\mathbf{v}^T P_{\mathbb{W}} \mathbf{v} &= \mathbf{v}^T P_{\mathbb{W}} P_{\mathbb{W}} \mathbf{v} = \mathbf{v}^T (P_{\mathbb{W}})^T P_{\mathbb{W}} \mathbf{v} = \mathbf{w}^T \mathbf{w} \geq 0 \\ \mathbf{v}^T P_{\mathbb{W}^\perp} \mathbf{v} &= \mathbf{v}^T P_{\mathbb{W}^\perp} P_{\mathbb{W}^\perp} \mathbf{v} = \mathbf{v}^T (P_{\mathbb{W}^\perp})^T P_{\mathbb{W}^\perp} \mathbf{v} = (\mathbf{w}^\perp)^T \mathbf{w}^\perp \geq 0.\end{aligned}$$

■

Enunciamos a definição que sela as condições suficientes para que uma matriz seja de projeção ortogonal.

Definição A.1.4. Seja P uma matriz quadrada e real. P é uma matriz de projeção ortogonal se $P = P^2$ e $P = P^T$. □

Por exemplo, a matriz $P = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$ é idempotente, mas não é simétrica. Portanto, trata-se de uma matriz que projeta um vetor obliquamente.

Proposição A.1.5. Seja $\mathbb{W}$ um subespaço não-degenerado em $\mathbb{R}^n$ e seja $\mathbf{0} \neq \mathbf{v} \in \mathbb{R}^n$. Então, existem vetores únicos $\mathbf{w} \in \mathbb{W}$ e $\mathbf{w}^\perp \in \mathbb{W}^\perp$ tais que $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$. —

Demonstração. Vide [Poole 2010]. ■

Segundo [Poole 2010], a expressão “menor distância” está relacionada com a noção de projeção ortogonal: se $\mathbb{W}$ é um subespaço em $\mathbb{R}^n$, então $\mathbf{w} = \text{proj}_{\mathbb{W}}(\mathbf{v})$ é o vetor em $\mathbb{W}$ que está mais próximo de $\mathbf{v}$, em outras palavras, $\mathbf{w} = \text{proj}_{\mathbb{W}}(\mathbf{v})$ é a melhor aproximação para $\mathbf{v}$ em $\mathbb{W}$.

Consideremos que $\mathbf{w}_a$ seja um vetor em $\mathbb{W}$ tal que $\mathbf{w}_a \neq \mathbf{w}$, então $\|\mathbf{v} - \mathbf{w}\| < \|\mathbf{v} - \mathbf{w}_a\|$. A Figura A.1.4 ilustra esse resultado o qual é formalizado na Proposição A.1.6 vista a seguir:

Figura A.1.4: Ilustração para a Proposição A.1.6.

Proposição A.1.6. Sejam $\mathbb{W}$ é um subespaço em $\mathbb{R}^n$ e $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{w} = \text{proj}_{\mathbb{W}}(\mathbf{v})$ é a melhor aproximação para $\mathbf{v}$ em $\mathbb{W}$. —

Demonstração. Seja $\mathbf{w}_a \neq \mathbf{w}$ um vetor em $\mathbb{W}$. Então, $\mathbf{w}_b = \mathbf{w} - \mathbf{w}_a$ também está em $\mathbb{W}$ e $\mathbf{w}^\perp = \mathbf{v} - \mathbf{w}$ é ortogonal a $\mathbf{w}_b$. Pelo teorema de Pitágoras, temos:

$$\begin{aligned} \|\mathbf{w}^\perp\|^2 + \|\mathbf{w}_b\|^2 &= \|\mathbf{w}^\perp + \mathbf{w}_b\|^2 = \|\mathbf{v} - \mathbf{w}_a\|^2 \Rightarrow \\ \|\mathbf{w}^\perp\|^2 &< \|\mathbf{w}^\perp\|^2 + \|\mathbf{w}_b\|^2 < \|\mathbf{v} - \mathbf{w}_a\|^2 \Rightarrow \\ \|\mathbf{v} - \mathbf{w}\|^2 &< \|\mathbf{v} - \mathbf{w}_a\|^2 \Rightarrow \\ \|\mathbf{v} - \text{proj}_{\mathbb{W}}(\mathbf{v})\|^2 &< \|\mathbf{v} - \mathbf{w}_a\|^2 \Rightarrow \\ \|\mathbf{v} - \text{proj}_{\mathbb{W}}(\mathbf{v})\| &< \|\mathbf{v} - \mathbf{w}_a\|. \quad \blacksquare \end{aligned} \tag{A.1.7}$$

A.2 Aplicação e propriedades da projeção ortogonal em subespaços do $\mathbb{R}^m$

Sem perda de generalidade, suponhamos que temos uma sequência finita de pontos no $\mathbb{R}^2$ não todos colineares quando tomados três a três. Esses pontos definem o conjunto $\mathbb{Y} = \{(x_{1_1}, x_{2_1}), \dots, (x_{1_t}, x_{2_t}), \dots, (x_{1_m}, x_{2_m})\}$. Por não serem colineares, não existe uma reta r que contém todos esses pontos, mas podemos pensar em uma reta r que esteja “o mais próximo possível” de passar pelos pontos de $\mathbb{Y}$. Segundo [Poole 2010], para qualquer reta r , medimos a distância vertical de cada ponto dado até r , ou seja, consideramos distância ε_t do ponto (x_{1_t}, x_{2_t}) , $t = 1, \dots, m$ à reta r definida na direção Ox_2 . E então determinamos r que minimize o *erro total* $\varepsilon = (\varepsilon_1, \dots, \varepsilon_t, \dots, \varepsilon_m)^T$. A Figura A.2.1 apresenta a disposição desses pontos no plano cartesiano.

Figura A.2.1: A reta r que está o mais próximo possível de passar pelos pontos de $\mathbb{Y}$.

O objetivo é determinar r tal que ε seja o menor possível na norma euclidiana. Assim sendo, temos o seguinte problema de minimização:

$$\text{Min} \quad \frac{1}{2} \langle \varepsilon, \varepsilon \rangle. \tag{A.2.1}$$

Segundo [Poole 2010], o termo “quadrados mínimos” relaciona-se diretamente ao problema (A.2.1), pois precisamos determinar a menor soma de quadrados, definido pela expressão $\langle \varepsilon, \varepsilon \rangle = \|\varepsilon\|^2$. O número $\|\varepsilon\|^2$ é chamado “erro quadrático da aproximação”, a reta r é chamada “reta de

quadrados mínimos ” ou “reta de melhor ajuste”. Assim, se r é a reta de quadrados mínimos para os pontos de $\mathbb{Y}$, temos o seguinte sistema:

$$\underbrace{\begin{pmatrix} x_{1_1} & 1 \\ \vdots & \vdots \\ x_{1_t} & 1 \\ \vdots & \vdots \\ x_{1_m} & 1 \end{pmatrix}}_{A \in \mathbb{R}^{m \times 2}} \underbrace{\begin{pmatrix} a_1 \\ a_0 \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} x_{2_1} \\ \vdots \\ x_{2_t} \\ \vdots \\ x_{2_m} \end{pmatrix}}_{\mathbf{b}}. \quad (\text{A.2.2})$$

Quando $\mathbf{x} \in \mathbb{R}^m$, temos, por analogia, que $A \in \mathbb{R}^{m \times n}$ e que $\mathbf{b} \in \mathbb{R}^m$. Para resolvermos o sistema (A.2.2), consideramos que $A \in \mathbb{R}^{m \times n}$ tem n colunas linearmente independentes, isto é, $\text{posto}(A) = n$ e que $m > n$. Portanto, a definição de reta de quadrados mínimos passa ser a de hiperplano de quadrados mínimos. Desde que $\varepsilon(\mathbf{x}) = \mathbf{b} - A\mathbf{x}$, então o problema (A.2.1) é redefinido de forma equivalente em (A.2.3):

$$\text{Min } \frac{1}{2} \|\mathbf{b} - A\mathbf{x}\|^2. \quad (\text{A.2.3})$$

O problema (A.2.3) é um problema de programação quadrática com função objetivo estritamente convexa. Assim, como $\mathbf{x}^*$ é um ponto de mínimo global para o problema (A.2.3), então devemos ter que:

$$\|\mathbf{b} - A\mathbf{x}^*\| \leq \|\mathbf{b} - A\mathbf{x}\| \quad \forall \mathbf{x} \in \mathbb{R}^n. \quad (\text{A.2.4})$$

Observe a semelhança entre as estruturas (A.2.4) e (A.1.7). Qualquer que seja $\mathbf{x} \in \mathbb{R}^n$, o vetor $A\mathbf{x} = \mathbf{y} \in \mathbb{R}^m$ está no conjunto $\text{Im}(A) = \{\mathbf{y} \in \mathbb{R}^m; \exists \mathbf{x} \in \mathbb{R}^n, \mathbf{y} = A\mathbf{x}\}$, o qual é denominado “espaço-imagem de A ”. Portanto, a expressão (A.2.4) é equivalente à (A.2.5):

$$\|\mathbf{b} - \mathbf{y}^*\| \leq \|\mathbf{b} - \mathbf{y}\|. \quad (\text{A.2.5})$$

Então, a solução do problema (A.2.3) é obtida ao encontrarmos um $\mathbf{y} \in \text{Im}(A)$ que esteja o mais próximo de $\mathbf{b}$. Pela Proposição A.1.6, temos:

$$A\mathbf{x}^* = \text{proj}_{\text{Im}(A)}(\mathbf{b}). \quad (\text{A.2.6})$$

Desde que $\varepsilon(\mathbf{x}^*) = \mathbf{b} - A\mathbf{x}^*$, então:

$$\varepsilon(\mathbf{x}^*) = \mathbf{b} - A\mathbf{x}^* \Rightarrow \varepsilon(\mathbf{x}^*) = \mathbf{b} - \text{proj}_{\text{Im}(A)}(\mathbf{b}) \Rightarrow \varepsilon(\mathbf{x}^*) \in (\text{Im}(A))^\perp. \quad (\text{A.2.7})$$

Como $\varepsilon(\mathbf{x}^*) \in (\text{Im}(A))^\perp$, temos que $\varepsilon(\mathbf{x}^*)$ é ortogonal a todas as n colunas de A . Ao definirmos $A = [A_1 \dots A_t \dots A_n]$, onde A_t é a t -ésima coluna de A , temos o sistema homogêneo (A.2.8):

$$\begin{cases} A_1^T \varepsilon(\mathbf{x}^*) = 0 \\ \vdots \\ A_t^T \varepsilon(\mathbf{x}^*) = 0 \\ \vdots \\ A_n^T \varepsilon(\mathbf{x}^*) = 0, \end{cases}$$

cuja forma matricial é:

$$A^T \varepsilon(\mathbf{x}^*) = \mathbf{0}. \quad (\text{A.2.8})$$

Ao compararmos (A.2.8) com (A.2.7) temos o importante resultado em (A.2.9):

$$(\text{Im}(A))^\perp = \ker(A^T), \quad (\text{A.2.9})$$

em que $\ker(A^T) = \{A^T \in \mathbb{R}^{n \times m} \text{ e } \mathbf{y} \in \mathbb{R}^m; A^T \mathbf{y} = \mathbf{0}\}$ é denominado “kernel de A^T ”. Assim, temos:

$$A^T \varepsilon(\mathbf{x}^*) = \mathbf{0} \Rightarrow A^T(\mathbf{b} - A\mathbf{x}^*) = \mathbf{0} \Rightarrow A^T A\mathbf{x}^* = A^T \mathbf{b}. \quad (\text{A.2.10})$$

A matriz $A^T A \in \mathbb{R}^{n \times n}$ é simétrica, pois $(A^T A)^T = A^T (A^T)^T = A^T A$ e definida positiva, pois para todo $\mathbf{x} \in \mathbb{R}^n$ temos: $\underbrace{\mathbf{x}^T A^T}_{\mathbf{y}^T \in \text{Im}(A)} \underbrace{A \mathbf{x}}_{\mathbf{y} \in \text{Im}(A)} = \langle \mathbf{y}, \mathbf{y} \rangle > 0$. Como $\text{posto}(A) = n$, então $A^T A$ possui inversa e a solução por quadrados mínimos

$$\mathbf{x}^* = (A^T A)^{-1} A^T \mathbf{b} \quad (\text{A.2.11})$$

é única. Com as informações obtidas em (A.2.6) e (A.2.11), enunciamos na Proposição A.2.1 uma expressão para a projeção de vetor para um subespaço.

Proposição A.2.1. *Sejam $A \in \mathbb{R}^{m \times n}$ cujas colunas formam uma base para $\text{Im}(A)$ e $\mathbf{x} \in \mathbb{R}^n$. Se $\mathbf{b} \in \mathbb{R}^m$, então a projeção de $\mathbf{b}$ em $\text{Im}(A)$ é o vetor $\mathbf{y} = \text{proj}_{\text{Im}(A)}(\mathbf{b}) = A(A^T A)^{-1} A^T \mathbf{b}$. A transformação linear $P_{\text{Im}(A)} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ que projeta $\mathbf{b} \in \mathbb{R}^m$ em $\text{Im}(A)$ tem $A(A^T A)^{-1} A^T$ como sua matriz canônica. —*

Demonstração. Como o $\text{posto}(A) = n$ e $\text{Im}(A)$ é um subespaço do $\mathbb{R}^m$, temos uma única solução determinada por quadrados mínimos para o sistema $A\mathbf{x} = \mathbf{b}$ dada por $\mathbf{x}^* = (A^T A)^{-1} A^T \mathbf{b}$. Da expressão (A.2.6), temos:

$$\mathbf{y}^* = \text{proj}_{\text{Im}(A)}(\mathbf{b}) = A\mathbf{x}^* = A(A^T A)^{-1} A^T \mathbf{b} \quad (\text{A.2.12})$$

Como a equação (A.2.12) é verdadeira para todo $\mathbf{b} \in \mathbb{R}^m$, então a verificação de $P_{\text{Im}(A)} = A(A^T A)^{-1} A^T$ é imediata. ■

A Proposição A.2.2 mostra que a matriz $P_{\text{Im}(A)}$ é de projeção ortogonal.

Proposição A.2.2. *Sejam $\text{Im}(A)$ o subespaço em $\mathbb{R}^m$ definido por n colunas linearmente independentes de uma matriz $A \in \mathbb{R}^{m \times n}$ e $P_{\text{Im}(A)} = A(A^T A)^{-1} A^T$ a matriz que projeta um vetor $\mathbf{b} \in \mathbb{R}^m$ em $\text{Im}(A)$. $P_{\text{Im}(A)}$ é uma matriz de projeção ortogonal. —*

Demonstração. Mostremos que $P_{\text{Im}(A)}$ é simétrica.

$$(P_{\text{Im}(A)})^T = (A(A^T A)^{-1} A^T)^T = (A^T)^T (A^T A)^{-1} A^T = A(A^T A)^{-1} A^T = P_{\text{Im}(A)}.$$

Mostramos que $P_{\text{Im}(A)}$ é idempotente.

$$\begin{aligned} (P_{\text{Im}(A)})^2 &= (A(A^T A)^{-1} A^T) (A(A^T A)^{-1} A^T) \Rightarrow \\ (P_{\text{Im}(A)})^2 &= A(A^T A)^{-1} \underbrace{A^T A(A^T A)^{-1} A^T}_{\text{Identidade}} \Rightarrow \\ (P_{\text{Im}(A)})^2 &= A(A^T A)^{-1} A^T = P_{\text{Im}(A)}. \end{aligned}$$

Portanto, $P_{\text{Im}(A)}$ é uma matriz de projeção ortogonal. ■

A Proposição A.2.3 apresenta uma matriz que projeta $\mathbf{b} \in \mathbb{R}^m$ em um único vetor $\mathbf{y}^\perp \in \ker(A^T)$.

Proposição A.2.3. *Sejam cujas colunas formam uma base para $\text{Im}(A)$. Se $\mathbf{b} \in \mathbb{R}^m$ e $P_{\text{Im}(A)} = A(A^T A)^{-1} A^T$ a matriz que projeta um vetor $\mathbf{b} \in \mathbb{R}^m$ em $\text{Im}(A)$, então a matriz $P_{\ker(A^T)} = I_m - A(A^T A)^{-1} A^T$, em que I_m é a matriz identidade de ordem m , é a transformação linear que projeta $\mathbf{b}$ em $\mathbf{y}^\perp \in \ker(A^T)$. Mais ainda: $P_{\ker(A^T)}$ é uma matriz de projeção ortogonal.* —

Demonstração. Desde que $(\text{Im}(A))^\perp = \ker(A^T)$ e que o posto $(A) = n$, então, pela equação (A.1.4), temos uma matriz que projeta $\mathbf{b} \in \mathbb{R}^m$ em um único vetor $\mathbf{y}^\perp \in \ker(A^T)$. Assim:

$$\begin{aligned} \mathbf{y}^\perp &= \mathbf{b} - \text{proj}_{\text{Im}(A)}(\mathbf{b}) = \mathbf{b} - A(A^T A)^{-1} A^T \mathbf{b} \Rightarrow \\ \mathbf{y}^\perp &= (I_m - A(A^T A)^{-1} A^T) \mathbf{b}. \end{aligned} \quad (\text{A.2.13})$$

Portanto, $P_{\ker(A^T)} = I_m - A(A^T A)^{-1} A^T$ é a matriz que projeta $\mathbf{b}$ em $\mathbf{y}^\perp \in \ker(A^T)$. Mostremos que $P_{\ker(A^T)}$ é simétrica.

$$(P_{\ker(A^T)})^T = (I_m - P_{\text{Im}(A)})^T = I_m^T - (P_{\text{Im}(A)})^T = I_m - P_{\text{Im}(A)} = P_{\ker(A^T)}.$$

Mostremos que $P_{\ker(A^T)}$ é idempotente.

$$(P_{\ker(A^T)})^2 = (I_m - P_{\text{Im}(A)})(I_m - P_{\text{Im}(A)}) = I_m - P_{\text{Im}(A)} - P_{\text{Im}(A)} + (P_{\text{Im}(A)})^2 = I_m - P_{\text{Im}(A)} = P_{\ker(A^T)}.$$

Assim, $P_{\ker(A^T)}$ é uma matriz de projeção ortogonal. ■

Geometricamente, as projeções e os subespaços do $\mathbb{R}^m$ são apresentados na Figura A.2.2.

Figura A.2.2: O efeito das matrizes de projeção ortogonal em subespaços do $\mathbb{R}^m$.

E são confirmadas algebricamente pelo Teorema de Pitágoras conforme (A.2.14):

$$\begin{aligned}
\|\mathbf{y}\|^2 + \|\mathbf{y}^\perp\|^2 &= \|\mathbf{v}\|^2 \Rightarrow \\
\|P_{\text{Im}(A)}\mathbf{v}\|^2 + \|P_{\text{ker}(A^T)}\mathbf{v}\|^2 &= \|\mathbf{v}\|^2 \Rightarrow \\
\mathbf{v}^T (P_{\text{Im}(A)})^T P_{\text{Im}(A)} \mathbf{v} + \mathbf{v}^T (P_{\text{ker}(A^T)})^T P_{\text{ker}(A^T)} \mathbf{v} &= \mathbf{v}^T \mathbf{v} \Rightarrow \\
\mathbf{v}^T (P_{\text{Im}(A)} + P_{\text{ker}(A^T)}) \mathbf{v} &= \mathbf{v}^T \mathbf{v} \Rightarrow \\
\mathbf{v}^T (A(A^T A)^{-1} A^T + I_m - A(A^T A)^{-1} A^T) \mathbf{v} &= \mathbf{v}^T \mathbf{v} \\
\mathbf{v}^T \mathbf{v} &= \mathbf{v}^T \mathbf{v}
\end{aligned} \tag{A.2.14}$$

A.3 Projeções ortogonais e oblíquas em subespaços do $\mathbb{R}^n$

De acordo com a Definição A.1.4, uma matriz real e quadrada P é uma matriz de projeção ortogonal se esta é idempotente e simétrica. Adicionalmente, a Proposição A.1.4 nos assegura que esta deve ser semidefinida positiva para todo $\mathbf{x} \in \mathbb{R}^n$. Entretanto, a matriz $P = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$ é um contra-exemplo de que nem toda idempotente é uma matriz de projeção ortogonal.

As consequências de assumirmos a assimetria de uma matriz idempotente é a de que ela pode não ser semidefinida positiva para todo $\mathbf{x} \in \mathbb{R}^n$ mesmo que seus autovalores assumam apenas valores 1 ou 0. Com efeito, ao escolhermos o vetor $\mathbf{x}_2 = (3, -2)^T$, então facilmente verificamos que $\mathbf{x}_2^T P \mathbf{x}_2 = -3 < 0$, logo, P não é semidefinida positiva. A matriz de projeção P também não pode ser exclusivamente semidefinida negativa. De fato, seja $\mathbf{x}_1 = (3, -1)^T$, então verificamos que $\mathbf{x}_1^T P \mathbf{x}_1 = 3 > 0$. Por fim, para um vetor não-nulo $\mathbf{x}_3$, verificamos por meio de $\mathbf{x}_3 = (0, 2)^T$ que a possibilidade de $\mathbf{x}_3^T P \mathbf{x}_3$ ser nulo existe.

Qualquer matriz $M \in \mathbb{R}^{n \times n}$ pode ser decomposta como uma soma envolvendo uma matriz simétrica e uma antissimétrica¹. Isso é verificado em (A.3.1).

$$M = \frac{1}{2}(M + M^T) + \frac{1}{2}(M^T - M) = \frac{M + M^T}{2} + \frac{M - M^T}{2}, \tag{A.3.1}$$

em que a parte simétrica de M é $\frac{1}{2}(M + M^T)$ e a parte antissimétrica de M é $\frac{1}{2}(M - M^T)$. Assim sendo, temos para qualquer $\mathbf{x} \in \mathbb{R}^n$ que:

$$\mathbf{x}^T M \mathbf{x} = \mathbf{x}^T \left(\frac{M + M^T}{2} + \frac{M - M^T}{2} \right) \mathbf{x} = \mathbf{x}^T \left(\frac{M + M^T}{2} \right) \mathbf{x} + \mathbf{x}^T \left(\frac{M - M^T}{2} \right) \mathbf{x}.$$

Mas,

$$\mathbf{x}^T \left(\frac{M - M^T}{2} \right) \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n (M_{ij} - M_{ji}) x_i x_j = 0,$$

isto é, a parte antissimétrica não influencia no cálculo de $\mathbf{x}^T M \mathbf{x}$. Portanto, para qualquer matriz $M \in \mathbb{R}^{n \times n}$, temos a seguinte identidade:

$$\mathbf{x}^T M \mathbf{x} = \mathbf{x}^T \left(\frac{M + M^T}{2} \right) \mathbf{x} \tag{A.3.2}$$

¹Uma matriz $A \in \mathbb{R}^{n \times n}$ é antissimétrica se $A = -A^T$.

Por exemplo, consideremos novamente a matriz de projeção $P = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$. Então a forma quadrática associada à matriz P é dada em (A.3.3).

$$\mathbf{x}^T P \mathbf{x} = \frac{1}{2} \mathbf{x}^T \left(\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \right) \mathbf{x} = \mathbf{x}^T \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{x} \quad (\text{A.3.3})$$

Os autovalores de $\frac{1}{2}(P + P^T)$ são $v_1 = \frac{1}{2}(1 + \sqrt{5}) > 0$ e $v_2 = \frac{1}{2}(1 - \sqrt{5}) < 0$. Isso significa que a forma quadrática $\mathbf{x}^T P \mathbf{x}$ está associada a hipérboles em $\mathbb{R}^2$ e o porquê da existência de regiões do plano cujo produto $\mathbf{x}^T P \mathbf{x}$ assume valores positivo, negativo ou nulo.

Na Seção A.2 apresentamos o desenvolvimento das matrizes de projeção ortogonal em subespaços do $\mathbb{R}^m$ sem a necessidade do cálculo diferencial. Isso foi feito com a finalidade observarmos as relações entre espaço-imagem de A e de seu complemento ortogonal cujo principal resultado é apresentado na equação (A.2.9). Nesta seção utilizamos otimização via cálculo diferencial para obtermos resultados mais gerais. Entretanto, estes resultados são para projeção de vetores em subespaços do $\mathbb{R}^n$. Iniciamos nossa discussão pela seguinte definição.

Definição A.3.1. *Seja $P \in \mathbb{R}^{n \times n}$ uma matriz arbitrária. P é uma matriz de projeção se $P = P^2$.* $\square$

Note que a matriz P identicamente nula é a única que projeta todos os vetores para o conjunto cujo único elemento é o vetor nulo e que matriz identidade é a única matriz que projeta qualquer vetor não-nulo em si mesmo.

Sejam $\mathbf{x} \in \mathbb{R}^n$ e $M \in \mathbb{R}^{n \times n}$ uma matriz simétrica e não-singular. Definimos por forma quadrática isotrópica regular² à função $\phi_M(\mathbf{x}) = \mathbf{x}^T M \mathbf{x}$, vide [Lam 2004]. Dentre os temas teóricos relacionados a esta forma quadrática, nos interessamos pelo resultado de que $\phi_M(\mathbf{x})$ é uma métrica semi-Riemanniana com assinatura³ (a_1, a_2) , isto é, M possui a_1 autovalores positivos e a_2 negativos (com todas as multiplicidades algébricas), de modo que $a_1 + a_2 = n$, vide [O'Neill 1983, Carmo 2008].

Como consequência, existem os seguintes subconjuntos em $\mathbb{R}^n$:

1. $\mathbb{E}_1 = \{\mathbf{0} \neq \mathbf{x}_1 \in \mathbb{R}^n; \phi_M(\mathbf{x}_1) > 0\}$, ϕ_M é definida positiva sobre $\mathbb{E}_1$;
2. $\mathbb{E}_2 = \{\mathbf{0} \neq \mathbf{x}_2 \in \mathbb{R}^n; \phi_M(\mathbf{x}_2) < 0\}$, ϕ_M é definida negativa sobre $\mathbb{E}_2$ e;
3. $\mathbb{E}_3 = \{\mathbf{0} \neq \mathbf{x}_3 \in \mathbb{R}^n; \phi_M(\mathbf{x}_3) = 0\}$, ϕ_M é indefinida, é isotrópico.

Foquemos agora aos problemas de otimização cujas soluções ótimas estão relacionadas às matrizes de projeção. Sejam $N \in \mathbb{R}^{m \times n}$ com $m < n$, posto $(N) = m$ e $\mathbf{v} \in \mathbb{R}^n$ um vetor conhecido, então o primeiro problema de otimização é apresentado em (A.3.4).

$$\begin{aligned} \text{Min} \quad & \frac{1}{2} \phi_M(\mathbf{v} - \mathbf{x}) = \frac{1}{2} (\mathbf{v} - \mathbf{x})^T M (\mathbf{v} - \mathbf{x}) \\ \text{s.a :} \quad & N \mathbf{x} = \mathbf{0}. \end{aligned} \quad (\text{A.3.4})$$

² $\phi_M(\mathbf{x})$ é uma forma quadrática anisotrópica regular, se M é definida positiva ($\sqrt{\phi_M(\mathbf{x})}$ define uma norma em $\mathbb{R}^n$ denominada por Métrica Riemanniana) ou se M é definida negativa.

³Em Física, as Métricas Lorentzianas são exemplos de métrica semi-Riemanniana que ocorrem para a descrição do espaço-tempo na Teoria da Relatividade Geral, vide [O'Neill 1983].

O objetivo do problema (A.3.4) é encontrar um $\mathbf{w}^\perp = \mathbf{x}^* \in \ker(N) = \{\mathbf{x} \in \mathbb{R}^n; N\mathbf{x} = \mathbf{0}\}$ cujo quadrado da “magnitude” do vetor $(\mathbf{v} - \mathbf{x})$ definida pela forma quadrática isotrópica $\phi_M(\mathbf{v} - \mathbf{x})$ com assinatura (a_1, a_2) seja a menor possível. O problema lagrangiano associado ao problema (A.3.4) é apresentado em (A.3.5):

$$\text{Min } L(\omega) = \frac{1}{2}\phi_M(\mathbf{v} - \mathbf{x}) + \lambda^T N\mathbf{x}, \quad (\text{A.3.5})$$

em que $\omega = (\mathbf{x}^T, \lambda^T)^T \in \mathbb{R}^{n+m}$ e $\lambda \in \mathbb{R}^m$ é o vetor de multiplicadores associado à restrições de igualdade $N\mathbf{x} = \mathbf{0}$. As condições de KKT, $\nabla L(\omega) = \mathbf{0}$, são apresentadas em (A.3.6):

$$\begin{pmatrix} M & N^T \\ N & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \lambda \end{pmatrix} = \begin{pmatrix} M\mathbf{v} \\ \mathbf{0} \end{pmatrix}. \quad (\text{A.3.6})$$

A única⁴ solução ótima $\mathbf{w}^\perp$ para o problema (A.3.4) é conforme (A.3.7):

$$\mathbf{x}^* = (I_n - M^{-1}N^T(NM^{-1}N^T)N)\mathbf{v}. \quad (\text{A.3.7})$$

Proposição A.3.1. *A matriz $P_{\ker(N)_M} = I_n - M^{-1}N^T(NM^{-1}N^T)N$ é idempotente, mas não é simétrica.* —

Demonstração. Mostremos que $P_{\ker(N)_M}$ é idempotente.

$$\begin{aligned} \left(P_{\ker(N)_M}\right)^2 &= (I_n - M^{-1}N^T(NM^{-1}N^T)N)(I_n - M^{-1}N^T(NM^{-1}N^T)N) \Rightarrow \\ \left(P_{\ker(N)_M}\right)^2 &= I_n - M^{-1}N^T(NM^{-1}N^T)N - M^{-1}N^T(NM^{-1}N^T)N + \\ &+ \underbrace{M^{-1}N^T(NM^{-1}N^T)NM^{-1}N^T(NM^{-1}N^T)N}_{\text{Identidade}} \Rightarrow \\ \left(P_{\ker(N)_M}\right)^2 &= I_n - M^{-1}N^T(NM^{-1}N^T)N - M^{-1}N^T(NM^{-1}N^T)N + M^{-1}N^T(NM^{-1}N^T)N \Rightarrow \\ \left(P_{\ker(N)_M}\right)^2 &= I_n - M^{-1}N^T(NM^{-1}N^T)N \Rightarrow \\ \left(P_{\ker(N)_M}\right)^2 &= P_{\ker(N)_M}. \end{aligned}$$

Mostremos que $P_{\ker(N)_M}$ não é simétrica:

$$\begin{aligned} \left(P_{\ker(N)_M}\right)^T &= (I_n - M^{-1}N^T(NM^{-1}N^T)N)^T = I_n - N^T(NM^{-1}N^T)NM^{-1} \Rightarrow \\ \left(P_{\ker(N)_M}\right)^T &\neq P_{\ker(N)_M}. \end{aligned} \quad \blacksquare$$

Pelo fato de que $P_{\ker(N)_M}$ ser idempotente e não-simétrica, segue a definição:

Definição A.3.2. *A transformação linear $P_{\ker(N)_M} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ que projeta $\mathbf{v} \in \mathbb{R}^n$ em $\mathbf{w}^\perp \in \ker(N)$ tem $I_n - M^{-1}N^T(NM^{-1}N^T)N$ como sua matriz canônica de acordo com a métrica semi-Riemanniana $\phi_M(\mathbf{x}) = \mathbf{x}^T M \mathbf{x}$ com assinatura (a_1, a_2) .* □

Para o caso particular em $M = I_n$, então, de acordo com a Definição A.1.4, temos que $P_{\ker(N)_{I_n}} = P_{\ker(N)}$ é transformação que projeta $\mathbf{v} \in \mathbb{R}^n$ ortogonalmente para o $\ker(N)$. Compare o resultado obtido com seu análogo para o $\mathbb{R}^m$ em (A.2.13).

⁴Pois $\text{posto}(N) = m$ e $\det(M) \neq 0$.

O segundo problema de otimização é apresentado em (A.3.8) a seguir:

$$\text{Min } \frac{1}{2} \phi_{M^{-1}} (\mathbf{v}^T - \mathbf{y}^T N) = \frac{1}{2} (\mathbf{v}^T - \mathbf{y}^T N) M^{-1} (\mathbf{v} - N^T \mathbf{y}). \quad (\text{A.3.8})$$

Agora, o vetor procurado é $\mathbf{w}^* = N^T \mathbf{y}^*$ que seja o mais próximo possível de $\mathbf{v}$. Está claro que este vetor procurado está no conjunto $\text{Im}(N^T) = \{\mathbf{w} \in \mathbb{R}^n; \exists \mathbf{y} \in \mathbb{R}^m, N^T \mathbf{y} = \mathbf{w}\}$. A solução deste problema generaliza o resultado apresentado na Proposição A.1.6. Como $\text{posto}(N) = m$ e $\det(M^{-1}) \neq 0$, então a única solução $\mathbf{y}^*$ é apresentada em (A.3.9).

$$\mathbf{y}^* = (NM^{-1}N^T)^{-1} NM^{-1} \mathbf{v} \Rightarrow \mathbf{w}^* = N^T (NM^{-1}N^T)^{-1} NM^{-1} \mathbf{v} \quad (\text{A.3.9})$$

Proposição A.3.2. A matriz $P_{\text{Im}(N^T)_{M^{-1}}} = N^T (NM^{-1}N^T)^{-1} NM^{-1}$ é idempotente, mas não é simétrica. —

Demonstração. Verifiquemos que $P_{\text{Im}(N^T)_{M^{-1}}}$ é idempotente.

$$\begin{aligned} \left(P_{\text{Im}(N^T)_{M^{-1}}}\right)^2 &= N^T (NM^{-1}N^T)^{-1} \underbrace{NM^{-1}N^T (NM^{-1}N^T)^{-1}}_{\text{Identidade}} NM^{-1} \\ &= N^T (NM^{-1}N^T)^{-1} NM^{-1} \\ &= P_{\text{Im}(N^T)_{M^{-1}}}. \end{aligned}$$

Verifiquemos que $P_{\text{Im}(N^T)_{M^{-1}}}$ não é simétrica.

$$\left(P_{\text{Im}(N^T)_{M^{-1}}}\right)^T = \left(N^T (NM^{-1}N^T)^{-1} NM^{-1}\right)^T = M^{-1} N^T (NM^{-1}N^T)^{-1} N \neq P_{\text{Im}(N^T)_{M^{-1}}}. \quad \blacksquare$$

Pelo fato de que $P_{\text{Im}(N^T)_{M^{-1}}}$ também ser idempotente, mas não simétrica, segue a seguinte definição.

Definição A.3.3. A transformação linear $P_{\text{Im}(N^T)_{M^{-1}}} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ que projeta $\mathbf{v} \in \mathbb{R}^n$ em $\mathbf{w} \in \text{Im}(N^T)$ tem $N^T (NM^{-1}N^T)^{-1} NM^{-1}$ como sua matriz canônica de acordo com métrica semi-Riemanniana $\phi_{M^{-1}}(\mathbf{x}) = \mathbf{x}^T M^{-1} \mathbf{x}$ com assinatura (a_1, a_2) . $\square$

Para o caso particular em $M^{-1} = I_n$, então a matriz $P_{\text{Im}(N^T)_{I_n}} = P_{\text{Im}(N^T)}$ é o operador linear que projeta $\mathbf{v} \in \mathbb{R}^n$ ortogonalmente para $\text{Im}(N^T)$. Compare este resultado com o seu análogo para o $\mathbb{R}^m$ em (A.2.12).

Com o auxílio de (A.3.7), estabelecemos uma propriedade de interesse, a qual é apresentada em (A.3.10) a seguir.

$$\begin{aligned} \mathbf{w}^* &= P_{\text{Im}(N^T)_{M^{-1}}} \mathbf{v} \Rightarrow \\ -M^{-1} \mathbf{w}^* &= -\left(M^{-1} N^T (NM^{-1}N^T)^{-1} N\right) M^{-1} \mathbf{v} \Rightarrow \\ -M^{-1} \mathbf{w}^* &= -\left(P_{\text{Im}(N^T)_{M^{-1}}}\right)^T M^{-1} \mathbf{v} \Rightarrow \\ M^{-1} \mathbf{v} - M^{-1} \mathbf{w}^* &= M^{-1} \mathbf{v} - \left(P_{\text{Im}(N^T)_{M^{-1}}}\right)^T M^{-1} \mathbf{v} \Rightarrow \\ M^{-1} (\mathbf{v} - \mathbf{w}^*) &= \left(I_n - \left(P_{\text{Im}(N^T)_{M^{-1}}}\right)^T\right) M^{-1} \mathbf{v} \Rightarrow \\ M^{-1} (\mathbf{v} - \mathbf{w}^*) &= P_{\text{ker}(N)_M} M^{-1} \mathbf{v}. \end{aligned} \quad (\text{A.3.10})$$

Ao utilizarmos o produto interno usual em $\mathbb{R}^n$, então, para qualquer $0 \neq \mathbf{v} \in \mathbb{R}^n$, temos:

$$\begin{aligned} \langle P_{\ker(N)_M} \mathbf{v}, P_{\text{Im}(N^T)_{M^{-1}}} \mathbf{v} \rangle &= \mathbf{v}^T \left(I_n - N^T (NM^{-1}N^T)^{-1} NM^{-1} \right) N^T (NM^{-1}N^T)^{-1} NM^{-1} \mathbf{v} \\ &= \mathbf{v}^T \left(N^T (NM^{-1}N^T)^{-1} NM^{-1} - N^T (NM^{-1}N^T)^{-1} NM^{-1} \right) \mathbf{v} \\ &= 0. \end{aligned} \quad (\text{A.3.11})$$

A expressão apresentada em (A.3.11) nos remete imediatamente para um dos principais resultados desta seção.

$$\ker(N) = (\text{Im}(N^T))^\perp. \quad (\text{A.3.12})$$

Este resultado pode ser comparado com seu análogo para o $\mathbb{R}^m$ em (A.2.9).

Consideremos as matrizes $N = \begin{pmatrix} 1 & 1 \end{pmatrix}$ e $M = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ e o vetor $\mathbf{v}_1 = \begin{pmatrix} 3 & 2 \end{pmatrix}^T$. Com esses elementos, afirmamos que $P_{\ker(N)_M} = \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix}$ e $P_{\text{Im}(N^T)_{M^{-1}}} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$; $P_{\ker(N)} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ e $P_{\text{Im}(N^T)} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, para $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. A Figura A.3.1 ilustra a projeção oblíqua de $\mathbf{v}_1$ em $\mathbf{w}_1^\perp = \begin{pmatrix} -2 & 2 \end{pmatrix}^T$ e em $\mathbf{w}_1 = \begin{pmatrix} 3 & 3 \end{pmatrix}^T$ e a projeção ortogonal de $\mathbf{v}_1$ em $\mathbf{w}_2^\perp = \frac{1}{2} \begin{pmatrix} 1 & -1 \end{pmatrix}^T$ e em $\mathbf{w}_2 = \frac{1}{2} \begin{pmatrix} 5 & 5 \end{pmatrix}^T$.

Figura A.3.1: Projeções oblíquas e ortogonais de $\mathbf{v}_1$ em subespaços do $\mathbb{R}^2$.

Por fim, consideremos o problema de quadrados mínimos dado em (A.3.13) a seguir:

$$\begin{aligned} \text{Min} \quad & \frac{1}{2} \phi_M(\mathbf{q}) = \frac{1}{2} \mathbf{q}^T M \mathbf{q} \\ \text{s.a :} \quad & N \mathbf{q} = \mathbf{b} \end{aligned} \quad (\text{A.3.13})$$

em que $\mathbf{b} \in \mathbb{R}^m$ é conhecido. Isto é, desejamos encontrar $\mathbf{q}^* \in \mathbb{R}^n$ com menor magnitude segundo a métrica semi-Riemanniana $\phi_M(\mathbf{q})$ de modo que $N\mathbf{q}^* = \mathbf{b}$. Chamamos atenção para o fato de que as colunas de N são vetores linearmente dependentes. Portanto, há infinitas soluções $\bar{\mathbf{q}} \in \mathbb{R}^n$ tal que $N\bar{\mathbf{q}} = \mathbf{b}$, mas a única solução $\mathbf{q}^* = \mathbf{q}_\mathbf{b}$ tal que resolve o problema (A.3.13) é dada em (A.3.14).

$$\mathbf{q}_\mathbf{b} = M^{-1}N^T(NM^{-1}N^T)^{-1}\mathbf{b} \quad (\text{A.3.14})$$

Ao particularizarmos $N \in \mathbb{R}^{n \times n}$ com $\det(N) \neq 0$, $\mathbf{b} \in \mathbb{R}^n$ e $M = I_n$, então $\mathbf{q}_\mathbf{b} = N^{-1}\mathbf{b}$. Nesse sentido, a transformação linear $M^{-1}N^T(NM^{-1}N^T)^{-1} \in \mathbb{R}^{n \times m}$ generaliza o conceito de matriz inversa. Essa generalização é conhecida como pseudo-inversa de Moore-Penrose⁵ para $M = I_n$.

⁵Vide os originais em [Moore 1920] e [Penrose 1955].

Regularização, lei de inércia de Sylvester e pseudo-espço euclidiano

B.1 Regularização

Iniciamos esta seção comentando sobre a possibilidade de aproximarmos uma matriz $M \in \mathbb{C}^{n \times n}$ qualquer arbitrariamente por uma matriz $\hat{M} \in \mathbb{C}^{n \times n}$ inversível. Isso é mostrado através da Proposição B.1.1 a seguir.

Proposição B.1.1. *Seja $M \in \mathbb{C}^{n \times n}$ qualquer e $B \in \mathbb{C}^{n \times n}$ inversível. Então existem escalares α_1 e α_2 com $0 < \alpha_1 \leq \alpha_2$ que dependem de M e de B tais que $\hat{M} = M + \beta B$ é inversível para todo $\beta \in \mathbb{C}$ com $0 < |\beta| < \alpha_1$ ou para todo $\beta \in \mathbb{C}$ com $|\beta| > \alpha_2$.* —

Demonstração. Como B é inversível, então podemos escrever $\hat{M} = (MB^{-1} + \beta I_n) B$ o que implica diretamente que $\hat{M}$ terá inversa se, e somente, se $(MB^{-1} + \beta I_n)$ admitir inversa. Seja $f(v) = \det(MB^{-1} - vI_n)$ o polinômio característico de MB^{-1} e $\mathbb{S} = \{v_1, \dots, v_j, \dots, v_n\} \subset \mathbb{C}$ o conjunto de autovalores de MB^{-1} . Se $v_j = 0$, para todo $j = 1, \dots, n$, então podemos escolher $\alpha_1 = \alpha_2$ quaisquer. Caso contrário, seja $\alpha_1 = \min\{|v_j|, v_j \neq 0\}$ e $\alpha_2 = \max\{|v_1|, \dots, |v_j|, \dots, |v_n|\}$. Nesses casos, podemos escolher seguramente um $\beta \in \mathbb{C}$ que satisfaça $0 < |\beta| < \alpha_1$ ou $|\beta| > \alpha_2$, pois, para esses valores de β , garantimos que $\beta \notin \mathbb{S}$ e, por consequência, que $f(v) = \det(MB^{-1} + (\beta - v)I_n) \neq 0$. Logo, para todo $\beta \in \mathbb{C}$ que satisfaça $0 < |\beta| < \alpha_1$ ou $|\beta| > \alpha_2$, $\hat{M}$ possui inversa. ■

Notemos que quanto menor a magnitude de β , mais próxima $\hat{M}$ estará de M . De fato:

$$0 \leq \|M + \beta B\| \leq \|M\| + |\beta| \|B\|.$$

Então, ao fazermos $|\beta| \rightarrow 0$, então o resultado segue imediatamente. Assim, fica claro que qualquer matriz $M \in \mathbb{C}^{n \times n}$, inversível ou não, pode ser arbitrariamente aproximada pela matriz $M + \beta I_n$ por exemplo, desde que tomemos um $\beta \in \mathbb{C}$ conveniente com magnitude suficientemente pequena, mas não nula. Portanto, feitas estas considerações, temos:

$$M = \lim_{|\beta| \rightarrow 0} (M + \beta I_n). \quad (\text{B.1.1})$$

Agora, consideremos a possibilidade de que a matriz simétrica $\theta_k \in \mathbb{R}^{n \times n}$, expressão (5.5.26), ser singular em alguma iteração k do método PDRNLI. Neste caso, a Proposição B.1.1 torna-se um importante instrumento de correção para este inconveniente. Como os autovalores de θ_k são números reais, [Domingues 2009, p. 262], então podemos restringir o número β_k apenas para o campo dos números reais. Logo, podemos arbitrariamente aproximar θ_k por uma matriz $\hat{\theta}_k$ inversível segundo (B.1.2):

$$\hat{\theta}_k = \theta_k + \beta_k I_n, \quad (\text{B.1.2})$$

para um $|\beta_k|$ conveniente e suficientemente pequeno.

Em relação às demais matrizes inversíveis, temos algumas significativas vantagens em escolhermos a matriz identidade I_n para representar a matriz B . A primeira dessas vantagens é que a matriz I_n é a mais simples e intuitiva; a segunda é que podemos verificar facilmente que βI_n perturba diretamente os autovalores de θ_k , logo, favorece o melhoramento do número de condicionamento θ_k .

Com efeito, sejam θ_k singular e $\mathbf{v}^k \in \mathbb{R}^n$ não-nulo. Assim, temos a seguinte forma quadrática:

$$(\mathbf{v}^k)^T \hat{\theta}_k \mathbf{v}^k = (\mathbf{v}^k)^T (\theta_k + \beta_k I_n) \mathbf{v}^k. \quad (\text{B.1.3})$$

Como θ_k é simétrica, então ela é diagonalizável por uma matriz $Q_k \in \mathbb{R}^{n \times n}$ ortogonal cuja j -ésima coluna é o autovetor normalizado associado ao respectivo autovalor v_{j_k} , $j = 1, \dots, n$, [Domingues 2009, p. 263]. Logo, ao fazermos a transformação $\mathbf{v}^k = Q_k \mathbf{y}^k$, $\mathbf{0} \neq \mathbf{y}^k \in \mathbb{R}^n$, temos:

$$\begin{aligned} (\mathbf{v}^k)^T (\theta_k + \beta_k I_n) \mathbf{v}^k &= (\mathbf{v}^k)^T \theta_k \mathbf{v}^k + \beta_k (\mathbf{v}^k)^T \mathbf{v}^k \\ &= (\mathbf{y}^k)^T Q_k^T \theta_k Q_k \mathbf{y}^k + \beta_k (\mathbf{y}^k)^T Q_k^T Q_k \mathbf{y}^k \\ &= (\mathbf{y}^k)^T D_k \mathbf{y}^k + \beta_k (\mathbf{y}^k)^T \mathbf{y}^k, \end{aligned}$$

em que:

$$D_k = \text{diag}(v_{1_k}, \dots, v_{p_k}, v_{p+1_k}, \dots, v_{q_k}, v_{q+1_k}, \dots, v_{s_k}) \quad (\text{B.1.4})$$

e com p indicando o número de autovalores positivos, q o número de autovalores negativos e s o número de autovalores nulos de θ_k . Finalizando, temos:

$$\begin{aligned} (\mathbf{y}^k)^T D_k \mathbf{y}^k + \beta_k (\mathbf{y}^k)^T \mathbf{y}^k &= \\ (\mathbf{y}^k)^T (D_k + \beta_k I_n) \mathbf{y}^k. \end{aligned} \quad (\text{B.1.5})$$

De (B.1.5), constatamos que os s autovalores nulos são eliminados. Logo, segue imediatamente que a operação $(\theta_k + \beta_k I_n)$ oferece uma correção sobre o posto de θ_k e, além disso, temos que (B.1.3) define uma forma quadrática isotrópica regular $\phi_{\hat{\theta}_k}(\mathbf{v}^k)$. Isso significa que $\phi_{\hat{\theta}_k}(\mathbf{v}^k)$ define uma semi-métrica Riemanniana, a qual está em função de quão pequeno é $|\beta_k|$. Sendo assim, seja a_{1_k} e a_{2_k} , respectivamente, o número de autovalores positivos e negativos de $\hat{\theta}_k$ apresentada em (B.1.3), então, temos:

$$\hat{D}_k = \text{diag}(\hat{v}_{1_k}, \dots, \hat{v}_{a_{1_k}}, \hat{v}_{a_{1+1_k}}, \dots, \hat{v}_{a_{2_k}}) \quad (\text{B.1.6})$$

e, por definição, que a assinatura de $\phi_{\hat{\theta}_k}(\mathbf{v}^k)$ é dada pelo par (a_{1_k}, a_{2_k}) .

B.2 Lei de inércia de Sylvester

Agora, consideremos a Definição B.2.1 a seguir:

Definição B.2.1. Seja $M \in \mathbb{C}^{n \times n}$ hermitiana. A tripla (a_1, a_2, a_3) , em que a_1 é o número de autovalores positivos de M , a_2 é o número de autovalores negativos de M e a_3 é o número de autovalores nulos de M (contando com todas as multiplicidades algébricas) é denominada inércia de M . $\square$

De imediato, temos que a inércia de $\hat{\theta}_k$ é a tripla $(a_1, a_2, 0)$. Por simplicidade, a inércia da matriz $\hat{\theta}_k$ é indicada apenas pelo par (a_{1_k}, a_{2_k}) . Em 1852, Sylvester¹ descobriu que a inércia de M é um invariante sob as transformações de congruência, [Meyer 2001, p. 568].

Definição B.2.2. Sejam $M \in \mathbb{C}^{n \times n}$ e $N \in \mathbb{C}^{n \times n}$. Dizemos que M e N são congruentes se existe uma matriz $C \in \mathbb{C}^{n \times n}$ inversível tal que a transformação de congruência $M = C^*NC$ é verificada. $\square$

O par (M, N) de matrizes congruentes é denotado pelo símbolo $M \cong N$ (lê-se: M é cônica a N). Vamos mostrar que a transformação de congruência é uma relação de equivalência na proposição 3.2.

Proposição B.2.1. A transformação de congruência é uma relação de equivalência sobre o conjunto das matrizes complexas quadradas de dimensão n . —

Demonstração. Sejam $M \in \mathbb{C}^{n \times n}$, $N \in \mathbb{C}^{n \times n}$ e $R \in \mathbb{C}^{n \times n}$. Para garantirmos a relação de equivalência, é necessário que a transformação de congruência seja reflexiva, simétrica e transitiva.

1. Reflexiva: $M \cong M$.

De fato, seja I_n a matriz identidade em $\mathbb{C}^{n \times n}$, então:

$$M = I_n M I_n \Rightarrow M \cong M.$$

2. Simetria: $M \cong N \Leftrightarrow N \cong M$.

($\rightarrow$) Assumamos que $M \cong N$. Como $M \cong N$, então existe uma matriz $C \in \mathbb{C}^{n \times n}$ inversível tal que $M = C^*NC$. Assim: $N = (C^*)^{-1}MC^{-1} = (C^{-1})^*MC^{-1}$. Se definirmos $H = C^{-1}$, então temos $N = H^*MH$ e, assim, $M \cong N \Rightarrow N \cong M$.

($\leftarrow$) É realizado de forma análoga.

3. Transitividade: Se $M \cong N$ e $N \cong R$, então $M \cong R$.

De fato, como $M \cong N$, então existe uma matriz $C_1 \in \mathbb{C}^{n \times n}$ inversível tal que $M = C_1^*NC_1$. Além disso, temos, também, que $N \cong R$, então existe uma matriz $C_2 \in \mathbb{C}^{n \times n}$ inversível tal que $N = C_2^*RC_2$. Como C_1 e C_2 são inversíveis, então C_1C_2 também é. Logo:

$$\begin{aligned} M &= C_1^*NC_1 \\ &= C_1^*(C_2^*RC_2)C_1 \\ &= (C_2C_1)^*RC_2C_1. \end{aligned}$$

Ao fazermos $J = C_2C_1$, então $M = J^*RJ$. Portanto, $M \cong R$. $\blacksquare$

¹ James Joseph Sylvester (1814–1887) é o fundador da *American Journal of Mathematics*, o primeiro jornal de matemática dos Estados Unidos, [Poole 2010].

Proposição B.2.2. Se $M \in \mathbb{C}^{n \times n}$ qualquer é côngrua a uma matriz $N \in \mathbb{C}^{n \times n}$ hermitiana, então M é hermitiana. —

Demonstração. De fato: $M \cong N$ implica que existe $C \in \mathbb{C}^{n \times n}$ tal que $M = C^*NC$. Logo, temos:

$$\begin{aligned} (M)^* &= (C^*NC)^* = C^*N^*(C^*)^* = C^*NC = M \Rightarrow \\ (M)^* &= M. \end{aligned} \quad \blacksquare$$

Como a transformação de congruência é uma relação de equivalência, então temos que somente matrizes hermitianas são côngruas a matrizes hermitianas. Então, dado uma matriz M hermitiana qualquer, então existe uma matriz C inversível tal que $D = C^*MC$ é uma matriz diagonal. O Exemplo B.2.1 ilustra a importância desta afirmação.

Exemplo B.2.1. Suponha que $\theta_k \in \mathbb{R}^{n \times n}$, expressão (5.5.26), é definida positiva. Então $\theta_k \cong I_n$. $\square$

Seja $Q_k \in \mathbb{R}^{n \times n}$ ortonormal tal que $D_k = Q_k^T \theta_k Q_k$ é uma matriz diagonal, com $D_k = \text{diag}(v_{jk})_{j=1}^n$. Como θ_k é definida positiva implica que os autovalores v_{jk} são estritamente positivos para todo $j = 1, \dots, n$, [Meyer 2001, p. 559]. Como $v_{jk} > 0$, obviamente $\sqrt{v_{jk}}$ é um número real. Assim, considere a transformação de congruência

$$D_k = \left(\sqrt{D_k}\right)^T I_n \sqrt{D_k} \Leftrightarrow I_n = \left(\sqrt{D_k}\right)^{-T} D_k \left(\sqrt{D_k}\right)^{-1},$$

em que $\sqrt{D_k} = \text{diag}(\sqrt{v_{jk}})_{j=1}^n$. Portanto, pela propriedade de transitividade da relação de equivalência, temos:

$$\begin{aligned} D_k &= Q_k^T \theta_k Q_k \Rightarrow \\ \left(\sqrt{D_k}\right)^{-T} D_k \left(\sqrt{D_k}\right)^{-1} &= \left(Q_k \left(\sqrt{D_k}\right)^{-1}\right)^T \theta_k Q_k \left(\sqrt{D_k}\right)^{-1} \Rightarrow \\ I_n &= \left(Q_k \left(\sqrt{D_k}\right)^{-1}\right)^T \theta_k Q_k \left(\sqrt{D_k}\right)^{-1}, \end{aligned} \quad (\text{B.2.1})$$

logo, $\theta_k \cong I_n$.

Com procedimentos análogos ao que foi apresentado no Exemplo B.2.1, é fácil de provarmos que toda $M \in \mathbb{C}^{n \times n}$ hermitiana e definida positiva é côngrua à I_n . Enunciamos o Teorema da lei de inércia de Sylvester na Proposição B.2.3. O teorema de Sylvester estabelece a relação que há entre a transformação de congruência com a inércia de uma matriz.

Proposição B.2.3 (Lei de Inércia de Sylvester). Sejam $M \in \mathbb{C}^{n \times n}$ e $N \in \mathbb{C}^{n \times n}$ hermitianas. A lei de inércia de Sylvester assegura que: $M \cong N$ se, e somente, se M e N possuem a mesma inércia, ou seja, M e N possuem exatamente a mesma quantidade de autovalores positivos, negativos e nulos (contando as multiplicidades algébricas). —

Demonstração. Vide [Meyer 2001, p. 568] $\blacksquare$

B.3 O pseudo-espaço euclidiano

Retomemos à matriz $\hat{\theta}_k \in \mathbb{R}^{n \times n}$ apresentada em (B.1.2) cuja assinatura é (a_{1k}, a_{2k}) . Chamamos a atenção que a assinatura pode ser diferente de uma iteração para outra se o problema de otimização é não-convexo.

Vamos supor que a matriz $\hat{\theta}_k$ está definida em $\mathbb{C}^{n \times n}$ (hermitiana). Desta forma, consideremos um vetor $\mathbf{v}^k \in \mathbb{C}^n$ e um $\beta_k \in \mathbb{R}$ suficientemente pequeno². Então a forma quadrática isotrópica regular $\phi_{\hat{\theta}_k}(\mathbf{v}^k)$, é representada por

$$\phi_{\hat{\theta}_k}(\mathbf{v}^k) = (\mathbf{v}^k)^* \hat{\theta}_k \mathbf{v}^k = \sum_{j=1}^n \sum_{i=1}^n \hat{\theta}_k(i, j) \bar{v}_{i_j} v_{j_k}.$$

Seja $Q_k \in \mathbb{C}^{n \times n}$ uma matriz ortonormal tal que a matriz diagonal $\hat{D}_k$ seja expressa por $\hat{D}_k = Q_k^* \hat{\theta}_k Q_k$. Pelo Teorema de Sylvester, Proposição B.2.3, constatamos que a matriz $\hat{D}_k \cong \hat{\theta}_k$, logo, $\hat{D}_k$ contém a_{1_k} autovalores v_k positivos e a_{2_k} negativos v_k de $\hat{\theta}_k$. Seja $\mathbf{v}^k = Q_k \mathbf{y}^k$, com $\mathbf{y}^k \in \mathbb{C}^n$. Então, temos:

$$\phi_{\hat{\theta}_k}(Q_k \mathbf{y}^k) = \sum_{i=1}^{a_1} (v_{i_k} + \beta_k) |y_{i_k}|^2 + \sum_{j=1}^{a_2} (v_{j_k} + \beta_k) |y_{j_k}|^2,$$

com $0 < |\beta_k| < \min_{j=1, \dots, n} \{|v_{j_k}|\}$.

Como estamos trabalhando no espaço vetorial complexo, então tem sentido extrairmos a raiz quadrada dos números $(v_{j_k} + \beta_k) < 0$, para $j = 1, \dots, a_2$. Então, ao fazermos a mudança de variável $\mathbf{t}^k = \left(\sqrt{\hat{D}_k}\right)^{-1} \mathbf{y}^k$ com:

$$\sqrt{\hat{D}_k} = \begin{pmatrix} \sqrt{v_{1_k} + \beta_k} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{v_{a_{1_k}} + \beta_k} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & i\sqrt{-(v_{a_{1+1_k}} + \beta_k)} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & i\sqrt{-(v_{a_{2_k}} + \beta_k)} \end{pmatrix}, \quad (\text{B.3.1})$$

em que $i^2 = -1$. Então:

$$\begin{aligned} \phi_{\hat{\theta}_k}\left(Q_k \left(\sqrt{\hat{D}_k}\right)^{-1} \mathbf{t}^k\right) &= (\mathbf{t}^k)^* \left(\sqrt{\hat{D}_k}\right)^{-*} Q_k^* \hat{\theta}_k Q_k \left(\sqrt{\hat{D}_k}\right)^{-1} \mathbf{t}^k \\ &= (\mathbf{t}^k)^* \left(\sqrt{\hat{D}_k}\right)^{-*} \hat{D}_k \left(\sqrt{\hat{D}_k}\right)^{-1} \mathbf{t}^k \\ &= \sum_{s=1}^{a_1} \frac{(v_{s_k} + \beta_k)}{(\sqrt{(v_{s_k} + \beta_k)})^2} \bar{t}_{s_k} t_{s_k} + \sum_{j=1}^{a_2} \frac{i\sqrt{-(v_{j_k} + \beta_k)}(-i\sqrt{-(v_{j_k} + \beta_k)})}{(v_{j_k} + \beta_k)^2} (v_{j_k} + \beta_k) \bar{t}_{j_k} t_{j_k} \\ &= \sum_{s=1}^{a_1} \frac{(v_{s_k} + \beta_k)}{(\sqrt{(v_{s_k} + \beta_k)})^2} \bar{t}_{s_k} t_{s_k} + \sum_{j=1}^{a_2} \frac{ii(\sqrt{-(v_{j_k} + \beta_k)})^2 (v_{j_k} + \beta_k)}{(v_{j_k} + \beta_k)^2} \bar{t}_{j_k} t_{j_k} \end{aligned}$$

logo, temos:

$$\phi_{\hat{\theta}_k}\left(Q_k \left(\sqrt{\hat{D}_k}\right)^{-1} \mathbf{t}^k\right) = \sum_{i=1}^{a_1} |t_{i_k}|^2 - \sum_{j=1}^{a_2} |t_{j_k}|^2. \quad (\text{B.3.2})$$

Pela lei de inércia de Sylvester, temos que (B.3.2) é a forma mais simples de representarmos $\phi_{\hat{\theta}_k}(\mathbf{v}^k)$ com assinatura (a_1, a_2) através de transformações de congruência. Notemos que os coeficientes de cada parcela $|t_{i_k}|^2$ e $|t_{j_k}|^2$ são os números 1 e -1 respectivamente. Assim, definimos a matriz

²Autovalores de matrizes hermitianas são sempre números reais, [Domingues 2009, p. 262].

de inércia $\mathcal{I}_n$ de ordem n conforme (B.3.3):

$$\mathcal{I}_n = \begin{pmatrix} I_{a_1} & 0 \\ 0 & -I_{a_2} \end{pmatrix}, \quad (\text{B.3.3})$$

em que I_{a_1} é a matriz identidade de ordem a_1 e I_{a_2} é a matriz identidade de ordem a_2 .

Antes de prosseguirmos, consideremos o caso particular em que a matriz $\hat{\theta}_k$ com assinatura (a_{1k}, a_{2k}) está em $\mathbb{R}^{n \times n}$. Assim, a partir de (B.1.5), segue que:

$$\phi_{\hat{\theta}_k}(\mathbf{Q}_k \mathbf{y}^k) = \sum_{i=1}^{a_1} (v_{i_k} + \beta_k) y_{i_k}^2 + \sum_{j=1}^{a_2} (v_{j_k} + \beta_k) y_{j_k}^2$$

com $0 < |\beta_k| < \min_{j=1, \dots, n} \{|v_{j_k}|\}$ e $\mathbf{y}^k \in \mathbb{R}^n$. Agora, não tem sentido em extrair a raiz quadrada dos números negativos $(v_{j_k} + \beta_k) < 0$, para $j = 1, \dots, a_2$, mas, de fato, $-(v_{j_k} + \beta_k) > 0$. Então, podemos definir $\sqrt{\hat{D}_k} \in \mathbb{R}^{n \times n}$ da seguinte forma:

$$\sqrt{\hat{D}_k} = \begin{pmatrix} \sqrt{v_{1_k} + \beta_k} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{v_{a_{1k}} + \beta_k} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \sqrt{-(v_{a_{1k}+1_k} + \beta_k)} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \sqrt{-(v_{a_{2k}} + \beta_k)} \end{pmatrix}. \quad (\text{B.3.4})$$

Logo, ao fazermos a mudança de variável $\mathbf{v}^k = \mathbf{Q}_k \left(\sqrt{\hat{D}_k} \right)^{-1} \mathbf{t}^k$ e aplicarmos a forma quadrática $\phi_{\hat{\theta}_k} \left(\mathbf{Q}_k \left(\sqrt{\hat{D}_k} \right)^{-1} \mathbf{t}^k \right)$, então obtemos a equação (B.3.2).

Definição B.3.1. Sejam o vetor $\mathbf{v} \in \mathbb{C}^n$ e a matriz de inércia $\mathcal{I}_n$ segundo (B.3.3), cuja assinatura é (a_1, a_2) . Considere que $a_1 \neq 0$ e $a_2 \neq 0$ com $a_1 + a_2 = n$. Definimos por forma quadrática isotrópica regular em sua forma canônica por:

$$\phi_{\mathcal{I}_n}(\mathbf{v}) = \mathbf{v}^* \mathcal{I}_n \mathbf{v} = \sum_{i=1}^{a_1} |v_i|^2 - \sum_{j=1}^{a_2} |v_j|^2. \quad (\text{B.3.5})$$

□

Agora, consideremos a seguinte definição:

Definição B.3.2. Seja $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ e um escalar $\alpha \in \mathbb{R}$. A função f é denominada “forma bilinear” se f satisfaz as propriedades (i) – (iv) de (B.3.6). Além disso, f é denominada “forma bilinear simétrica” se f satisfaz a propriedade (v) e f é definida positiva se satisfaz a propriedade (vi).

$$\begin{aligned} (i) \quad & f(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) = f(\mathbf{u}_1, \mathbf{v}) + f(\mathbf{u}_2, \mathbf{v}) \\ (ii) \quad & f(\alpha \mathbf{u}, \mathbf{v}) = \alpha f(\mathbf{u}, \mathbf{v}) \\ (iii) \quad & f(\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{u}, \mathbf{v}_1) + f(\mathbf{u}, \mathbf{v}_2) \\ (iv) \quad & f(\mathbf{u}, \alpha \mathbf{v}) = \alpha f(\mathbf{u}, \mathbf{v}) \\ (v) \quad & f(\mathbf{u}, \mathbf{v}) = f(\mathbf{v}, \mathbf{u}) \\ (vi) \quad & f(\mathbf{u}, \mathbf{u}) > 0, \forall \mathbf{0} \neq \mathbf{u} \in \mathbb{R}^n. \end{aligned} \quad (\text{B.3.6})$$

□

Se $\mathbb{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_j, \dots, \mathbf{e}_n\}$ define uma base qualquer para o $\mathbb{R}^n$, então $\mathbf{u} = \sum_{j=1}^n u_j \mathbf{e}_j$ e $\mathbf{v} = \sum_{j=1}^n v_j \mathbf{e}_j$, onde u_j e v_j são, respectivamente, as componentes do vetor $\mathbf{u}$ e $\mathbf{v}$ na base $\mathbb{B}$. Logo, ao fazermos uso exclusivo das propriedades (i) – (iv) de (B.3.6) diversas vezes, temos:

$$\begin{aligned} f(\mathbf{u}, \mathbf{v}) &= f\left(\sum_{i=1}^n u_i \mathbf{e}_i, \mathbf{v}\right) \\ &= \sum_{i=1}^n f(u_i \mathbf{e}_i, \mathbf{v}) \\ &= \sum_{i=1}^n u_i f(\mathbf{e}_i, \mathbf{v}) \\ &= \sum_{i=1}^n u_i f\left(\mathbf{e}_i, \sum_{j=1}^n v_j \mathbf{e}_j\right) \\ &= \sum_{i=1}^n u_i \sum_{j=1}^n f(\mathbf{e}_i, v_j \mathbf{e}_j) \\ &= \sum_{i=1}^n u_i \sum_{j=1}^n v_j f(\mathbf{e}_i, \mathbf{e}_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n f(\mathbf{e}_i, \mathbf{e}_j) u_i v_j \end{aligned}$$

ou, ainda:

$$f(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T M \mathbf{v}. \quad (\text{B.3.7})$$

A matriz M é denominada “matriz da forma bilinear f na base $\mathbb{B}$ ”, cujo elemento $M(i, j)$ é $f(\mathbf{e}_i, \mathbf{e}_j)$. Se a propriedade (v) em (B.3.6) é verificada, temos que $M = M^T$. A propriedade (vi) em (B.3.6) exige que a forma bilinear simétrica $f(\mathbf{u}, \mathbf{u}) = \mathbf{u}^T M \mathbf{u} > 0$ para todo $\mathbf{u} \neq \mathbf{0}$, em outras palavras, a propriedade exige $M \cong I_n$.

É imediato constatar-mos que a forma quadrática $\phi_M(\mathbf{u})$ de assinatura $(n, 0, 0)$ é um caso particular de forma bilinear e que está associada à elipsoides em $\mathbb{R}^n$, pois $f(\mathbf{u}, \mathbf{u}) = \phi_M(\mathbf{u})$. Além disso, dado uma forma quadrática $\phi_M(\mathbf{u})$ de assinatura (a_1, a_2, a_3) , podemos estabelecer uma forma bilinear simétrica $f(\mathbf{u}, \mathbf{v})$ de forma única a partir de uma forma quadrática, [Shafarevich e Remizov 2012, p. 196].

Antes de prosseguirmos, considere $f_1(\mathbf{u}, \mathbf{v})$ uma forma bilinear qualquer. Obtemos uma forma bilinear simétrica a partir de $f_1(\mathbf{u}, \mathbf{v})$ ao fazermos:

$$f(\mathbf{u}, \mathbf{v}) = \frac{1}{2} (f_1(\mathbf{u}, \mathbf{v}) + f_1(\mathbf{v}, \mathbf{u})),$$

logo, os procedimentos a seguir são feitos ao assumirmos uma forma bilinear simétrica.

Sejam $\mathbf{u} + \mathbf{v}$ e $f(\mathbf{u}, \mathbf{v})$ uma forma bilinear simétrica. Assim sendo, temos:

$$\begin{aligned} \phi_M(\mathbf{u} + \mathbf{v}) &= f(\mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}) \\ &= f(\mathbf{u}, \mathbf{u} + \mathbf{v}) + f(\mathbf{v}, \mathbf{u} + \mathbf{v}) \\ &= f(\mathbf{u}, \mathbf{u}) + f(\mathbf{u}, \mathbf{v}) + f(\mathbf{v}, \mathbf{u}) + f(\mathbf{v}, \mathbf{v}). \end{aligned}$$

Como f é simétrica, então:

$$\begin{aligned} \phi_M(\mathbf{u} + \mathbf{v}) &= f(\mathbf{u}, \mathbf{u}) + f(\mathbf{u}, \mathbf{v}) + f(\mathbf{v}, \mathbf{u}) + f(\mathbf{v}, \mathbf{v}) \\ &= \phi_M(\mathbf{u}) + 2f(\mathbf{u}, \mathbf{v}) + \phi_M(\mathbf{v}). \end{aligned}$$

Logo:

$$f(\mathbf{u}, \mathbf{v}) = \frac{1}{2} (\phi_M(\mathbf{u} + \mathbf{v}) - \phi_M(\mathbf{u}) - \phi_M(\mathbf{v})). \quad (\text{B.3.8})$$

Uma definição formal e simples de espaço euclidiano segue de acordo com [Shafarevich e Remizov 2012, p. 213].

Definição B.3.3. *Um espaço euclidiano é um espaço vetorial real sobre o qual é definida uma forma bilinear simétrica fixa cuja forma quadrática associada é definida positiva.* $\square$

De acordo com [Shafarevich e Remizov 2012, p. 265], muito dos teoremas provados para espaços euclidianos em $\mathbb{R}^n$ permanecem válidos se abrirmos mão da condição $f(\mathbf{u}, \mathbf{v}) > 0$ para todo $\mathbf{0} \neq \mathbf{u} \in \mathbb{R}^n$ ou se a substituirmos por uma condição mais fraca. Sem a condição de que M seja definida positiva, o produto interno $f(\mathbf{u}, \mathbf{v}) = \langle \mathbf{u}, \mathbf{v} \rangle_M$ não difere em nada a partir de uma arbitrária forma bilinear simétrica. Como a equação (B.3.8) mostra, este produto interno é definido de modo único através da forma quadrática $\phi_M(\mathbf{u})$.

Então se munirmos o $\mathbb{R}^n$ com uma forma bilinear simétrica não-degenerada, isto é, M não possui autovalores nulos, então a fórmula (B.3.5) e a Proposição B.2.3 mostram que, a menos de um isomorfismo, um espaço é unicamente definido pelo posto de M e pela assinatura (a_1, a_2) da forma quadrática $\phi_M(\mathbf{u})$, [Shafarevich e Remizov 2012, p. 266].

Definição B.3.4 (Pseudo-espaço euclidiano). *Um pseudo-espaço euclidiano é um espaço vetorial real sobre o qual é definida uma forma bilinear simétrica fixa cuja forma quadrática ϕ_M associada é isotrópica regular.* $\square$

Pela lei de inércia de Sylvester, se M tem assinatura (a_1, a_2) com $a_1 + a_2 = n$, $M \cong \mathcal{I}_n$, então o produto interno canônico de um pseudo-espaço euclidiano em $\mathbb{R}^n$ é dado em (B.3.9).

$$f(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T \mathcal{I}_n \mathbf{v} = \sum_{i=1}^{a_1} u_i v_i - \sum_{j=1}^{a_2} u_j v_j. \quad (\text{B.3.9})$$

O quadrado da magnitude de um vetor $\mathbf{u} \in \mathbb{R}^n$ é dado pela métrica semi-Riemanniana, a qual é formalizada através da forma quadrática isotrópica regular canônica $\phi_{\mathcal{I}_n}(\mathbf{u})$ apresentada em (B.3.10).

$$\phi_{\mathcal{I}_n}(\mathbf{u}) = \mathbf{u}^T \mathcal{I}_n \mathbf{u} = \sum_{i=1}^{a_1} (u_i)^2 - \sum_{j=1}^{a_2} (u_j)^2. \quad (\text{B.3.10})$$

O mais conhecido exemplo de aplicação do pseudo-espaço euclidiano é relacionado com a teoria da relatividade geral, na qual muni-se o $\mathbb{R}^4$ com o produto interno $f(\mathbf{u}, \mathbf{v})$ cuja forma quadrática tem assinatura $(3, 1)$. Tal pseudo-espaço é denominado “espaço de Minkowsky”, [Shafarevich e Remizov 2012, p. 268].

Agora, mostramos uma identidade simples que nos permite associar $\phi_{\mathcal{I}_n}(\mathbf{u})$ com produto interno usual em um espaço euclidiano $\langle \mathbf{p} + \mathbf{q}, \mathbf{p} - \mathbf{q} \rangle$, onde $\mathbf{p}$ e $\mathbf{q}$ são vetores em $\mathbb{R}^n$ que dependem de $\mathbf{u}$. Consideremos um $\mathbf{u} \in \mathbb{R}^n$ em um espaço vetorial euclidiano cujas coordenadas são na base canônica, então:

$$\mathbf{u} = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1 & 0 \\ 0 & \mathbf{u}_2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{p} + \mathbf{q},$$

onde $\mathbf{p} = \begin{pmatrix} (\mathbf{u}_1)^T & 0 \end{pmatrix}^T \in \mathbb{R}^n$ e $\mathbf{q} = \begin{pmatrix} 0 & (\mathbf{u}_2)^T \end{pmatrix}^T \in \mathbb{R}^n$, então, ao aplicarmos (B.3.10) segue que:

$$\phi_{\mathcal{I}_n}(\mathbf{u}) = \|\mathbf{u}_1\|^2 - \|\mathbf{u}_2\|^2 = \|\mathbf{p}\|^2 - \|\mathbf{q}\|^2 = \langle \mathbf{p} + \mathbf{q}, \mathbf{p} - \mathbf{q} \rangle,$$

mas, $\mathbf{p} + \mathbf{q} = \mathbf{u}$, obtemos a seguinte identidade:

$$\phi_{\mathcal{I}_n}(\mathbf{u}) = \langle \mathbf{u}, \mathbf{p} - \mathbf{q} \rangle. \quad (\text{B.3.11})$$

Com (B.3.11), podemos interpretar geometricamente o significado de um vetor $\mathbf{u} \neq \mathbf{0}$ ser isotrópico, isto é, ser ortogonal a si mesmo.

Exemplo B.3.1. Seja $\mathbb{R}^2$ munido com produto interno definido por $f(\mathbf{u}, \mathbf{v}) = u_1v_1 - u_2v_2$ cuja forma quadrática é $\phi_{\mathcal{I}_2}(\mathbf{u}) = \mathbf{u}^T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{u} = u_1^2 - u_2^2$. Determine algebricamente: a condição para que $\phi_{\mathcal{I}_2}(\mathbf{u})$ seja definida positiva; a condição para que $\phi_{\mathcal{I}_2}(\mathbf{u})$ seja definida negativa; a condição para que $\phi_{\mathcal{I}_2}(\mathbf{u})$ seja isotrópica; Represente-os geometricamente. Além disso, represente geometricamente as pseudo-esferas definidas pelas equações $\phi_{\mathcal{I}_2}(\mathbf{u}) = 1$ e $\phi_{\mathcal{I}_2}(\mathbf{u}) = -1$. $\square$

Notamos que $\mathbf{p} = (u_1, 0)^T$ e $\mathbf{q} = (0, u_2)^T$. A condição para que $\phi_{\mathcal{I}_2}(\mathbf{u})$ seja definida positiva é simples e é dada por:

$$\phi_{\mathcal{I}_2}(\mathbf{u}) = \|\mathbf{p}\|^2 - \|\mathbf{q}\|^2 = u_1^2 - u_2^2 > 0,$$

portanto, a condição é satisfeita quando $u_1^2 > u_2^2$.

Por analogia, a condição para que $\phi_{\mathcal{I}_2}(\mathbf{u})$ seja definida negativa é dada por:

$$\phi_{\mathcal{I}_2}(\mathbf{u}) = \|\mathbf{p}\|^2 - \|\mathbf{q}\|^2 = u_1^2 - u_2^2 < 0,$$

ou seja, $u_1^2 < u_2^2$.

A isotropia ocorre quando:

$$\phi_{\mathcal{I}_2}(\mathbf{u}) = \|\mathbf{p}\|^2 - \|\mathbf{q}\|^2 = u_1^2 - u_2^2 = 0,$$

isto é, $u_1^2 = u_2^2$. A ilustração geométrica é conforme apresentado na Figura B.3.1.

Figura B.3.1: Representação geométrica de um Pseudo-Espaço Euclidiano em $\mathbb{R}^2$.