

Universidade Estadual Paulista “Júlio de Mesquita Filho”
Faculdade de Engenharia

**Método Primal-Dual Previsor-Corretor de Pontos
Interiores e Exteriores com Estratégias de Correção de
Inércia e Suavização Hiperbólica Aplicado ao Problema
de Despacho Econômico com Ponto de Carregamento
de Válvula e Representação da Transmissão**

Diego Nunes da Silva

**Bauru – SP
Dezembro – 2014**

Universidade Estadual Paulista “Júlio de Mesquita Filho”
Faculdade de Engenharia

Método Primal-Dual Previsor-Corretor de Pontos Interiores e Exteriores com Estratégias de Correção de Inércia e Suavização Hiperbólica Aplicado ao Problema de Despacho Econômico com Ponto de Carregamento de Válvula e Representação da Transmissão

Diego Nunes da Silva

Orientador: Prof. Dr. Antonio Roberto Balbo

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Faculdade de Engenharia de Bauru da Universidade Estadual Paulista “Júlio de Mesquita Filho” – Campus de Bauru, como um dos requisitos para obtenção do título de Mestre em Engenharia Elétrica.

Bauru – SP
Dezembro – 2014

Silva, Diego Nunes da.

Método primal-dual previsor-corretor de pontos interiores e exteriores com estratégias de correção de inércia e suavização hiperbólica aplicado ao problema de despacho econômico com ponto de carregamento de válvula e representação da transmissão / Diego Nunes da Silva, 2014
261f.

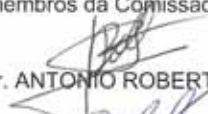
Orientador: Antonio Roberto Balbo.

Dissertação (Mestrado) - Universidade Estadual Paulista. Faculdade de Engenharia, Bauru, 2014

1. Despacho econômico com ponto de carregamento de válvula. 2. Otimização não-linear. 3. Método de pontos interiores. 4. Função barreira logarítmica modificada. 5. Suavização hiperbólica. I. Universidade Estadual Paulista. Faculdade de Engenharia. II. Título.

ATA DA DEFESA PÚBLICA DA DISSERTAÇÃO DE MESTRADO DE DIEGO NUNES DA SILVA, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, DO(A) FACULDADE DE ENGENHARIA DE BAURU.

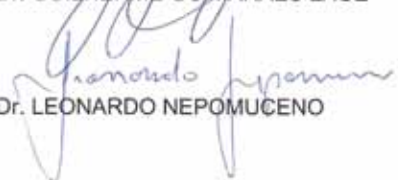
Aos 17 dias do mês de dezembro do ano de 2014, às 09:00 horas, no(a) Anfiteatro do S.T.I., da Faculdade de Engenharia de Bauru, reuniu-se a Comissão Examinadora da Defesa Pública, composta pelos seguintes membros: Prof. Dr. ANTONIO ROBERTO BALBO do(a) Departamento de Matemática / Faculdade de Ciências de Bauru, Prof. Dr. GUILHERME GUIMARÃES LAGE do(a) Departamento de Engenharia Elétrica/UFSCar, Prof. Dr. LEONARDO NEPOMUCENO do(a) Departamento de Engenharia Elétrica / Faculdade de Engenharia de Bauru, sob a presidência do primeiro, a fim de proceder a arguição pública da DISSERTAÇÃO DE MESTRADO de DIEGO NUNES DA SILVA, intitulado "MÉTODO PRIMAL-DUAL PREVISOR-CORRETOR DE PONTOS INTERIORES E EXTERIORES COM ESTRATÉGIAS DE CORREÇÃO DE INÉRCIA E SUAVIZAÇÃO HIPERBÓLICA APLICADO AO PROBLEMA DE DESPACHO ECONÔMICO COM PONTO DE VÁLVULA E REPRESENTAÇÃO DA TRANSMISSÃO". Após a exposição, o discente foi arguido oralmente pelos membros da Comissão Examinadora, tendo recebido o conceito final: Aprovado. Nada mais havendo, foi lavrada a presente ata, que, após lida e aprovada, foi assinada pelos membros da Comissão Examinadora.



Prof. Dr. ANTONIO ROBERTO BALBO



Prof. Dr. GUILHERME GUIMARÃES LAGE



Prof. Dr. LEONARDO NEPOMUCENO

PROPOSTA DE ALTERAÇÃO DO TÍTULO

A COMISSÃO EXAMINADORA PROPÕE A ALTERAÇÃO DO TÍTULO DO TRABALHO DO ALUNO:
DIEGO NUNES DA SILVA

DE: "MÉTODO PRIMAL-DUAL PREVISOR-CORRETOR DE PONTOS INTERIORES E EXTERIORES
COM ESTRATÉGIAS DE CORREÇÃO DE INÉRCIA E SUAVIZAÇÃO HIPERBÓLICA APLICADO AO
PROBLEMA DE DESPACHO ECONÔMICO COM PONTO DE VÁLVULA E REPRESENTAÇÃO DA
TRANSMISSÃO"

PARA:

"MÉTODO PRIMAL-DUAL *** COM PONTO DE CARREGAMENTO
DE VÁLVULA E REPRESENTAÇÃO DA TRANSMISSÃO"

Bauru, 17 de dezembro de 2014.


Prof. Dr. Antônio Roberto Balbo
Orientador

*Aos meus pais, Darci e Eunice,
por me apoiarem em todas as minhas decisões,
ao meu irmão, Danilo, e minha cunhada, Isabela,
por compreenderem as ausências,
e aos meus sobrinhos, Gustavo, Davi e Pedro,
por serem alegria na minha vida.*

Dedico.

Agradecimentos

Primeiramente, agradeço à Deus por me dar saúde, paciência e perseverança para conseguir terminar este trabalho.

Agradecer a todas as pessoas que contribuíram para este trabalho é relativamente difícil. Por isso, quero começar agradecendo àqueles que me ajudaram e incentivaram durante o mestrado, mas que devido à minha curta memória não forem citados a seguir.

Ao professor Dr. Antonio Roberto Balbo, pela orientação precisa, amizade e ensinamentos. Muito mais do que um exemplo de ser humano e de profissional, o senhor me ajudou a levantar em um momento em que nem eu acreditava em mim mesmo, mostrando-me um novo caminho.

À minha família, por me incentivarem a prosseguir e continuar tentando, mesmo quando as coisas não davam certo.

À professora Dra. Edméa Cássia Baptista, ao professor Dr. Leonardo Nepomuceno e ao professor Dr. Guilherme Guimarães Lage, por participarem das bancas de avaliação, pelos ensinamentos e sugestões que tanto enriqueceram este trabalho.

Às minhas duas irmãs acadêmicas, Elis e Gabriela. Obrigado pelas discussões e sugestões e, acima de tudo, pelos passeios no shopping e idas ao cinema, quando precisávamos refrescar as ideias.

Aos amigos do LOEESP e do “Labore”, tanto os que já terminaram o mestrado quanto os que continuam na luta, por propiciarem um agradável ambiente de trabalho.

Aos professores do Programa de Pós-Graduação em Engenharia Elétrica e da Licenciatura em Matemática da UNESP - Bauru, que contribuíram para minha formação acadêmica.

Aos funcionários da Seção Técnica de Pós-Graduação da Faculdade de Engenharia e do Departamento de Matemática, pela presteza e atenção concedida.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), pelo apoio financeiro.

*“Foi o tempo que você dedicou à sua rosa,
que a fez tão importante.”*

Antoine de Saint-Exupéry

*“Duas estradas bifurcavam numa árvore,
Eu trilhei a menos percorrida,
E isto fez toda a diferença.”*

Robert Lee Frost

Resumo

Em um problema de despacho econômico, objetiva-se determinar a geração de potência ativa das unidades geradoras, de modo que a demanda seja atendida e os custos dos combustíveis utilizados na geração sejam minimizados, ao mesmo tempo que as restrições de operação do sistema sejam atendidas. A função objetivo de custo deste problema pode ser modelada de diversas maneiras, porém uma forma mais realista da mesma inclui os efeitos dos pontos de carregamento de válvula. Quando os pontos de carregamento de válvula são considerados, a função objetivo se torna não-diferenciável, o que impossibilita a aplicação direta de métodos clássicos de otimização na resolução do problema. Por este motivo, a maioria dos trabalhos disponíveis na literatura para a resolução do Problema de Despacho Econômico com Ponto de Carregamento de Válvula (PDE-PV) são baseados em heurísticas. Neste trabalho, apresenta-se um método Primal-Dual Previsor-Corretor Barreira Logarítmica Modificada (PDPCBLM) associado a uma estratégia de correção de inércia e uma técnica de suavização hiperbólica para a resolução de problemas de despacho econômico com ponto de carregamento de válvula. A estratégia de correção de inércia visa garantir a convergência do método para mínimos locais do problema. A técnica de suavização hiperbólica, quando aplicada ao PDE-PV e suas variantes – o Problema de Despacho Econômico com representação das Perdas e Ponto de Carregamento de Válvula (PDE-PPV) e o Problema de Despacho Econômico com Ponto de Carregamento de Válvula e Representação da Transmissão (PDE-PV-RT) –, possibilita a utilização do método PDPCBLM para sua resolução. Uma implementação desses métodos, realizada em Matlab, mostra que tais métodos são eficientes na resolução do PDE-PV e suas variantes, obtendo resultados de boa qualidade, quando comparados aos disponíveis na literatura.

Palavras-chave: despacho econômico com ponto de carregamento de válvula, otimização não-linear, método de pontos interiores, função barreira logarítmica modificada, suavização hiperbólica.

Abstract

In the economic dispatch problem, the objective is to determine the active power output of the generating units in order to supply the demand and minimize the fuel costs associated with power generation, while taking into account system operational constraints. The cost objective function of this problem can be defined in many different ways, but a more realistic form of representation includes valve-point load effects. When valve-point load effects are considered, the objective function becomes non-differentiable, which precludes the direct application of classical optimization methods to solve the problem. For this reason, most proposals to solve the Economic Dispatch Problem with Valve-Point Load Effects (EDP-VPE) available in the literature are based on heuristics. In this work, we present a Primal-Dual Predictor-Corrector Modified Logarithmic Barrier method (PDPCMLB) associated with an inertia correction strategy and a hyperbolic smoothing technique to solve economic dispatch problems with valve-point load effects. The purpose of the inertia correction is to ensure that the proposed method converges to a local minimum of the problem. The hyperbolic smoothing technique, when applied to the EDP-VPE and its variations – the Economic Dispatch Problem with Loss Representation and Valve-Point Load Effects (EDP-LVPE) and the Network Constrained Economic Dispatch with Valve-Point Load Effects (NC-EDP-VPE) –, allows the use of the PDPCMLB method to solve it. An implementation of these methods, performed in Matlab, shows that they are efficient to solve the EDP-VPE and its variations, obtaining good results when compared to others available in the literature.

Keywords: economic dispatch with valve-point load effects, non-linear optimization, interior point methods, modified logarithmic barrier function, hyperbolic smoothing.

Sumário

1	Introdução	1
2	Geração Termelétrica	5
2.1	Eletricidade e suas formas de geração	7
2.1.1	Geração termelétrica	7
2.1.2	Geração nuclear	7
2.1.3	Geração hidrelétrica	8
2.1.4	Geração eólica	9
2.1.5	Geração solar	10
2.1.6	Bioeletricidade	11
2.2	Geração termelétrica: Brasil e mundo	11
2.2.1	Geração termelétrica no Brasil	11
2.2.2	Geração termelétrica no mundo	14
2.3	Aspectos técnicos e econômicos da geração termelétrica	16
3	Despacho Econômico	19
3.1	Geração termelétrica: custos e perdas	19
3.2	Despacho Econômico e problemas relacionados	22
3.2.1	Fluxo de Potência Ótimo	22
3.2.2	<i>Unit Commitment</i>	23
3.2.3	Despacho Hidrotérmico	24
3.3	Modelos Matemáticos do PDE	25
3.3.1	Função Objetivo	25
3.3.2	Problema de Despacho Econômico com Ponto de Carregamento de Válvula	28
3.3.3	Problema de Despacho Econômico com Representação das Perdas e Ponto de Carregamento de Válvula	29
3.3.4	Problema de Despacho Econômico com Ponto de Carregamento de Válvula e Representação da Transmissão	30
3.4	Revisão Bibliográfica	31

4	Métodos de Otimização Restrita	47
4.1	Introdução à otimização	47
4.2	Métodos de otimização	51
4.2.1	Método Dual-Lagrangiano	51
4.2.2	Método de Penalidade	53
4.2.3	Método de Barreira	54
4.2.4	Método de Rescalamento Não-Linear e de Barreira Modificada . . .	57
4.2.4.1	Função Barreira Logarítmica Modificada	60
4.2.4.2	Sobre a atualização dos estimadores dos multiplicadores de Lagrange	62
5	Metodologia de Resolução	65
5.1	O Método Primal-Dual Previsor-Corretor Barreira Logarítmica Modificada com Correção de Inércia	67
5.1.1	O sistema de direções de busca	67
5.1.1.1	Linearização da equação (5.8a)	69
5.1.1.2	Linearização da equação (5.8b)	70
5.1.1.3	Linearização da equação (5.8c)	71
5.1.1.4	Linearização da equação (5.8d)	72
5.1.1.5	O sistema de direções	72
5.1.2	Correção de Inércia	73
5.1.2.1	Decomposição de Matrizes Simétricas Indefinidas	75
5.1.2.2	Algoritmo da Correção de Inércia	78
5.1.3	Cálculo das Direções	80
5.1.3.1	Procedimento Previsor	80
5.1.3.2	Procedimento Corretor	83
5.1.4	Atualização do parâmetro de barreira, cálculo do novo ponto e do comprimento do passo	84
5.1.5	Critério de parada	87
5.1.6	Algoritmo do Método PDPCBLM	88
5.2	Suavização Hiperbólica	88
5.2.1	A técnica de suavização hiperbólica	92
5.2.2	Propriedades da suavização hiperbólica	93
6	Resultados Numéricos	95
6.1	PDE-PV	96
6.1.1	Caso 3 Geradores – Parte I	96
6.1.2	Caso 3 Geradores – Parte II	99
6.1.3	Caso 3 Geradores – Parte III	102

6.1.4	Caso 13 Geradores – Parte I	106
6.1.5	Caso 13 Geradores – Parte II	110
6.1.6	Caso 19 Geradores	114
6.2	PDE-PPV	123
6.2.1	Sistema 6 Barras de Wood e Wollenberg	124
6.2.2	Sistema IEEE-14 Barras	129
6.2.3	Sistema IEEE-30 Barras	130
6.3	PDE-PV-RT	139
6.3.1	Sistema 6 Barras de Wood e Wollenberg – Parte I	139
6.3.2	Sistema 6 Barras de Wood e Wollenberg – Parte II	140
6.3.3	Sistema IEEE-14 Barras – Parte I	147
6.3.4	Sistema IEEE-14 Barras – Parte II	150
6.3.5	Sistema IEEE-30 Barras – Parte I	156
6.3.6	Sistema IEEE-30 Barras – Parte II	157
7	Conclusões e Perspectivas Futuras	169
8	Trabalhos Publicados	173
A	Cálculo dos Coeficientes-B	191
B	Revisão da Teoria de Otimização	199
B.1	Otimização irrestrita	199
B.1.1	Conceitos básicos	199
B.1.2	Convexidade e suas implicações	201
B.1.3	Condições de otimalidade	202
B.2	Otimização restrita	206
B.2.1	Restrições de igualdade	207
B.2.1.1	Condições de otimalidade para restrições de igualdade	210
B.2.2	Restrições de desigualdade	213
B.2.2.1	Condições de otimalidade para restrições de desigualdade	216
B.2.3	Restrições de igualdade e desigualdade	218
B.3	A função Lagrangiana	220
B.3.1	Jogos e a noção de dualidade	221
B.3.2	Dualidade Lagrangiana	222
C	Banco de Dados dos Sistemas	229
C.1	Sistema de 6 Barras de Wood e Wollenberg	229
C.2	IEEE-14 Barras	230
C.3	IEEE-30 Barras	232

Lista de Figuras

2.1	Potência gerada pelos diversos tipos de equipamentos.	6
2.2	Diagrama de funcionamento de um sistema de potência com geração hidrelétrica.	9
2.3	Participação das diversas formas de geração de eletricidade, no mundo, em 2011.	14
2.4	Curva de entrada-saída idealizada.	16
3.1	Possibilidades para representação das curvas da taxa de calor incremental.	26
3.2	Funções custo e suas derivadas, com representação de ponto de carregamento de válvula e sem representação de ponto de carregamento válvula.	27
3.3	Modelo de um transformador em-fase.	31
5.1	Função ϑ para diferentes valores de η	92
5.2	Gráfico da função $f(x_1, x_2)$ com suas curvas de nível, em suas versões suavizadas.	94
6.1	Sequência de potências geradas pelo método PDPCBLM para o PDE-PV de 3 geradores (Parte I).	100
6.2	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PV de 3 geradores (Parte I)	100
6.3	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PV de 3 geradores (Parte I).	101
6.4	Trajetória dos pontos $(P_{G,1}, P_{G,2})$ – PDE-PV: 3 geradores (Parte II).	104
6.5	Sequência de potências geradas pelo método PDPCBLM para o PDE-PV de 3 geradores (Parte II).	104
6.6	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PV de 3 geradores (Parte II)	105
6.7	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PV de 3 geradores (Parte II).	105
6.8	Sequência de potências geradas pelo método PDPCBLM para o PDE-PV de 3 geradores (Parte III).	106

6.9	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PV de 3 geradores (Parte III)	108
6.10	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PV de 3 geradores (Parte III).	108
6.11	Sequência de potências geradas pelo método PDPCBLM para o PDE-PV de 13 geradores (Parte I).	112
6.12	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PV de 13 geradores (Parte I)	112
6.13	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PV de 13 geradores (Parte I).	113
6.14	Sequência de potências geradas pelo método PDPCBLM para o PDE-PV de 13 geradores (Parte II).	116
6.15	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PV de 13 geradores (Parte II)	116
6.16	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PV de 13 geradores (Parte II).	117
6.17	Sequência de potências geradas pelo método PDPCBLM para o PDE-PV de 19 geradores.	120
6.18	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PV de 19 geradores.	120
6.19	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PV de 19 geradores.	122
6.20	Esquema de resolução do PDE-PPV.	124
6.21	Sequência de potências geradas pelo método PDPCBLM para o PDE-PPV de 6 Barras.	127
6.22	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PPV de 6 Barras.	127
6.23	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PPV de 6 Barras.	128
6.24	Sequência de potências geradas pelo método PDPCBLM para o PDE-PPV de 14 Barras.	132
6.25	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PPV de 14 Barras.	132
6.26	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PPV de 14 Barras.	133
6.27	Sequência de potências geradas pelo método PDPCBLM para o PDE-PPV de 30 Barras.	136
6.28	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PPV de 30 Barras.	136

6.29	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PPV de 30 Barras.	137
6.30	Sequência de potências geradas pelo método PDPCBLM para o PDE-PV-RT de 6 Barras (Parte I)	142
6.31	Sequência de ângulos de tensão nas barras (diferentes da <i>slack</i>) geradas pelo método PDPCBLM para o PDE-PV-RT de 6 Barras (Parte I).	142
6.32	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PV-RT de 6 Barras (Parte I).	143
6.33	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PV-RT de 6 Barras (Parte I).	143
6.34	Sequência de potências geradas pelo método PDPCBLM para o PDE-PV-RT de 6 Barras (Parte II)	145
6.35	Sequência de ângulos de tensão nas barras (diferentes da <i>slack</i>) geradas pelo método PDPCBLM para o PDE-PV-RT de 6 Barras (Parte II).	145
6.36	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PV-RT de 6 Barras (Parte II).	146
6.37	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PV-RT de 6 Barras (Parte II).	146
6.38	Comparação entre os preços de energia obtidos ao resolver o PDE-PPV, PDE-PV-RT (Parte I) e PDE-PV-RT (Parte II) para o sistema 6 Barras. .	147
6.39	Sequência de potências geradas pelo método PDPCBLM para o PDE-PV-RT de 14 Barras (Parte I)	150
6.40	Sequência de ângulos de tensão nas barras (diferentes da <i>slack</i>) geradas pelo método PDPCBLM para o PDE-PV-RT de 14 Barras (Parte I).	151
6.41	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PV-RT de 14 Barras (Parte I).	151
6.42	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PV-RT de 14 Barras (Parte I).	152
6.43	Sequência de potências geradas pelo método PDPCBLM para o PDE-PV-RT de 14 Barras (Parte II)	154
6.44	Sequência de ângulos de tensão nas barras (diferentes da <i>slack</i>) geradas pelo método PDPCBLM para o PDE-PV-RT de 14 Barras (Parte II).	154
6.45	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PV-RT de 14 Barras (Parte II).	155
6.46	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PV-RT de 14 Barras (Parte II).	155
6.47	Comparação entre os preços de energia obtidos ao resolver o PDE-PPV, PDE-PV-RT (Parte I) e PDE-PV-RT (Parte II) para o sistema 14 Barras.	157

6.48	Sequência de potências geradas pelo método PDPCBLM para o PDE-PV-RT de 30 Barras (Parte I)	158
6.49	Sequência de ângulos de tensão nas barras (diferentes da <i>slack</i>) geradas pelo método PDPCBLM para o PDE-PV-RT de 30 Barras (Parte I).	160
6.50	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PV-RT de 30 Barras (Parte I).	160
6.51	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PV-RT de 30 Barras (Parte I).	162
6.52	Sequência de potências geradas pelo método PDPCBLM para o PDE-PV-RT de 30 Barras (Parte II)	164
6.53	Sequência de ângulos de tensão nas barras (diferentes da <i>slack</i>) geradas pelo método PDPCBLM para o PDE-PV-RT de 30 Barras (Parte II).	164
6.54	Evolução dos multiplicadores de Lagrange das restrições de igualdade e desigualdade no caso PDE-PV-RT de 30 Barras (Parte II).	165
6.55	Sequência de parâmetros de correção de inércia no método PDPCBLM para o PDE-PV-RT de 30 Barras (Parte II).	165
6.56	Comparação entre os preços de energia obtidos ao resolver o PDE-PPV, PDE-PV-RT (Parte I) e PDE-PV-RT (Parte II) para o sistema 30 Barras.	167
B.1	Representação geométrica do teorema B.5.	204
B.2	Curva da região factível e vetores tangentes em $(0, 1)^t$, no exemplo B.3.	209
B.3	Gráfico da função $\mathcal{L}_*(\lambda)$ no exemplo B.4.	225

Lista de Tabelas

2.1	Produção de hidreletricidade	10
2.2	Consumo mundial de gás natural (milhões de tep)	15
6.1	Características do sistema – PDE-PV: 3 geradores	97
6.2	Inicialização – PDE-PV: 3 geradores (Parte I)	97
6.3	Resultados (Iterações) – PDE-PV: 3 geradores (Parte I)	98
6.4	Resultados – PDE-PV: 3 geradores (Parte I)	99
6.5	Inicialização – PDE-PV: 3 geradores (Parte II)	101
6.6	Resultados (Iterações) – PDE-PV: 3 geradores (Parte II)	103
6.7	Resultados – PDE-PV: 3 geradores (Parte II)	104
6.8	Resultados (Iterações) – PDE-PV: 3 geradores (Parte III)	107
6.9	Resultados – PDE-PV: 3 geradores (Parte III)	108
6.10	Características do sistema – PDE-PV: 13 geradores	109
6.11	Inicialização – PDE-PV: 13 geradores (Parte I)	110
6.12	Resultados (Iterações) – PDE-PV: 13 geradores (Parte I)	111
6.13	Resultados – PDE-PV: 13 geradores (Parte I)	113
6.14	Inicialização – PDE-PV: 13 geradores (Parte II)	114
6.15	Resultados (Iterações) – PDE-PV: 13 geradores (Parte II)	115
6.16	Resultados – PDE-PV: 13 geradores (Parte II)	117
6.17	Características do sistema – PDE-PV: 19 geradores	118
6.18	Inicialização – PDE-PV: 19 geradores	119
6.19	Resultados (Iterações) – PDE-PV: 19 geradores	121
6.20	Resultados – PDE-PV: 19 geradores	122
6.21	Características do sistema – PDE-PPV: 6 Barras	125
6.22	Inicialização – PDE-PPV: 6 Barras	125
6.23	Resultados (Iterações) – PDE-PPV: 6 Barras	126
6.24	Resultados – PDE-PPV: 6 Barras	128
6.25	Magnitudes e Ângulos de Tensão – PDE-PPV: 6 Barras	128
6.26	Características do sistema – PDE-PPV: 14 Barras	129
6.27	Inicialização – PDE-PPV: 14 Barras	130
6.28	Resultados (Iterações) – PDE-PPV: 14 Barras	131

6.29	Resultados – PDE-PPV: 14 Barras	132
6.30	Magnitudes e Ângulos de Tensão – PDE-PPV: 14 Barras	133
6.31	Características do sistema – PDE-PPV: 30 Barras	134
6.32	Inicialização – PDE-PPV: 30 Barras	134
6.33	Resultados (Iterações) – PDE-PPV: 30 Barras	135
6.34	Resultados – PDE-PPV: 30 Barras	137
6.35	Magnitudes e Ângulos de Tensão – PDE-PPV: 30 Barras	138
6.36	Inicialização – PDE-PV-RT: 6 Barras (Parte I)	140
6.37	Resultados (Iterações) – PDE-PV-RT: 6 Barras (Parte I)	141
6.38	Resultados – PDE-PV-RT: 6 Barras (Parte I)	143
6.39	Resultados (Iterações) – PDE-PV-RT: 6 Barras (Parte II)	144
6.40	Resultados – PDE-PV-RT: 6 Barras (Parte II)	146
6.41	Inicialização – PDE-PV-RT: 14 Barras (Parte I)	148
6.42	Resultados (Iterações) – PDE-PV-RT: 14 Barras (Parte I)	149
6.43	Resultados – PDE-PV-RT: 14 Barras (Parte I)	152
6.44	Resultados (Iterações) – PDE-PV-RT: 14 Barras (Parte II)	153
6.45	Resultados – PDE-PV-RT: 14 Barras (Parte II)	156
6.46	Inicialização – PDE-PV-RT: 30 Barras (Parte I)	158
6.47	Resultados (Iterações) – PDE-PV-RT: 30 Barras (Parte I)	159
6.48	Resultados – PDE-PV-RT: 30 Barras (Parte I)	161
6.49	Resultados (Iterações) – PDE-PV-RT: 30 Barras (Parte II)	163
6.50	Resultados – PDE-PV-RT: 30 Barras (Parte II)	166
C.1	Dados de Barra – 6 Barras	229
C.2	Dados de Linha – 6 Barras	230
C.3	Dados de Barra – 14 Barras	230
C.4	Dados de Linha – 14 Barras	231
C.5	Dados de Barra – 30 Barras	232
C.6	Dados de Linha – 30 Barras	233

Capítulo 1

Introdução

A ideia de estabelecer a representação de algum fenômeno ou problema por meio da linguagem matemática tem possibilitado a análise e a compreensão dos mesmos, além de permitir a aproximação da matemática e das situações do mundo real. Esta abordagem aos problemas reais, utilizando recursos matemáticos, é o que comumente denominamos de modelagem matemática. Modelos matemáticos permitem o planejamento e auxiliam na tomada de decisões em diversas áreas, desde a engenharia até a medicina.

Na última década, o campo da programação matemática, e juntamente a ele a área de otimização, evoluiu de maneira significativa, impulsionado pelas transformações e exigências sócio-econômico-culturais. Novas abordagens e a implementação eficiente de algoritmos permitiram o tratamento e solução de problemas cada vez mais complexos, especialmente aqueles para os quais a obtenção de soluções analíticas é inviável. Devido às diversas características de tais problemas (como não-linearidade, não-convexidade, não-diferenciabilidade ou integralidade de variáveis), o desenvolvimento de métodos numéricos eficientes e com convergência satisfatória se tornou um importante campo de pesquisa.

Em engenharia elétrica, diversos problemas podem ser modelados matematicamente, em particular aqueles relacionados à área de sistemas de energia. No Brasil, a matriz energética elétrica se ampara, essencialmente, nas usinas hidrelétricas, que correspondem a cerca de 80% de toda energia elétrica produzida em território nacional (Brasil, 2012). Contudo, o racionamento de energia elétrica ocorrido no ano de 2001, em virtude da escassez de chuvas e da precariedade dos investimentos em geração e transmissão de energia, levou o governo brasileiro a investir mais em usinas que não dependem do ciclo de águas, em particular termelétricas, movidas a gás, carvão e óleo. Além disso, a situação brasileira difere bastante da maioria dos países no mundo. Um relatório da *International Energy Agency* (2013) indica que mais de 60% da geração de energia elétrica mundial é obtida através de termelétricas convencionais.

Dessa forma, o aumento da demanda por energia elétrica é uma questão que tem preocupado diversas nações, uma vez que isso gera novos desafios no que diz respeito ao aumento das perdas e dos custos de geração, bem como à segurança e confiabilidade do sistema elétrico. Uma maneira de auxiliar na tomada de decisão em geração de energia é a resolução do Problema de Despacho Econômico (PDE), que consiste na programação ótima das unidades geradoras de modo que a demanda por energia elétrica seja atendida e as restrições do sistema sejam respeitadas, ao mesmo tempo que o custo associado ao uso de combustíveis empregados na geração termelétrica é minimizado.

O PDE pode ser modelado de diversas maneiras, e o seu grau de complexidade depende das características do modelo. Em sua formulação clássica, o PDE é modelado como um problema de programação quadrática convexa, sendo que a resolução deste modelo já foi amplamente abordada na literatura, como nos trabalhos de Samed (2004), Rodrigues (2007), Balbo et al. (2012) e Stanzani et al. (2014). Apesar da exploração exaustiva na literatura, o PDE clássico, proposto por Steinberg & Smith (1943), não inclui em sua modelagem outras características intrínsecas às unidades geradoras e ao sistema elétrico, dentre as quais pode-se citar os efeitos de ponto de carregamento de válvula, as perdas na transmissão e restrições de segurança relativas ao fluxo máximo de potência nos ramos do sistema.

Neste trabalho, a atenção será dirigida a três variantes do PDE clássico, obtidas considerando as características supracitadas:

- Problema de Despacho Econômico com Ponto de Carregamento de Válvula (PDE-PV): obtido a partir do PDE ao introduzir na função objetivo termos que modelam os pontos de válvula das unidades geradoras. Estes termos tornam a função objetivo não-convexa e não-diferenciável;
- Problema de Despacho Econômico com Perdas e Ponto de Carregamento de Válvula (PDE-PPV): o qual é obtido a partir do PDE-PV, ao introduzir na restrição de atendimento de demanda termos quadráticos que modelam as perdas. Os coeficientes da função de perdas são os coeficientes-B do sistema para algum estado da rede, de modo a representar a transmissão de maneira implícita;
- Problema de Despacho Econômico com Ponto de Carregamento de Válvula e Representação da Transmissão (PDE-PV-RT): obtido a partir do PDE-PV, de modo a considerar a representação explícita da transmissão. Neste caso, a restrição de atendimento de demanda é substituída pelas equações de balanço de potência ativa por barra do sistema. Além disso, acrescentam-se restrições de fluxo, relativas ao fluxo máximo de potência nos ramos. Os controles reativos são fixados e as variáveis de controle estão relacionadas à potência ativa (geração ativa e ângulos de tensão).

Todos os problemas dos quais este trabalho trata (PDE-PV, PDE-PPV e PDE-PV-RT) são, portanto, caracterizados pela não-convexidade e não-diferenciabilidade. Devido ao inconveniente causado pela não-diferenciabilidade, métodos clássicos de otimização não podem ser aplicados diretamente para a resolução destes problemas. Isto porque tais métodos usualmente empregam informações relacionadas ao gradiente e à matriz hessiana das funções, a fim de calcular as direções de busca. Além disso, a não-convexidade é outro fator que limita o uso de métodos determinísticos, uma vez que estes usualmente convergem para mínimos locais. Por estes motivos, tais problemas e outras variantes do PDE-PV foram predominantemente explorados na literatura utilizando métodos heurísticos ou heurísticas hibridizadas com métodos determinísticos, como os trabalhos de Walters & Sheble (1993), Bhagwan Das & Patvardhan (1998), Sinha et al. (2003), Samed (2004), Victoire & Jeyakumar (2004), Victoire & Jeyakumar (2005), Ravi et al. (2006), Rodrigues (2007), Oliveira et al. (2008), He et al. (2008), Arul et al. (2011), Yasar & Özyön (2011), Hardiansyah (2013).

Assim, tendo em vista a escassez de abordagens determinísticas para resolução do PDE-PV, o objetivo deste trabalho é apresentar um método de otimização determinístico, para a sua resolução. A metodologia proposta consiste de um procedimento que utiliza o método primal-dual de pontos interiores barreira logarítmica modificada, bem como duas estratégias auxiliares – suavização hiperbólica e correção de inércia.

O método de pontos interiores aqui apresentado foi desenvolvido com base nos trabalhos de métodos primais-duais propostos por Kojima et al. (1989a), Monteiro et al. (1990) e Mizuno (1994), e no procedimento previsor-corretor apresentado nos trabalhos de Mehrotra (1992) e Wu et al. (1994), além de se explorar a esparsidade da matriz do sistema de direções, conforme proposto por Karmarkar (1984).

Além disso, o método de pontos interiores desenvolvido utiliza a função barreira modificada, apresentada no trabalho de Polyak (1992) e que veio, posteriormente, a se tornar parte dos métodos de rescalamento não-linear (Polyak & Teboulle, 1997; Polyak, 2002; Griva & Polyak, 2006). Visando tornar o algoritmo mais robusto e que o mesmo convirja para mínimos locais, desenvolveu-se ainda uma estratégia de convergência através da correção da inércia da matriz do sistema de direções, mediante a investigação dos trabalhos de Shanno & Vanderbei (2000), Benson et al. (2004), Nocedal & Wright (2006) e Pinheiro (2012). Ademais, uma técnica de suavização hiperbólica de funções baseada em Junior (2010), Souza (2010) e Xavier & Xavier (2011) também é apresentada como forma de tratar a não-diferenciabilidade da função objetivo do PDE-PV. O método previsor-corretor primal-dual de pontos interiores com as estratégias apresentadas foi, então, implementado em Matlab e aplicado à resolução do PDE-PV, PDE-PPV e PDE-PV-RT.

O presente trabalho encontra-se dividido da seguinte forma: no Capítulo 2 são apresen-

tados alguns aspectos da geração termelétrica, no Brasil e no mundo, a fim de destacar a relevância que os problemas de despacho econômico assumem. No Capítulo 3 descrevemos a formulação matemática dos problemas de despacho econômico com ponto de carregamento de válvula e suas variantes (PDE-PV, PDE-PPV e PDE-PV-RT), os problemas a ele relacionado, bem como uma revisão da literatura sobre os métodos para a sua resolução. No Capítulo 4 é feita uma revisão dos métodos de otimização usualmente empregados em otimização não-linear e que apresentam forte relação com este trabalho. O Capítulo 5 dedica-se ao desenvolvimento de um método Primal-Dual Previsor-Corretor Barreira Logarítmica Modificada (PDPCBLM) para a resolução do PDE-PV. No mesmo capítulo, é apresentada uma estratégia de correção de inércia, que visa auxiliar o método na geração de direções de descida. Ainda no Capítulo 5, apresenta-se a técnica de suavização hiperbólica, utilizada no tratamento da função objetivo do PDE-PV. No Capítulo 6, aplicamos o método desenvolvido na resolução do PDE-PV e suas variantes. Finalmente, no Capítulo 7, são apresentadas as conclusões e delineadas algumas propostas de melhoria e continuidade deste trabalho. Os apêndices A, B e C incluem, respectivamente, o método para cálculo dos coeficientes-B, uma revisão teórica sobre otimização e os bancos de dados dos sistemas elétricos utilizados.

Capítulo 2

Geração Termelétrica

Desde os primórdios, a humanidade busca maneiras de evoluir e melhorar sua qualidade de vida. Essa busca incessante levou o homem a atingir patamares antes inimagináveis, através de conhecimentos que foram produzidos desde o início de sua existência. Dentre esses conhecimentos, a noção de energia, antes puramente empírica e de compreensão extremamente limitada, se consolidou e adquiriu caráter científico. Naturalmente, assim que percebeu o potencial energético existente na natureza, o homem passou a buscar meios para utilizá-lo, de modo a atender suas necessidades de sobrevivência.

Ao longo do tempo, o homem descobriu novas formas de energia e de como aproveitá-la. Na antiguidade, ela era obtida essencialmente através do sol e queima de biomassa (energia térmica e luminosa), moinhos, rodas d'água e força motriz animal (formas de energia mecânica). Entretanto, o consumo de energia ainda era relativamente baixo se comparado aos padrões de hoje. Foi com o advento da máquina a vapor, durante a Revolução Industrial que a forma de obtenção de energia mudou significativamente; o esforço muscular humano e/ou animal passou a ser progressivamente substituído por máquinas visando aumentar a produtividade em uma sociedade voltada cada vez mais para o capital. Essa transição, contudo, foi marcada por uma intensa dependência dos combustíveis fósseis, em especial do carvão mineral, necessários para o funcionamento do maquinário e que passaram a substituir a energia advinda da biomassa.

Pouco tempo depois, o mundo assistiu a introdução da energia elétrica, quando em 1882 foram autorizadas as primeiras estações de geração elétrica, movidas a queima de carvão. A flexibilidade da energia elétrica fez com que ela logo se destacasse entre as demais formas de energia. Desde então, diversas nações têm utilizado grandes parcelas de suas reservas de combustíveis fósseis para produção de eletricidade, bem como pesquisado o desenvolvimento de novas formas de obtenção da mesma – hidrelétricas, fissão nuclear, turbinas eólicas e células fotovoltaicas (Smil, 2004).

O desenvolvimento da energia elétrica teve grande impacto na vida humana. Na indústria, os motores elétricos significaram uma grande revolução e logo substituíram as antigas máquinas a vapor dotadas de diversas engrenagens e correias e que, conseqüentemente, apresentavam maiores chances de defeitos e acidentes. Nesse mesmo período também se desenvolveu o motor a combustão que, em conjunto com a energia elétrica, permitiu o surgimento dos primeiros automóveis, elevando a importância do petróleo enquanto combustível fóssil. Todos estes fatos culminaram na formação da sociedade contemporânea, marcada pelo alto consumo energético, em especial na forma de energia elétrica proveniente da queima de combustíveis fósseis. A Figura 2.1 apresenta a evolução da potência produzida através de alguns equipamentos ao longo do tempo.

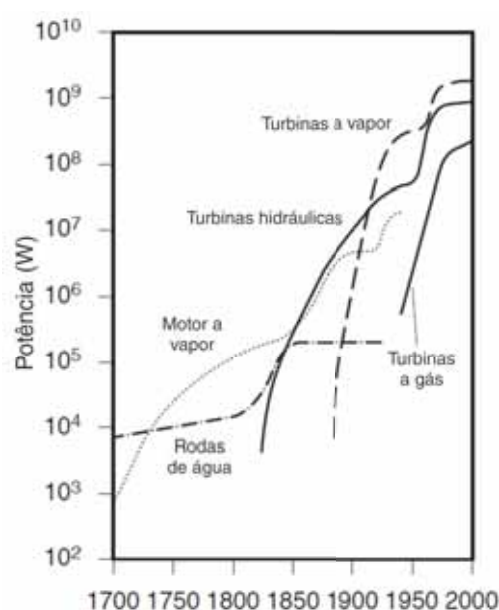


Figura 2.1: Potência gerada pelos diversos tipos de equipamentos.
Fonte: Smil (2004)

As consequências do consumo desenfreado dos combustíveis fósseis são várias. A primeira é o fato de que as reservas conhecidas de combustíveis fósseis estão se esgotando de maneira acelerada, impactando significativamente sobre o preço e disponibilidade dos mesmos. A segunda é de cunho ambiental: ainda que os combustíveis fósseis fossem recursos infindáveis, sua extração e queima podem causar impactos ambientais e para a vida humana, que vão de derramamentos de petróleo e aumento das concentrações de gases responsáveis pelo efeito estufa ao aumento da incidência de doenças, principalmente respiratórias.

Diante desse cenário, a questão dos custos envolvidos na geração de energia em termelétricas, a partir da queima de combustíveis fósseis, assume um papel socioeconômico de grande relevância. Assim, o objetivo deste capítulo é abordar aspectos da geração de eletricidade, em especial a termelétrica, que servirão como pano de fundo para os modelos

a serem apresentados no próximo capítulo.

2.1 Eletricidade e suas formas de geração

A energia elétrica pode ser gerada de diversas maneiras, através da conversão de outras formas de energia. Algumas formas de geração já estão bem consolidadas, enquanto outras ainda requerem pesquisas a fim de se tornarem viáveis. A seguir, apresenta-se algumas dessas formas de geração.

2.1.1 Geração termelétrica

É a forma mais comum de geração de energia elétrica no mundo. O funcionamento das usinas termelétricas consiste na queima de um combustível fóssil, como o carvão, óleo ou gás natural, visando transformar a água em vapor com o calor gerado na caldeira. Esse vapor, em alta pressão, movimenta a turbina que, por sua vez, aciona o rotor do gerador elétrico, convertendo energia mecânica em energia elétrica. A energia é, então, conduzida através de cabos, dos terminais do gerador até o transformador elevador, a fim de ter a tensão gerada aumentada para transmissão. Quando chega aos centros de consumo, a tensão é reduzida nos alimentadores primários do sistema de distribuição, e depois reduzida para níveis de consumo (alimentadores secundários). Por fim, na usina, o vapor é condensado, fechando o ciclo. Uma discussão mais detalhada acerca da geração termelétrica será realizada na Seção 2.2.

2.1.2 Geração nuclear

Pode ser considerada um tipo de geração termelétrica, uma vez que o calor produzido no processo de fissão nuclear é utilizado para produzir vapor, como nas usinas termelétricas convencionais. O combustível utilizado para geração do calor é que a difere das usinas termelétricas tradicionais: o urânio-235. A fissão é iniciada pelo bombardeamento com nêutrons e continua espontaneamente, numa reação em cadeia, uma vez que na divisão de um átomo de urânio-235 ocorre a liberação de 2 a 3 nêutrons, que atingirão outros núcleos de urânio-235. O minério de urânio é encontrado em pouca quantidade, de modo que é necessário um longo processo de extração e beneficiamento, até que ele possa ser utilizado nas usinas. Na América Latina, a única mina de urânio em funcionamento está em Caetité, na Bahia.

Dentre suas vantagens, estão: a operação das usinas é livre de emissões diretas de gases do efeito estufa, lançam poucos resíduos na atmosfera e o processo é bem conhecido do ponto de vista teórico. Por outro lado, como desvantagens, a geração nuclear tem altos custos de construção e operação devido aos sistemas de segurança necessários, não há solução definitiva sobre a alocação do resíduo radioativo e, em caso de acidentes, os resultados podem ser catastróficos.

Outra forma de geração de energia nuclear que vem sendo pesquisada é através da fusão nuclear, o mesmo fenômeno que ocorre no Sol, e que poderia gerar uma quantidade muito grande de energia limpa, por utilizar o hidrogênio como combustível. As dificuldades relacionadas à geração de energia nuclear por fusão residem no fato de que o processo ainda consome mais energia do que libera. Associado a isso, tem-se o problema da contenção eletromagnética do plasma gerado no processo de fusão, que exige equipamentos de tecnologia sofisticada.

2.1.3 Geração hidrelétrica

Consiste em utilizar a vazão dos rios e seus desníveis para a geração de energia. Uma usina hidrelétrica é composta, fundamentalmente, pela barragem, pelo sistema de captação e adução de água, casa de força, canal de fuga, que conduz a água captada até o leito do rio, e um vertedouro para evitar o comprometimento da segurança da usina, caso a água ultrapasse o limite máximo de volume armazenado. A Figura 2.2 ilustra de modo simplificado um sistema de potência com uma usina hidrelétrica.

De modo geral, a água é conduzida até a casa de força através de canais, a fim de passar por uma turbina hidráulica, onde a energia hidráulica é convertida em energia mecânica e, por conseguinte, em energia elétrica. O restante do processo de produção da energia elétrica é semelhante ao das usinas termelétricas.

Quanto ao reservatório, existem dois tipos: acumulação e fio d'água. Os de acumulação se caracterizam pelo alagamento de uma grande área, que é represada com auxílio da barragem, podendo estocar água para períodos de seca. Devido a essa característica, são construídas a montante das demais usinas, de modo a regular a vazão de água que fluirá para usinas a jusante. Já as usinas a fio d'água, salvo algumas exceções, em geral não represam água, de modo que possuem baixa capacidade regulatória. Este tipo de usina, contudo, pode represar a água para criar um maior desnível entre os espelhos d'água a montante e jusante, uma vez que a geração hidrelétrica depende da vazão e da altura de queda.

Apesar de causarem impactos ambientais durante suas construções, usinas hidrelétricas

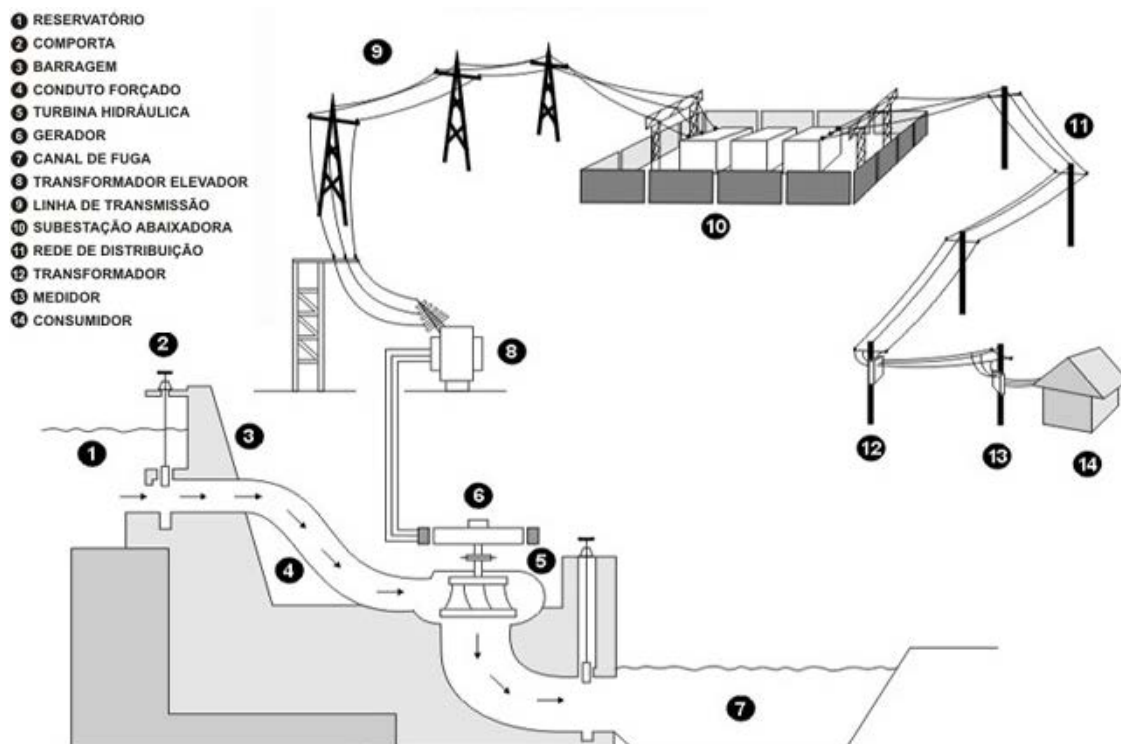


Figura 2.2: Diagrama de funcionamento de um sistema de potência com geração hidrelétrica.

Fonte: Furnas (2014)

são consideradas formas limpas de geração de eletricidade. Para tanto, porém, é necessária a retirada da matéria orgânica no local da construção da usina, já que a decomposição dessa matéria no fundo dos reservatórios pode liberar gases estufa. A eficiência de conversão da energia potencial em energia elétrica varia de 90% a 95%. No Brasil, a geração hidrelétrica é a mais desenvolvida e de maior proeminência. Dados de 2011 indicam que o Brasil era o segundo maior produtor de energia elétrica através desse tipo de geração (Tabela 2.1).

2.1.4 Geração eólica

De acordo com a ANEEL (2008), a energia eólica é caracterizada pela captação da energia cinética decorrente da migração de massas de ar causada por variações térmicas. O vento, “combustível” dessa forma de energia, faz girar as pás da turbina, produzindo a conversão de energia mecânica em eletricidade, através de um aerogerador. A geração de eletricidade a partir da energia eólica está sujeita a diversos fatores. Dentre eles, podemos citar:

- a densidade do ar, já que a energia cinética de um corpo depende de sua massa;
- a área de varrimento do rotor que determina quanta energia a turbina pode captar;
- a distribuição da pressão no rotor, já que conforme o ar se aproxima do rotor, sua

Tabela 2.1: Produção de hidreletricidade

Produtores	TWh	% do total mundial
China	699	19,6
Brasil	428	12,0
Canadá	376	10,5
Estados Unidos	345	9,7
Rússia	168	4,7
Índia	131	3,7
Noruega	122	3,4
Japão	92	2,6
Venezuela	84	2,3
Suécia	67	1,9
Restante do Mundo	1054	29,6
Mundo	3566	100,0

Fonte: *International Energy Agency* (2013)

pressão aumenta pois o rotor é uma barreira para a passagem do vento;

- Lei de Betz, segundo a qual um aerogerador somente pode converter menos de 16/27 (59%) da energia cinética em mecânica;
- variações do vento ao longo do tempo;
- tempo de início da geração do aerogerador, uma vez que ventos com baixa velocidade não têm energia suficiente para o acionamento das máquinas eólicas (somente funcionam a partir de uma determinada velocidade).

2.1.5 Geração solar

Tendo em vista as diversas preocupações ambientais, outra forma de energia renovável que também tem sido discutida é a solar fotovoltaica. Para a geração de energia elétrica a partir da energia solar, é necessária a utilização de elementos semicondutores fotossensíveis, cujos elétrons se excitam na presença de luz solar (ANEEL, 2008). Para tanto, usualmente são utilizadas células fotovoltaicas feitas de silício monocristalino.

Outra forma de produzir eletricidade a partir da energia solar é a criação de usinas solares. Este tipo de usina é composta por uma série de espelhos móveis, visando refletir a luz solar para um receptor. Tubos contendo água são aquecidos pela luz refletida, produzindo vapor que movimenta uma turbina.

As vantagens da energia solar são, basicamente, o fato de não haver consumo de combustível, de ser uma forma de energia limpa, de requerer pouca manutenção, além de evitar a volatilidade dos preços em mercados de combustíveis fósseis. Por outro lado, é uma forma de geração que somente funciona durante o dia, ainda é pouco eficiente e o custo

de instalação é elevado.

2.1.6 Bioeletricidade

De acordo com a ANEEL (2008), a bioeletricidade é uma forma de energia limpa e renovável, obtida em sistemas de cogeração a partir da biomassa, que pode ser de resíduos da cana-de-açúcar, restos de madeira, carvão vegetal, dentre outros. No Brasil, ainda que em pequena escala, a bioeletricidade é obtida essencialmente a partir do bagaço da cana-de-açúcar.

O Plano Nacional de Energia 2030 aponta que, em 2007, a potência instalada no Brasil para a produção de eletricidade a partir da biomassa da cana era de 2822MW, em mais de 250 usinas. Entretanto, a tecnologia de geração utilizada compreende ciclos de contrapressão, com caldeiras de baixa pressão e baixa eficiência (Brasil, 2007). A recuperação de parte da palha da cana também poderia contribuir para elevar a produção de bioeletricidade no Brasil.

2.2 Geração termelétrica: Brasil e mundo

Dentre as formas de geração apresentadas na Seção 2.1, as que mais têm sido estudadas através de modelos de despacho são a termelétrica, a hidrelétrica e a eólica. Nesta seção, apresentamos alguns aspectos da geração termelétrica no Brasil e no mundo.

2.2.1 Geração termelétrica no Brasil

O Brasil é um país com grande quantidade de recursos hídricos, de modo que ao longo dos anos conseguiu explorar boa parte desse potencial para geração de eletricidade, através de usinas hidrelétricas. As vantagens das hidrelétricas são, basicamente, os baixos custos operacionais tendo em vista que a água é um recurso gratuito, o fato da água ser uma fonte renovável de energia, a não produção de resíduos poluentes e a conversão relativamente eficiente da energia potencial gravitacional em energia elétrica nas usinas.

De acordo com Matsumura (2003), a reforma do setor elétrico ocorrida na década de 90 no Brasil visava criar incentivos ao investimento privado. Com isso, esperava-se que um número expressivo de termelétricas entrassem no mercado, uma vez que seus custos de produção eram competitivos quando comparados aos custos de novas hidrelétricas. Tais termelétricas seriam fundamentais para redução do risco de déficits energéticos futuros.

Entretanto, em 1999, concluiu-se que os investimentos em termelétricas esperados não estavam ocorrendo, levantando-se a possibilidade de um futuro racionamento de eletricidade. As causas para essa falta de investimento estavam vinculadas, essencialmente, às incertezas regulatórias, à variação cambial no preço dos combustíveis, bem como o risco hidrológico do sistema elétrico nacional.

As usinas hidrelétricas requerem planejamento uma vez que seu “combustível”, a água, é um recurso relacionado à sazonalidade das chuvas. Devido à natureza caótica dos fenômenos meteorológicos, o planejamento inadequado das hidrelétricas pode ocasionar a redução do nível dos reservatórios e, conseqüentemente, a quedas drásticas na geração de eletricidade. Essa foi a situação vivenciada pelo Brasil em 2001, durante o racionamento de energia, ocorrido tanto pela falta de planejamento quanto pela insuficiência dos investimentos em geração e transmissão.

Além da questão do planejamento, outras desvantagens das hidrelétricas são: os investimentos de grande monta e o tempo necessário para sua construção; a morte de várias espécies da fauna e flora, bem como a ruptura da cadeia alimentar e agressão às condições de abrigo, alimentação e reprodução animais; a necessidade de reassentamento de núcleos urbanos e ribeirinhos que se situam na zona afetada pela construção da usina; a perda de patrimônio cultural, bens históricos, arqueológicos e paisagísticos; e a distância das hidrelétricas aos centros consumidores, causando grandes perdas de energia na transmissão. Soma-se a isso o fato que, segundo o Plano Nacional de Energia (PNE) 2030, os prazos para obtenção das licenças ambientais tornam-se cada vez mais longos (Brasil, 2007).

Tendo em vista a crescente demanda e as experiências oferecidas pela crise pela qual passou, o Brasil tem buscado ampliar e diversificar sua matriz energética elétrica. Uma das possibilidades, em curto prazo, para efetuar a expansão do setor elétrico e suprir a demanda é a instalação de usinas termelétricas tradicionais, que funcionam através da queima de diesel ou gás natural. Daí o fato de que, nos anos 2000, o governo introduziu o Programa Estratégico Emergencial de Energia Elétrica (PEEE), dentro do qual se situava o Programa Prioritário de Termelétricas (PPT).

Dentre os objetivos do PEEE, estavam a expansão da oferta de energia, a maximização da produtividade das fontes geradoras instaladas e a diversificação da matriz energética elétrica nacional, de modo a reduzir a dependência do regime hidrológico (Brasil, 2001). O PPT visava ampliar a participação das termelétricas tradicionais no parque energético, aumentando de 12,2% em 2001 para 19,4% em 2003. A participação percentual deveria diminuir para 18,1% em 2005, não pela redução da produtividade das termelétricas, mas sim devido ao aumento proporcional de outras fontes.

A principal dificuldade para a efetivação das ações do PPT estava relacionada ao preço

do gás e do petróleo, cotado em dólar norte-americano e sujeito às operações de câmbio, e como reajustá-lo visando as tarifas do consumidor, fixadas em moeda nacional. Esse impasse veio a ser resolvido através de uma medida provisória que, pouco tempo depois, foi transformada em lei (Brasil, 2001).

O cenário previsto pelo plano decenal de expansão de energia 2022, é de que a capacidade instalada para produção de energia elétrica irá se expandir no Brasil, com destaque para a ampliação percentual das fontes renováveis de energia – biomassa, PCH e eólica. A concretização do plano, entretanto, está sujeita à obtenção de licenças prévias ambientais. Caso contrário, uma expansão concentrada em projetos termelétricos, preferencialmente movidos a gás natural, poderá constituir alternativa de atendimento à demanda, frente eventuais atrasos dos projetos (Brasil, 2007).

Naturalmente, como toda forma de geração de eletricidade, a implantação de termelétricas apresenta vantagens e desvantagens. Suas principais desvantagens estão relacionadas ao elevado custo dos combustíveis fósseis, bem como a emissão de gases do efeito estufa e material particulado, que podem causar danos à saúde humana, à fauna e à flora. Em contrapartida, as usinas termelétricas também possuem aspectos vantajosos. Dentre eles, pode-se citar o fato de que podem ser construídas próximas dos centros consumidores, evitando grandes perdas da energia elétrica produzida. Além disso, sua construção é mais simples e rápida quando comparada a hidrelétricas. No que tange ao aspecto social, unidades termelétricas podem contribuir para a introdução da energia elétrica em regiões isoladas, não atendidas pelo sistema interligado (Stanzani, 2012).

A geração termelétrica a carvão sempre foi pouco expressiva no Brasil. Uma das razões para isso são as características do carvão nacional, de baixa qualidade, de modo que sua aplicação para fins de geração de energia elétrica é bastante limitada. Devido a esse motivo, muitos estudos e prospecção desse recurso foram interrompidos. Mesmo assim, o Brasil é o país com a décima maior reserva do mundo.

Quanto ao gás natural, as reservas nacionais eram vistas como pouco expressivas para o atendimento do potencial de mercado, situação essa que mudou com a importação de gás boliviano. Devido às vantagens da geração termelétrica a gás, propôs-se que metade do gás importado da Bolívia fosse empregado na geração de energia elétrica. Essa decisão também foi incentivada pela crise de energia em 2001. Dessa forma, o ritmo de consumo nacional de gás natural tem crescido nos últimos anos.

De acordo com a ANEEL (2005), no Brasil, a principal função das usinas termelétricas é a de atendimento da demanda de pico do sistema elétrico e o suprimento de eletricidade para municípios e comunidades isoladas. Caso a hidrologia não seja favorável, as usinas termelétricas podem operar na base de carga, de modo que sua capacidade tende a se

aproximar do fator de disponibilidade máximo. Por outro lado, em situações de hidrologia favorável, o despacho das usinas termelétricas tende a ser mínimo, porém limitado pelas condições operativas/comerciais mínimas.

2.2.2 Geração termelétrica no mundo

A realidade brasileira, contudo, difere significativamente da maioria das outras nações. Ao redor do mundo, a principal forma de geração de energia elétrica para suprir a base da demanda ainda é através de termelétricas tradicionais, conforme indica a Figura 2.3.

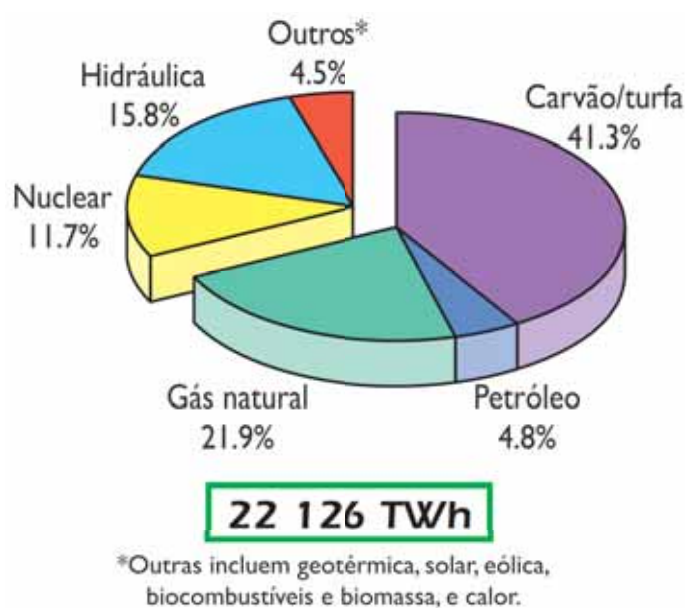


Figura 2.3: Participação das diversas formas de geração de eletricidade, no mundo, em 2011.

Fonte: *International Energy Agency* (2013)

Nota-se, portanto, que as formas de geração térmica tradicionais, através de carvão, petróleo e gás natural somam mais de 65% da geração de eletricidade mundial. Isto significa que, do ponto de vista econômico, estratégias que permitam a redução do custo dos combustíveis queimados são de grande valia.

No que diz respeito à geração termelétrica, o petróleo perdeu grande parte de sua importância, cedendo espaço para o carvão e, principalmente, para o gás natural. Este fato fica evidenciado, considerando que em 1973 o petróleo era o combustível responsável pela geração de 24,6% da eletricidade no mundo e, em 2011, por apenas 4,8%. Em contrapartida, o gás natural aumentou sua participação, passando de 12,2% para 21,9%, ao passo que o carvão aumentou de 38,3% para 41,3%, no mesmo período (*International Energy Agency*, 2013).

A justificativa para esse acontecimento é simples: com a crise do petróleo na década de

70, o carvão assumiu maior expressividade no cenário de geração de eletricidade. Fontes alternativas, como o gás natural, também passaram a se mostrar como uma possibilidade viável para substituição das usinas movidas a derivados do petróleo. A partir daí, a capacidade instalada de termelétricas a óleo foi reduzida e diversas plantas foram modificadas para empregar outros combustíveis, principalmente gás natural. Com o encarecimento do petróleo, esse tipo de usina passou, na maioria dos países, a ser utilizada para suprir a demanda de pico (ANEEL, 2005).

Na produção de energia elétrica, apesar das diversas pressões ambientais, o carvão ainda é o principal insumo utilizado para geração de eletricidade, e essa realidade não deverá mudar muito cedo, mesmo com a introdução e desenvolvimento de fontes renováveis. Isso porque o carvão é a fonte de energia primária mais abundante no planeta. Dados do PNE 2030 apontam que, apesar de concentradas principalmente em três regiões (Ásia e Oceania, Eurásia e América do Norte), as reservas provadas de carvão são suficientes para suportar o consumo mundial por mais de 160 anos (Brasil, 2007). Ademais, os preços do carvão apresentam volatilidade significativamente menor do que outros combustíveis fósseis.

Nos países desenvolvidos tem-se verificado uma maior restrição à expansão da geração termelétrica a carvão, devido às questões ambientais, em especial as emissões de dióxido de carbono. Contudo, essa realidade não é acompanhada por diversos outros países, principalmente aqueles em desenvolvimento, como a China, a Índia e a Rússia.

O segundo combustível mais utilizado para produzir energia elétrica é o gás natural. Segundo a ANEEL (2005), o desenvolvimento das turbinas a gás ocorreu depois da Segunda Guerra Mundial, o que significa que esse tipo de forma de geração é recente. Apesar disso, é uma forma de geração de eletricidade que apresentou grande expansão nos últimos anos, conforme se verifica na Tabela 2.2.

	1973	2003	Δ ao ano
Oferta Total (energia primária)	979,1	2244,1	2,8%
Geração de Energia Elétrica	160,0	468,6	3,7%
Cogeração	50,9	275,4	5,8%
Geração de Calor	0,7	87,7	17,5%
Outros Usos	96,1	220,5	2,8%
Uso Final	671,4	1191,9	1,9%

Fonte: Brasil (2007)

Devido à crise do petróleo da década de 70 e com a introdução das usinas de ciclo combinado, o gás natural tornou-se um recurso estratégico e economicamente atrativo. De acordo com Samed (2004), as turbinas a gás são compactas, leves, baratas e de instalação mais simples, além de se encontrarem em estágio maduro de confiabilidade e eficiência.

A principal dificuldade relacionada à utilização do gás natural em turbinas operando em ciclo simples é a baixa eficiência do ciclo termodinâmico e a perda de energia nos gases de exaustão. Por isso, as usinas mais modernas que utilizam gás natural operam em ciclo combinado: os gases de escape da turbina a gás, devido à temperatura, aquecem a água transformando-a em vapor, assim como ocorre numa usina termelétrica convencional. Dessa forma, combinam-se os ciclos de turbina a gás e turbina a vapor. Normalmente, usinas de ciclo simples apresentam rendimento de até 35% enquanto as de ciclo combinado podem apresentar rendimento superior a 45% (Brasil, 2007).

2.3 Aspectos técnicos e econômicos da geração termelétrica

Os conceitos apresentados nessa seção se baseiam no trabalho de Samed (2004). Segundo Miller (1987), a relação ou taxa entre a entrada de combustível e a saída de energia elétrica em uma unidade térmica, medida sob diversas situações de carga, permite a obtenção de curvas que expressam a eficiência das mesmas. Tais curvas são denominadas curvas de entrada-saída.

Comumente, as curvas de entrada-saída são idealizadas como convexas e suaves, como a Figura 2.4. Em geral, cada unidade geradora possui uma curva de entrada-saída, que é obtida através de cálculos de projetos ou testes de taxa de calor. Entretanto, se testes de taxa de calor são utilizados, os pontos encontrados para determinar a curva não resultam em uma curva suave.

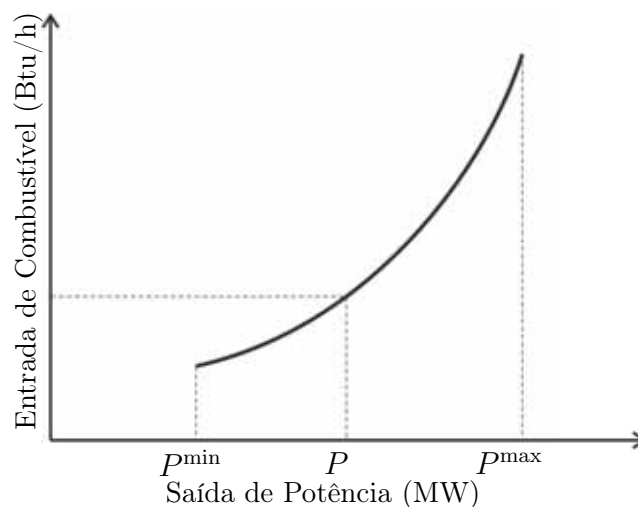


Figura 2.4: Curva de entrada-saída idealizada.
Fonte: Samed (2004)

A fim de calcular o custo de combustíveis empregados por uma unidade geradora, multiplica-

se a entrada de combustível, expressa em Btu/h, pelo equivalente custo de combustível, em \$/Btu, em que \$ é a unidade monetária.

Tendo em vista que o carregamento das unidades geradoras afeta o custo dos combustíveis, pode-se considerar que existe uma função que relaciona o custo e a potência gerada. Outros fatores que afetam o custo dos combustíveis são a combinação de usinas hidráulicas e térmicas, as variações na carga e a compra e venda de energia. Neste trabalho, o foco principal está nas usinas termelétricas, e a questão dos custos envolvidos na aquisição de combustíveis será abordada no próximo capítulo.

Capítulo 3

Despacho Econômico

Neste capítulo, apresenta-se o Problema de Despacho Econômico (PDE) e suas variantes. É feita, ainda, uma revisão bibliográfica desse problema e suas estratégias de resolução. O PDE é um problema de otimização que permite minimizar o custo dos combustíveis empregados na geração termelétrica tendo em vista as restrições operativas e, portanto, é de grande valia para os operadores do sistema elétrico.

3.1 Geração termelétrica: custos e perdas

O problema de despacho econômico, proposto por Steinberg & Smith (1943), pode ser definido como o processo de alocação ótima das demandas de eletricidade entre as unidades geradoras disponíveis, de modo que as restrições operacionais sejam satisfeitas ao mesmo tempo em que se minimizam os custos de geração.

O PDE clássico é modelado como um problema de otimização matemática não-linear, cujas restrições estão relacionadas essencialmente aos limitantes de potência gerada para cada uma das unidades geradoras e à satisfação da demanda. Modelos mais complexos do PDE podem incluir outras restrições, como restrições de transmissão, relacionadas ao fluxo de potência nos ramos do sistema.

A questão econômica da geração, entretanto, não é recente. De acordo com Happ (1977), o despacho econômico data do início dos anos 1920, ou até mesmo antes, quando engenheiros preocupavam-se com a alocação econômica da geração ou como dividir adequadamente a carga entre as unidades geradoras disponíveis.

Até o período anterior ao ano de 1930, vários métodos para despacho eram utilizados. Dentre eles, destaca-se o carregamento por ordem de mérito, que se baseava no carrega-

mento sucessivo das unidades mais eficientes (Happ, 1977). Isto é, segundo esse critério para despacho, solicita-se potência do gerador mais eficiente até que este atinja o limite de geração e, então, se passa para o próximo gerador mais eficiente. A falha deste método para minimizar os custos reside no fato de que se desconsidera que a eficiência de uma unidade geradora é uma função da potência gerada.

No início de 1930, entretanto, o método supracitado passou a ser substituído pelo critério dos custos incrementais iguais, devido aos resultados mais econômicos que este último apresentava. A utilização do custo incremental considera que a eficiência é função da potência gerada, e que o ponto de menor custo é aquele em que os custos incrementais de todas as unidades geradoras são iguais. Isto é, a ideia do método dos custos incrementais é que o carregamento das unidades geradores seja realizado de modo que o próximo aumento incremental de carga no sistema deve ser feito alocando-o na unidade geradora com menor custo incremental, o qual é determinado medindo a derivada da curva de custo.

Por volta de 1931, a ideia do critério dos custos incrementais iguais já estava bem estabelecida, de modo que se concluiu que para que a geração fosse econômica, os custos incrementais de todas as unidades geradoras deveria ser igual. No intuito de compreender do ponto de vista teórico essa escolha para as unidades geradoras, um passo foi dado no trabalho de Steinberg & Smith (1934), quando investigaram e mostraram que este critério era válido para a obtenção do custo mínimo, no caso de duas unidades geradoras. Com o propósito didático, a demonstração desse fato será reproduzida abaixo. Para tanto, é necessário assumir as hipóteses apresentadas por Steinberg & Smith (1934):

- As curvas de entrada-saída são contínuas e diferenciáveis;
- As derivadas de tais curvas são contínuas;
- O valor das taxas incrementais (derivadas) sempre aumenta conforme a saída aumenta.

Critério dos Custos Incrementais Iguais para 2 geradores

Mostrar-se-á que a entrada mínima para uma saída obtida através da combinação de 2 unidades geradoras é obtida quando as taxas incrementais são iguais. Para tanto, considera-se a seguinte notação:

O_t	é saída total das duas unidades (valor fixo);
O_1	é saída da unidade número 1;
O_2	é saída da unidade número 2;
I_t	é entrada total das duas unidades;
I_1	é entrada da unidade número 1;

I_2 é entrada da unidade número 2;

Conforme visto na Seção 2.3, pode-se assumir que a entrada é uma função da saída desejada, isto é, $I_1 = F_1(O_1)$ e $I_2 = F_2(O_2)$ são curvas de entrada-saída para as unidades geradoras.

Por hipótese, as derivadas $\frac{dI_1}{dO_1}$ e $\frac{dI_2}{dO_2}$ são contínuas e aumentam conforme O_1 e O_2 aumentam, respectivamente. Portanto, pode-se concluir que as funções I_1 e I_2 são convexas¹. Dessa forma a entrada total, $I_t = I_1 + I_2$, é uma função convexa. A saída total é expressa por $O_t = O_1 + O_2$.

O objetivo, portanto, é minimizar I_t de modo que a saída total seja obtida. Isto é, devemos ter $O_2 = O_t - O_1$. Logo, $I_t = F_1(O_1) + F_2(O_t - O_1)$ é uma função apenas de O_1 . Tendo em vista a convexidade de I_t , tal função assume valor mínimo quando $\frac{dI_t}{dO_1} = 0$.

Desde que $\frac{dI_t}{dO_1} = \frac{dI_1}{dO_1} + \frac{dI_2}{dO_2} \cdot \frac{dO_2}{dO_1}$, pela regra da cadeia tem-se que $\frac{dI_t}{dO_1} = \frac{dI_1}{dO_1} + \frac{dI_2}{dO_2} \cdot \frac{dO_2}{dO_1}$. Além disso, $\frac{dO_2}{dO_1} = \frac{d}{dO_1}(O_t - O_1) = -1$, de modo que $\frac{dI_t}{dO_1} = \frac{dI_1}{dO_1} - \frac{dI_2}{dO_2}$. Desde que para obtenção do mínimo deve-se ter $\frac{dI_t}{dO_1} = 0$, segue que $\frac{dI_1}{dO_1} - \frac{dI_2}{dO_2} = 0$ se, e somente se, $\frac{dI_1}{dO_1} = \frac{dI_2}{dO_2}$.

Entretanto, os autores destacam no mesmo artigo que caso as hipóteses apresentadas não sejam válidas, pode ser que exista mais de um mínimo dentro dos limites de operação, de modo que a computação de todos é necessária para concluir qual é o melhor. A compreensão acerca desse critério veio a ser aprofundada pelos mesmos autores, em 1943 (Steinberg & Smith, 1943).

A questão das perdas na transmissão também foi outra preocupação que desde cedo passou a ser considerada no despacho de geração. Em Steinberg & Smith (1934) é apresentada uma forma de considerar as perdas em termos da eficiência incremental. Entretanto, o cálculo das perdas era lento e tinha que ser repetido para cada situação nova do sistema.

Um grande avanço no sentido de computar as perdas foi obtido por George (1943). Em seu trabalho, o autor apresentou uma forma de calcular as perdas, bem como hipóteses subjacentes a esses cálculos. A expressão quadrática e as hipóteses por ele apresentadas são, com poucas modificações, utilizadas até os dias de hoje. O trabalho de George veio a ser aperfeiçoado, nas duas décadas seguintes, em uma série de quatro artigos produzidos por Kron. Nestes trabalhos, Kron apresentou uma modelagem das perdas de modo que as hipóteses em que se baseava eram bastante claras, originando a representação quadrática das perdas através dos coeficientes-B, também conhecida como *Fórmula de Kron*. Outros

¹A definição de convexidade de funções é apresentada no apêndice B. Este resultado decorre de um teorema de Análise, segundo o qual uma função real diferenciável em um intervalo aberto (a, b) é convexa se, e somente se, sua derivada f' é crescente em (a, b) .

métodos para tratar as perdas além dos coeficientes-B foram desenvolvidos, mas essa fórmula para representação das perdas continuou sendo a predominante (Happ, 1977).

Com a introdução e resolução dos problemas de fluxo de carga em computadores, no final dos anos 1950, a representação das perdas também evoluiu. Antes representadas por formas quadráticas, passaram a ser expressas através de equações não-lineares e não-convexas, dependentes das magnitudes e ângulos de tensão nas barras do sistema. O fluxo de carga visa determinar as magnitudes e ângulos de tensão nas barras, de modo que as equações de balanço de potência ativa e reativa sejam satisfeitas. Para tanto, considera a existência de três tipos de barra: barras de carga, barras de geração e barra *slack*. Em uma barra de carga, a potência ativa e reativa injetadas são conhecidas, enquanto a magnitude de tensão e o ângulo de tensão são desconhecidos; em uma barra de geração, a potência ativa injetada e a magnitude de tensão são conhecidas, enquanto a injeção de potência reativa e o ângulo de tensão devem ser determinados; por fim, na barra *slack*, a magnitude de tensão e o ângulo de tensão são conhecidos, restando determinar as injeções de potência ativa e reativa. Juntamente com a resolução do problema de fluxo de carga, surgiu outro problema de otimização denominado Fluxo de Potência Ótimo (FPO).

3.2 Despacho Econômico e problemas relacionados

O PDE pode ser entendido como um subproblema de determinados problemas mais complexos da área de sistemas de potência. Dentre eles, destacam-se o FPO, o *Unit Commitment* e o Despacho Hidrotérmico. A seguir, abordaremos de modo sucinto estes problemas.

3.2.1 Fluxo de Potência Ótimo

De acordo com Soler (2011), o objetivo da resolução do problema de FPO é determinar o ponto de operação do sistema elétrico, de modo a otimizar um critério da rede, ao mesmo tempo em que suas restrições físicas e operacionais não sejam violadas. O critério a ser otimizado pode variar, sendo alguns deles: minimização do desvio de tensão, minimização das perdas ativas na transmissão, minimização do corte de carga, maximização da capacidade de transferência simultânea entre áreas do sistema elétrico e, naturalmente, a minimização dos custos. Daí o fato de o PDE ser considerado um subproblema do FPO.

O FPO foi introduzido por Carpentier (1962) como uma generalização para o PDE. A principal contribuição do trabalho de Carpentier é a formulação do problema em bases matemáticas sólidas (Happ, 1977). Em seus trabalhos sobre FPO, Carpentier propôs a

solução do problema através da utilização das condições de otimalidade de Karush-Kuhn-Tucker.

De acordo com Samed (2004), o problema de FPO pode ser representado matematicamente por:

$$\begin{aligned} & \text{minimizar} && f(\mathbf{x}) \\ & \text{sujeito a} && \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ & && \mathbf{h}(\mathbf{x}) \leq \mathbf{0} \end{aligned} \tag{3.1}$$

em que:

- $\mathbf{x}$ é o vetor das variáveis de estado (magnitudes de tensão, ângulos de tensão, *taps* dos transformadores);
- $f(\mathbf{x})$ é a função objetivo, que traduz o critério a ser otimizado;
- $\mathbf{g}(\mathbf{x})$ são as equações de balanço de potência ativa e reativa;
- $\mathbf{h}(\mathbf{x})$ são as restrições funcionais e os limites operacionais.

O procedimento de resolução empregado por Carpentier foi o método de Gauss-Seidel. Entretanto, segundo Happ (1977), o comportamento de convergência mostrou-se difícil e mais irregular do que se pensava inicialmente. As formulações mais modernas do FPO são ainda mais complexas e também incluem variáveis discretas (Soler, 2011).

3.2.2 *Unit Commitment*

O PDE é um problema que supõe que as unidades geradoras ligadas e desligadas em cada usina são conhecidas *a priori*. Em contrapartida, os problemas de *Unit Commitment* devem não apenas determinar a potência a ser gerada pelas unidades, mas também aquelas que deverão ser ligadas e desligadas. Comumente, isso é feito em base horária, num horizonte de 24 horas, a fim de garantir que o despacho realizado seja capaz de atender às flutuações da demanda ao longo do dia. Para tanto, modelos de *Unit Commitment* incluem, além das restrições de satisfação de demanda presentes no Despacho Econômico, o custo de ligamento e desligamento das unidades geradoras na função objetivo (*start-up* e *shut-down*), além de restrições intertemporais (restrições de rampa).

A escolha do estado ligado/desligado usualmente é modelada através de variáveis binárias – *on* (1) e *off* (0) –, o que torna o *Unit Commitment* um problema de programação inteira mista (sendo ele NP-difícil, segundo Viana & Pedroso (2013)). Devido à sua natureza combinatória, esse problema freqüentemente é linearizado, a fim de que métodos de programação linear inteira, como *branch-and-bound* e *branch-and-cut*, possam ser utilizados. Métodos heurísticos também são empregados em sua resolução.

O Despacho Econômico, portanto, pode ser entendido como uma “fotografia”, isto é, um único intervalo do *Unit Commitment* multiperíodo, em que as unidades geradoras ligadas e desligadas são conhecidas *a priori*.

3.2.3 Despacho Hidrotérmico

O Despacho Hidrotérmico é a extensão do Despacho Econômico, de modo a incluir a geração hidrelétrica. Este tipo de despacho possui grande relevância para países como o Brasil, cuja produção de eletricidade é fundamentada na geração hidrelétrica.

Um sistema hidrotérmico de geração pode ser definido como o sistema elétrico em que usinas termelétricas e hidrelétricas funcionam em conjunto com o propósito de suprir a demanda. A resolução do despacho hidrotérmico é mais complexa que o de apenas termelétricas pois, conforme já destacado, a água é um insumo para geração de energia cuja disponibilidade é desconhecida em um horizonte de planejamento. Além disso, uma vez que muitas usinas estão conectadas à mesma cascata hidráulica, decisões tomadas em uma usina a montante pode afetar outras usinas a jusante.

No caso do despacho hidrotérmico, deve ocorrer o balanço entre a geração hidrelétrica e termelétrica, visando atender a demanda. Usualmente, esse balanço é expresso através da restrição de igualdade

$$\sum_{i \in \mathcal{T}} P_{G,i} + \sum_{j \in \mathcal{H}} P_{G,j} = D \quad (3.2)$$

em que:

$P_{G,i}$	é a potência ativa produzida pela unidade geradora i ;
$P_{G,j}$	é a potência ativa produzida pela unidade geradora j ;
$\mathcal{T}$	é o conjunto das unidades termelétricas de geração;
$\mathcal{H}$	é o conjunto das unidades hidrelétricas de geração;
D	é a potência demandada.

Diferentemente das termelétricas, os geradores hidrelétricos não possuem uma função de custo bem definida, uma vez que a água é um recurso gratuito. Por isso, diversos trabalhos são apresentados na literatura como forma de precificar a água, como a função custo de oportunidade apresentada em Vergílio (2011) que mede, economicamente, o quanto uma usina hidrelétrica pode perder quando não estiver operando em seu ponto de máxima produtividade.

3.3 Modelos Matemáticos do PDE

Como já foi dito, o PDE consiste em atribuir os níveis de produção adequados às unidades geradoras, de modo que a demanda seja suprida e as condições operacionais respeitadas, ao mesmo tempo em que se minimiza os custos associados ao consumo de combustível. Nesta seção, apresentamos os modelos de despacho econômico de interesse para esse trabalho. No que segue, iremos considerar que os modelos estão relacionados apenas à geração termelétrica.

3.3.1 Função Objetivo

Em sistemas elétricos, a demanda comumente é atendida por diversas unidades geradoras, cujos custos de combustíveis e manutenção são expressos por diferentes funções. Assim, a função de custo total é obtida pela soma dos custos individuais das unidades geradoras. Devido às dificuldades inerentes aos problemas não-lineares, a maioria das pesquisas realizadas para resolução de problemas de despacho econômico aproximam as funções de custo dos geradores através de uma função quadrática, isto é, uma função convexa e suave, de modo que a função objetivo de custo total assume a forma apresentada em (3.3):

$$C_T(\mathbf{P}_G) = \sum_{k \in \mathcal{G}} C_k(P_{G,k}) = \sum_{k \in \mathcal{G}} (a_k P_{G,k}^2 + b_k P_{G,k} + c_k) \quad (3.3)$$

em que:

- $C_T(\mathbf{P}_G)$ é a função objetivo de custo total;
- $P_{G,k}$ é a potência gerada pela unidade geradora k ;
- $\mathcal{G}$ é o conjunto das unidades geradoras;
- $C_k(P_{G,k})$ é a função custo individual da unidade geradora k ;
- a_k, b_k, c_k são os coeficientes de custo da unidade geradora k .

Entretanto, de acordo com Samed (2004), algumas formas diferentes foram utilizadas para representar as curvas de entrada-saída de unidades geradoras. Os dados obtidos em testes da taxa de calor foram ajustados de diversas maneiras: curvas polinomiais (incluindo quadráticas), segmentos de reta e segmentos quadráticos. Estas diferentes representações, naturalmente, resultam em distintas características de taxa de calor incremental e, conseqüentemente, diferentes funções de custo. A Figura 3.1 apresenta uma representação polinomial linear e linear por partes hipotética da taxa de calor incremental associada à saída de potência.

Na Figura 3.1, a linha contínua representa uma curva de taxa de calor incremental asso-

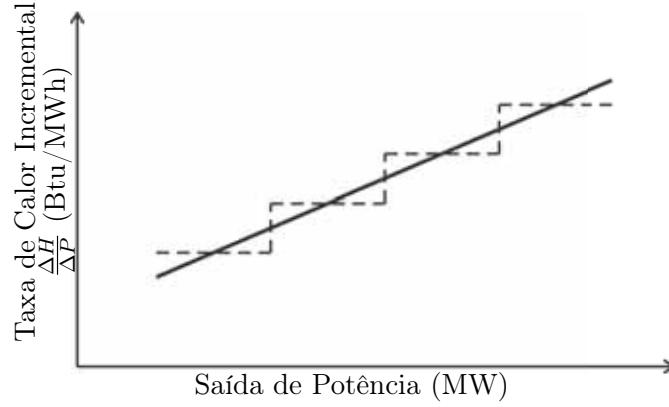


Figura 3.1: Possibilidades para representação das curvas da taxa de calor incremental.
Fonte: Samed (2004)

ciada a uma curva de entrada-saída quadrática (uma vez que a derivada de uma função quadrática é uma função linear). A linha tracejada representa uma curva de taxa de calor incremental associada a uma curva de entrada-saída linear por partes. Conclui-se, portanto, que as diferentes representações das taxas de calor incrementais conduzem a diferentes curvas de entrada-saída, de modo que métodos de resolução diferentes podem ser necessários para realizar o despacho. No caso supracitado, apenas o da linha contínua é caracterizado pela suavidade da função custo.

No caso de geradores com turbinas a vapor, entretanto, a representação das curvas de entrada-saída através de uma função quadrática não é a mais adequada. Nestes tipos de geradores existem válvulas de admissão de vapor, cuja abertura influencia na saída da unidade. De acordo com Happ (1977), no problema de despacho econômico clássico os custos incrementais são representados como monotonamente crescentes. Esta caracterização ignora as perdas que ocorrem na passagem de vapor através de válvulas parcialmente abertas. O ponto de carregamento de válvula pode ser definido como o ponto imediatamente anterior ao levantamento da próxima válvula. Em tais pontos, teoricamente, não ocorrem estrangulamentos e, portanto, não há perdas a eles associadas.

A fim de obter uma função de custo total mais representativa e que considere os efeitos dos pontos de carregamento de válvula, a equação (3.3) pode ser modificada de modo a obter a função apresentada em (3.4):

$$C_T(\mathbf{P}_G) = \sum_{k \in \mathcal{G}} C_k(P_{G,k}) = \sum_{k \in \mathcal{G}} [a_k P_{G,k}^2 + b_k P_{G,k} + c_k + |e_k \sin(f_k(\underline{P}_{G,k} - P_{G,k}))|] \quad (3.4)$$

em que:

- $C_T(\mathbf{P}_G)$ é a função objetivo de custo total;
- $P_{G,k}$ é a potência gerada pela unidade geradora k ;
- $\mathcal{G}$ é o conjunto das unidades geradoras termelétricas;

$C_k(P_{G,k})$ é a função custo individual da unidade geradora k ;
 a_k, b_k, c_k, e_k, f_k são os coeficientes da função custo da unidade geradora k ;
 $\underline{P_{G,k}}$ é o limite mínimo de saída de potência para a unidade geradora k .

Os termos senoidais que aparecem na função de custos totais (3.4) são responsáveis por modelar os pontos de carregamento de válvula. Ao mesmo tempo em que esta nova função objetivo é mais completa, ela também introduz novas dificuldades ao problema de otimização que deve ser resolvido: diferentemente da função quadrática apresentada em (3.3), a função de custos totais (3.4) é não-convexa e não-diferenciável. Estas características podem ser notadas na Figura 3.2, em que $\overline{P_{G,k}}$ denota o limite máximo de potência da unidade geradora k .

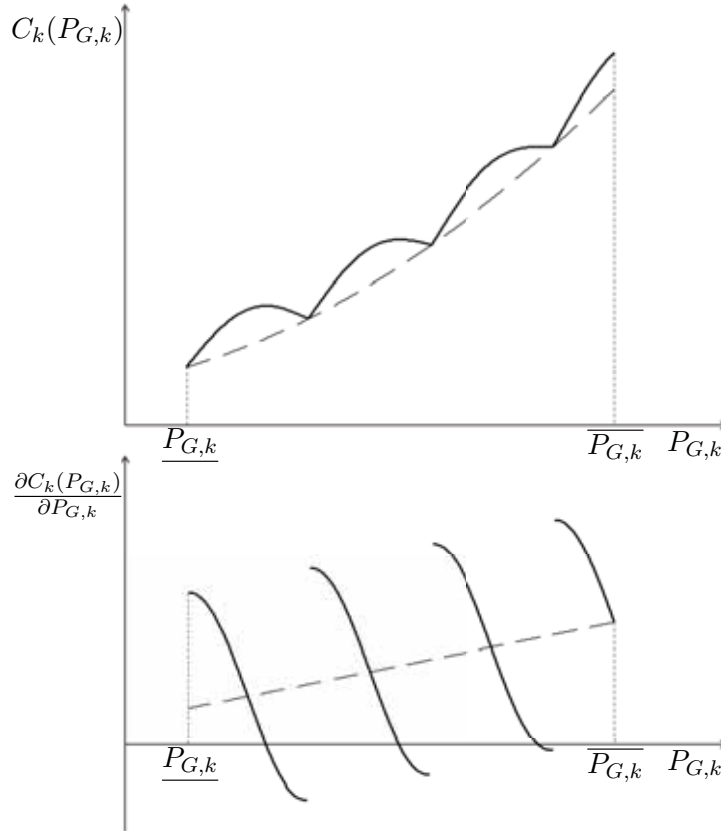


Figura 3.2: Funções custo e suas derivadas, com representação de ponto de carregamento de válvula (linha contínua) e sem representação de ponto de carregamento de válvula (linha tracejada).

Fonte: o autor.

Assim, como se pode notar, quanto à função objetivo do despacho econômico com ponto de carregamento de válvula, a principal dificuldade reside na não-diferenciabilidade da função objetivo, que não permite a aplicação direta de métodos clássicos de otimização. Os pontos de não-diferenciabilidade são aqueles em que

$$f_k(\underline{P_{G,k}} - P_{G,k}) = p\pi \text{ se, e somente se } P_{G,k} = \underline{P_{G,k}} - \frac{p\pi}{f_k}, \text{ com } p \in \mathbb{Z}.$$

3.3.2 Problema de Despacho Econômico com Ponto de Carregamento de Válvula

Tendo em vista a função custo apresentada em (3.4), o modelo de otimização para o Problema de Despacho Econômico com Ponto de Carregamento de Válvula (PDE-PV) é expresso por:

$$\text{minimizar } C_T(\mathbf{P}_G) = \sum_{k \in \mathcal{G}} C_k(P_{G,k}) = \sum_{k \in \mathcal{G}} [a_k P_{G,k}^2 + b_k P_{G,k} + c_k + |e_k \sin(f_k(\underline{P}_{G,k} - P_{G,k}))|] \quad (3.5a)$$

$$\text{sujeito a } \sum_{k \in \mathcal{G}} P_{G,k} = D \quad (3.5b)$$

$$\underline{P}_{G,k} \leq P_{G,k} \leq \overline{P}_{G,k}, \quad \forall k \in \mathcal{G} \quad (3.5c)$$

em que:

$C_T(\mathbf{P}_G)$	é a função objetivo de custo total;
$P_{G,k}$	é a potência gerada pela unidade geradora k ;
$\mathcal{G}$	é o conjunto das unidades geradoras termelétricas;
$C_k(P_{G,k})$	é a função custo individual da unidade geradora k ;
a_k, b_k, c_k, e_k, f_k	são os coeficientes de custo da unidade geradora k ;
$\underline{P}_{G,k}, \overline{P}_{G,k}$	são os limites mínimo e máximo, respectivamente, de saída de potência para a unidade geradora k ;
D	é a demanda de potência ativa.

A equação (3.5a) representa a função objetivo de custo total, incluindo efeitos de pontos de carregamento de válvula. Se esta função objetivo for substituída pela da equação (3.3), obtém-se o PDE clássico. A restrição de igualdade (3.5b) é a equação de balanço de potência ativa, isto é, visa garantir que as potências geradas consigam suprir a potência demandada. Por fim, as restrições de desigualdade em (3.5c) são restrições operacionais cujo objetivo é garantir o funcionamento das unidades geradoras dentro de seus limites mínimo e máximo de geração.

Cabe ressaltar que, neste modelo, a transmissão não é representada. O sistema é entendido de forma condensada, isto é, como se todas as unidades geradoras e cargas estivessem situadas em uma única barra e não houvesse nenhum tipo de perda.

3.3.3 Problema de Despacho Econômico com Representação das Perdas e Ponto de Carregamento de Válvula

Uma maneira de introduzir, de modo implícito, a transmissão do sistema elétrico no PDE-PV é a representação das perdas através da fórmula de Kron, que utiliza os coeficientes-B. Isto é, a potência gerada tem que ser suficiente para suprir a demanda, bem como as perdas, as quais são inerentes a qualquer sistema elétrico. Neste caso, obtém-se o seguinte problema, denominado Problema de Despacho Econômico com Perdas e Ponto de Carregamento de Válvula (PDE-PPV):

$$\text{minimizar } C_T(\mathbf{P}_G) = \sum_{k \in \mathcal{G}} C_k(P_{G,k}) = \sum_{k \in \mathcal{G}} [a_k P_{G,k}^2 + b_k P_{G,k} + c_k + |e_k \sin(f_k(\underline{P}_{G,k} - P_{G,k}))|] \quad (3.6a)$$

$$\text{sujeito a } \sum_{k \in \mathcal{G}} P_{G,k} = D + L_{\mathbf{P}^{\text{ref}}}(\mathbf{P}_G) \quad (3.6b)$$

$$\underline{P}_{G,k} \leq P_{G,k} \leq \overline{P}_{G,k}, \quad \forall k \in \mathcal{G} \quad (3.6c)$$

em que:

$L_{\mathbf{P}^{\text{ref}}}(\mathbf{P}_G)$ é a função quadrática que representa as perdas.

Apesar de incomum, utilizaremos a notação $L_{\mathbf{P}^{\text{ref}}}(\mathbf{P}_G)$ para denotar a função de perdas. Isto porque as perdas dependem das potências geradas ($\mathbf{P}_G$), bem como de uma situação possível de referência para o sistema elétrico ($\mathbf{P}^{\text{ref}}$), utilizada no cálculo dos coeficientes-B. Neste caso, as perdas são representadas pela fórmula de Kron, conforme a equação (3.7)

$$L_{\mathbf{P}^{\text{ref}}}(\mathbf{P}_G) = \sum_{k \in \mathcal{G}} \sum_{m \in \mathcal{G}} P_{G,k} B_{k,m} P_{G,m} + \sum_{k \in \mathcal{G}} B_{0,k} P_{G,k} + B_{0,0} \quad (3.7)$$

em que:

$B_{k,m}, B_{0,k}, B_{0,0}$ são os coeficientes-B relacionados às perdas, calculados a partir de uma situação factível para o sistema elétrico, com geração $\mathbf{P}^{\text{ref}}$.

Na literatura, diversos métodos são apresentados para o cálculo dos coeficientes-B. O cálculo destes coeficientes depende do conhecimento de um estado da rede elétrica, que pode ser determinado, por exemplo, através de um algoritmo de fluxo de potência. Neste trabalho, apresentamos uma maneira de realizar o cálculo dos coeficientes-B, baseada em Saadat (1999), a qual se encontra no Apêndice A. Outras formas de calcular tais coeficientes estão disponíveis nos trabalhos de Kothari & Dhillon (2006) e Dopazo et al. (1967).

3.3.4 Problema de Despacho Econômico com Ponto de Carregamento de Válvula e Representação da Transmissão

Até agora, os modelos apresentados ou não se preocupavam com a transmissão, ou representavam suas perdas de modo simplificado, através dos coeficientes-B. Isto significa que, apesar de realizarem o despacho de modo a suprir a demanda e, possivelmente, as perdas, o despacho realizado pode não ser factível no sentido de satisfazer as equações de balanço de potência por barra, expressas pelas Leis de Kirchhoff. Para tanto, é necessário que as equações de balanço de potência ativa sejam introduzidas como restrições do problema de otimização. Além disso, o fluxo de potência nos ramos do sistema não pode violar os limites mínimo ou máximo operacionais.

A seguir, apresenta-se um modelo de despacho, que visa incluir estas restrições. Tendo em vista que o problema está relacionado ao despacho de ativos, a parte reativa do problema será desconsiderada. Assume-se que as magnitudes de tensão estão fixas, e que o sistema possui suporte de reativos para manter as magnitudes de tensão constantes. Assim, o objetivo é determinar o valor da potência ativa a ser gerada pelas unidades, bem como os ângulos de tensão nas barras diferentes da *slack*.

Assim, no intuito de considerar a transmissão, de modo que a solução obtida se aproxime mais da realidade, tem-se o Problema de Despacho Econômico com Ponto de Carregamento de Válvula e Representação da Transmissão (PDE-PV-RT):

$$\text{minimizar } C_T(\mathbf{P}_G, \boldsymbol{\theta}) = \sum_{k \in \mathcal{G}} [a_k P_{G,k}^2 + b_k P_{G,k} + c_k + |e_k \sin(f_k(\underline{P}_{G,k} - P_{G,k}))|] \quad (3.8a)$$

$$\text{sujeito a } P_{G,k} - P_{L,k} = \sum_{m \in \Omega_k} P_{km}, \quad \forall k \in \mathcal{B} \quad (3.8b)$$

$$-f_{km} \leq P_{km} \leq f_{km}, \quad \forall k \in \mathcal{B}, \forall m \in \Omega_k \quad (3.8c)$$

$$\underline{P}_{G,k} \leq P_{G,k} \leq \overline{P}_{G,k}, \quad \forall k \in \mathcal{G} \quad (3.8d)$$

em que:

$C_T(\mathbf{P}_G, \boldsymbol{\theta})$ é a função custo total, incluindo efeitos de ponto de carregamento de válvula;

$\mathbf{P}_G$ é o vetor de potências ativas geradas;

$\boldsymbol{\theta}$ é o vetor de ângulos de tensão;

$\mathcal{G}$ é o conjunto das unidades geradoras termelétricas;

$\mathcal{B}$ é o conjunto das barras do sistema;

Ω_k é o conjunto das barras vizinhas à barra k ;

$P_{G,k}$ é a potência ativa gerada na barra k ;

$P_{L,k}$ é a demanda de potência ativa na barra k ;

a_k, b_k, c_k, e_k, f_k	são os coeficientes de custo da unidade geradora k ;
P_{km}	é o fluxo de potência ativa da barra k para a barra m ;
f_{km}	é o valor absoluto do fluxo de potência ativa máxima permitido na linha km ;
$\underline{P_{G,k}}, \overline{P_{G,k}}$	são os limites mínimo e máximo, respectivamente, de saída de potência para a unidade geradora k .

Assumindo que o ramo km tenha transformador em-fase, que o lado do transformador seja a barra k , com relação de transformação $1 : a_{km}$, e que o lado da impedância série que representa as perdas do transformador seja a barra m , conforme a Figura 3.3², as expressões de fluxo de potência ativa entre essas barras são dadas por:

$$P_{km} = (a_{km}V_k)^2 g_{km} - a_{km}V_kV_m[g_{km} \cos(\theta_k - \theta_m) + b_{km} \sin(\theta_k - \theta_m)] \quad (3.9)$$

$$P_{mk} = V_m^2 g_{km} - a_{km}V_kV_m[g_{km} \cos(\theta_m - \theta_k) + b_{km} \sin(\theta_m - \theta_k)] \quad (3.10)$$

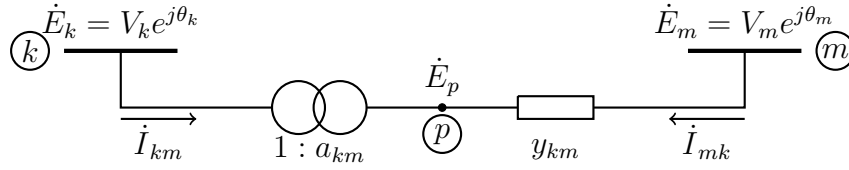


Figura 3.3: Modelo de um transformador em-fase.
Fonte: Monticelli (1983).

No caso do ramo km ser uma linha de transmissão, as expressões (3.9) e (3.10) continuam válidas, bastando tomar $a_{km} = 1$.

A equação (3.8a) é a função objetivo de custos de geração, incluindo o efeito de pontos de carregamento de válvula, (3.8b) são restrições relativas ao balanço de potência ativa para as barras do sistema, (3.8c) são restrições relativas ao limite mínimo e máximo para fluxo de potência ativa nos ramos do sistema e (3.8d) são restrições relativas aos limitantes mínimo e máximo de geração de potência ativa das unidades geradoras.

3.4 Revisão Bibliográfica

Nesta seção descrevemos alguns trabalhos que foram desenvolvidos visando solucionar o PDE e/ou problemas a ele relacionados, destacados na seção 3.2. O objetivo é apresentar

²Aqui, utilizamos o modelo de transformador apresentado em Monticelli (1983). Entretanto, os *taps* apresentados nos bancos de dados consideram transformadores em que a relação de transformação é $a_{km} : 1$, de modo que na implementação realizada o valor do banco de dados teve que ser invertido.

algumas metodologias – determinísticas, heurísticas e híbridas – que já foram aplicadas à resolução deste problema.

Conforme destacado, já em 1934, Steinberg & Smith (1934) apresentaram um método para alocar a geração de forma econômica entre as unidades geradoras, demonstrando com isso o critério dos custos incrementais iguais. Os autores, então, aplicaram o critério para divisão de carga em caldeiras e turbinas geradoras.

Posteriormente, Carpentier (1962) apresentou suas contribuições para formulação do PDE e FPO. Neste trabalho, o autor solucionou o problema de FPO através das condições de Karush-Kuhn-Tucker (KKT). O problema restrito é convertido em um problema irrestrito através da função lagrangiana. As condições de otimalidade foram aplicadas ao problema irrestrito, gerando um sistema não-linear que foi resolvido através do método de Gauss-Seidel.

No trabalho de Fahmideh-Vojdani & Galiana (1980), é apresentada uma maneira de resolver problemas de despacho econômico com função objetivo quadrática, com restrição de atendimento de demanda e limites de geração. Os autores apresentam um contra-exemplo mostrando que o procedimento de despacho realizado à época, fixando as variáveis que violam os limites no valor mínimo ou máximo, de acordo com a violação ocorrida, pode levar a um despacho de geração inadequado. Os autores ainda apresentam um teorema e um método que permite a realização do despacho de maneira ótima, bem como uma modificação do mesmo para o caso em que as perdas também são consideradas.

Em Aoki & Satoh (1984), três algoritmos para resolução do PDE clássico são apresentados, a saber: Programação Quadrática Paramétrica (PQP), Programação Quadrática Paramétrica Modificada (PQPM) e Programação Quadrática Recursiva (PQR). Enquanto o método PQP pode ser aplicado à função objetivo de custo aproximada por segmentos lineares, os outros dois métodos, PQPM e PQR, podem tratar de aproximações da função objetivo por segmentos quadráticos. Os autores aplicaram os métodos propostos a um sistema com 120 barras, 150 ramos e 31 geradores, sendo que utilizaram cinco segmentos para aproximar a função custo dos geradores. Além disso, os autores apresentam breves considerações a respeito dos autovalores da matriz de coeficientes-B utilizada.

No trabalho de Shaw et al. (1985), é apresentado um método para resolução de problemas de *unit commitment* e de despacho de grande porte. O modelo apresentado é caracterizado por ser um problema de programação inteira mista não-linear, com mais de 16800 variáveis inteiras, para um sistema que possui 100 unidades de geração termelétrica e 6 unidades de geração hidrelétrica, sob um horizonte de planejamento de uma semana. A formulação do problema inclui restrições de reserva e a função objetivo inclui os custos de *start-up* e *shut-down* para as unidades termelétricas. O problema é decomposto em problemas

menores, envolvendo um problema de programação dinâmica e um problema de controle ótimo.

Em Lee et al. (1988), um método para operação ótima de sistemas de grande porte, visando minimizar o custo dos combustíveis é apresentado. Os autores apresentam uma formulação ativa-reativa do FPO e propõem a decomposição do mesmo em dois problemas: o problema-P e o problema-Q. O problema-P é equivalente ao despacho econômico convencional, visando alocar de maneira ótima a geração de potência ativa entre os geradores. O problema-Q visa determinar, de maneira ótima, o despacho de potência reativa, bem como o estado de outras variáveis de controle, como os *taps* dos transformadores. O problema-P apresentado pelo autor, é semelhante ao PDE-PV-RT. Para resolver os problemas, o autor emprega o método do gradiente projetado.

Dois anos depois, Somuah & Khunaizi (1990) resolveram o problema de despacho econômico dinâmico. A formulação dos autores propunha a utilização de limites de rampa como restrições intertemporais, restrições de reserva girante e restrições de segurança visando evitar a sobrecarga de algumas linhas monitoradas e de maior importância. Para resolver o problema, os autores relaxaram as restrições de rampa, solucionando um problema de despacho econômico estático para cada intervalo, através de programação quadrática. Em seguida, um problema de programação linear é formado e resolvido, de modo a considerar as restrições de limite de rampa, mediante a linearização em torno das soluções do caso estático, baseando-se na decomposição de Dantzig-Wolfe.

Lin et al. (1992) apresentaram um método para resolver o despacho econômico em tempo real, utilizando uma formulação alternativa para a matriz Jacobiana considerando restrições do sistema. Uma aproximação para as perdas na transmissão é feita e as restrições de desigualdade são tratadas de modo direto, de modo que se o valor do redespacho violar algum limite, então ele é mantido nesse limitante crítico. O problema é resolvido pelo método de Newton-Raphson e as perdas na transmissão são consideradas de forma aproximada.

No trabalho de Walters & Sheble (1993), o problema de despacho econômico com ponto de carregamento de válvula e perdas representadas através de coeficientes-B é resolvido através de um algoritmo genético. Para minimizar a função, os autores empregaram os operadores genéticos de reprodução, *crossover* e mutação. Os autores destacam que a escolha adequada dos parâmetros é importante para a execução do algoritmo, a fim de que ele possa explorar adequadamente o espaço de busca. Em particular, o tamanho da população deve ser grande o suficiente para que haja informações genéticas que permitam uma varredura eficiente desse espaço.

No ano seguinte, Granville (1994) resolveu o problema de despacho ótimo de reativos,

utilizando um método primal-dual de pontos interiores. Para tanto, o autor fixou os controles de ativos, visando otimizar a injeção de reativos e o custo das perdas de potência ativa do sistema, nas variáveis relacionadas a potência reativa – magnitudes de tensão, *taps* e banco de capacitores. O método proposto foi aplicado a dois sistemas teste de grande porte: um com 1832 barras (o sistema Sul/Sudeste Brasileiro) e outro com 3467 barras (derivado do sistema de transmissão norte-americano).

No mesmo período, Wu et al. (1994) propuseram um método primal-dual previsor-corretor para resolução do problema de FPO. Técnicas de esparsidade são aplicadas para resolver o sistema derivado das condições necessárias de KKT. Os autores empregaram duas funções objetivo para o FPO: minimização dos custos de geração de ativos e minimização das perdas na transmissão. No caso da função custo, funções quadráticas foram utilizadas. Os testes realizados variaram entre sistemas de 9 e 2423 barras. Os resultados mostraram que o método previsor-corretor teve melhor desempenho que o método de ponto interior sem este procedimento, na resolução do problema de FPO.

Redes neurais artificiais foram utilizadas por Djukanovic et al. (1996) para resolução do despacho dinâmico em tempo real de unidades geradoras térmicas. A função objetivo de custos considerada era quadrática, e as restrições de limite de rampa para as unidades geradoras foram consideradas. Os autores também consideraram restrições ambientais relativas à emissão de poluentes atmosféricos. O método proposto foi testado em dois sistemas: IEEE-30 barras e o sistema da Iugoslávia, com 35 geradores e 180 barras. Os resultados mostraram que ganhos de tempo puderam ser obtidos, quando comparados a algoritmos tradicionais de despacho.

No trabalho de Irisarri et al. (1998) o problema de despacho econômico com restrições relativas à transmissão linearizadas, restrições intertemporais de rampa e restrições de reserva girante é modelado como um problema de programação quadrática convexa, e resolvido através de um método de pontos interiores. Os resultados obtidos a partir de três sistemas – um em que a topologia é desconsiderada, um de 31 barras e outro de 30 barras – mostraram que o número de iterações necessárias para resolver o sistema linearizado de KKT não varia muito com a dimensão do problema.

No mesmo ano, Torres (1998) resolveu o problema de FPO utilizando duas abordagens: pontos interiores e pontos não-interiores. No caso do método de pontos interiores, quatro variantes de alta ordem foram consideradas: o método previsor-corretor, o método de Newton composto perturbado, o método previsor-corretor com múltiplas correções e o método com correções múltiplas de centralidade. No caso do método de pontos não-interiores, as restrições de complementaridade são tratadas de modo que as condições de positividade estrita não são necessárias em cada iteração. As restrições de complementaridade são tratadas através de uma função de complementaridade não-linear.

Ainda naquele ano, Bhagwan Das & Patvardhan (1998) resolveram o problema multiobjetivo de despacho econômico, visando minimizar os custos de geração incluindo o efeito de pontos de carregamento de válvula, minimizar o índice de emissão de poluentes e maximizar o índice de margem de segurança. Para tanto, utilizaram uma abordagem heurística denominada SEEDSS (*Secure Economic Emission Dispatch with Stochastic Search*), que consistia de uma abordagem híbrida, incorporando o *simulated annealing* no processo de seleção do algoritmo genético. A restrição de satisfação de demanda também foi incluída, de modo que o método escolhia uma das unidades para ter geração dependente das demais. Os testes foram realizados nos sistemas IEEE-30 barras e IEEE-118 barras, a fim de mostrar a eficiência do algoritmo.

Um algoritmo de busca de colônia de formigas foi utilizado em Song et al. (1999) para resolver o problema de despacho econômico envolvendo geração de potência ativa e calor. Para tanto, as unidades geradoras são divididas em três tipos: geração de potência ativa convencional, unidades de cogeração e unidades geradoras de apenas calor. A função objetivo visa minimizar os custos para geração das demandas de potência ativa e calor. O algoritmo de busca de colônia de formigas se baseia no comportamento de formigas reais, que utilizam os feromônios para se comunicar e decidir o caminho a ser tomado. Os testes foram realizados para um caso de despacho econômico puro, com 6 unidades termelétricas, e para um caso com 4 geradores visando despachar potência ativa e calor.

El-Gallad et al. (2002) utilizaram o método de otimização por enxame de partículas para resolução do problema de despacho econômico com função custo quadrática e restrições de reserva girante e zonas de operação proibidas. Modificações foram aplicadas ao método de enxame de partículas convencional, a fim de reduzir as chances do algoritmo acabar “preso” em uma solução local. Os testes foram realizados em um sistema de 15 unidades geradoras, com 4 unidades que possuíam zonas de operação proibidas. Os resultados mostraram que o método proposto obteve soluções melhores do que outros trabalhos que utilizaram redes neurais.

Lin et al. (2002) desenvolveram um algoritmo de busca tabu aperfeiçoado, para a resolução do problema de despacho com função objetivo descontínua e não-diferenciável. O método funciona através de uma lista tabu, que contém um conjunto de movimentos proibidos para obtenção de pontos teste. Um critério do método permite a reativação de determinados movimentos no espaço de busca, quando satisfeito. Assim como em algoritmos genéticos, o método também emprega operadores de mutação e recombinação, baseados em probabilidade. O critério de parada é dado pelo número máximo de gerações. Os testes foram realizados em duas instâncias do PDE-PPV, envolvendo 3 e 30 unidades geradoras, a fim de mostrar a eficiência do método.

No trabalho de Sinha et al. (2003), oito variantes de técnicas baseadas em programação

evolutiva são aplicados ao PDE-PV, a fim de testar sua *performance*. Segundo os autores, técnicas de programação evolutiva se diferenciam dos algoritmos genéticos essencialmente em dois aspectos: eles utilizam os parâmetros de controle (valores reais), mas não sua codificação como nos algoritmos genéticos tradicionais, e seus operadores são, essencialmente, a mutação e a seleção, mas não o *crossover* como nos algoritmos genéticos. Os testes foram realizados para casos com 3, 13 e 40 unidades geradoras. Os dados deste trabalho se tornaram, posteriormente, referências para comparações de outros autores.

No ano seguinte, Samed (2004) apresentou um Algoritmo Genético Híbrido Co-Evolutivo (AGHCOE) para resolução de problemas de despacho. Neste trabalho, a autora visou resolver o PDE-PV através de um algoritmo genético. Tendo em vista os problemas que precisavam ser contornados – o tratamento das restrições, parada prematura do método e ajuste dos parâmetros – a autora optou por uma estratégia híbrida. A codificação foi feita em ponto flutuante ao invés da codificação binária, e as restrições tratadas separadamente da função objetivo, através de uma função adaptação. Os operadores de cruzamento e mutação foram inspirados em regras determinísticas, enquanto a seleção é realizada através do método da roleta. Tendo em vista a dificuldade no controle dos parâmetros associados ao AGH, a autora introduziu um algoritmo genético para controle de parâmetros, denominado Co-Evolutivo, que interagia com a população do AGH. O algoritmo foi implementado em Fortran, e testado em casos de PDE-PV com 3 e 13 geradores, bem como em instâncias de 6 geradores de problemas de despacho econômico e ambiental.

No mesmo ano, outra estratégia híbrida foi apresentada no trabalho de Victoire & Jeyakumar (2004). Os autores associaram o método de otimização por enxame de partículas a um método de programação quadrática sequencial. A otimização por enxame de partículas foi utilizada como otimizador principal, visando encontrar regiões de ótimo local, enquanto o método de programação quadrática sequencial foi utilizado para refinar as soluções localmente. No método de programação quadrática sequencial, utiliza-se um aproximante de Taylor de segunda ordem para a função objetivo e efetua-se a linearização das restrições. A matriz hessiana é determinada e atualizada através de uma fórmula empregada em métodos quasi-Newton e o tamanho do passo aceito é determinado pela utilização de uma função mérito. Simulações foram realizadas utilizando os dados de 3, 13 e 40 geradores do PDE-PV apresentados por Sinha et al. (2003). Os mesmos autores introduziram, no ano seguinte, um método semelhante, utilizando programação evolutiva como heurística no lugar da otimização por enxame de partículas (Victoire & Jeyakumar, 2005).

Park et al. (2005) resolveram o PDE-PV utilizando um algoritmo modificado de otimização por enxame de partículas. Além de resolverem casos em que a função objetivo representa os pontos de carregamento de válvula, resolveram casos alternativos de função

custo, como aquele em que a mesma é definida por segmentos quadráticos devido à utilização de múltiplos combustíveis. Assim como no trabalho de Bhagwan Das & Patvardhan (1998), a restrição de igualdade do problema é tratada reduzindo o grau de liberdade em um. Uma estratégia para modificação da posição de um indivíduo, caso este viole os limitantes de geração, é apresentada. Os testes foram realizados para os casos de 3 e 40 geradores.

Em Ravi et al. (2006), o PDE-PPV é resolvido utilizando um algoritmo baseado em programação evolutiva. Para tanto, uma população inicial de vetores pais é determinada e, sobre esta população, vetores filhos são determinados pela escolha do melhor filho gerado por mutação Gaussiana ou mutação de Cauchy. Em um ciclo externo do método, os coeficientes-B e a função de perdas são atualizados utilizando um algoritmo de fluxo de potência. Os testes foram realizados com sistemas de 6, 14 e 30 barras, mostrando a eficiência do método proposto. No caso de 6 barras, diversas situações de carga foram exploradas.

Nos trabalho de Sousa (2006) e Sousa et al. (2012), um método de pontos interiores, baseado na função barreira logarítmica modificada, foi utilizado para resolver o problema de FPO reativo. A vantagem da função barreira modificada sobre a barreira clássica é a possibilidade de utilização de pontos exteriores à região factível original, bem como a redução de problemas de mau condicionamento que ocorrem em métodos baseados na função barreira clássica. Para resolução do FPO, a autora associou ao problema original uma função lagrangiana. As condições necessárias de otimalidade foram aplicadas, resultando em um sistema não-linear resolvido pelo método de Newton.

Rodrigues (2007) apresentou e implementou um algoritmo cultural para resolução do PDE e do despacho econômico e ambiental. Esse algoritmo pode ser entendido como uma metaheurística inspirada no processo de evolução cultural da humanidade. O processo evolutivo é decorrente da aquisição e propagação da cultura (conhecimentos) ao longo das gerações. De acordo com a autora, este tipo de algoritmo é híbrido por definição, sendo que no trabalho a hibridização foi realizada com um algoritmo genético. Em um algoritmo cultural, os componentes principais são o espaço populacional e o espaço de crença, sendo este último responsável por armazenar o conhecimento adquirido e guiar o algoritmo para melhores regiões do espaço de busca. Para o PDE, casos de 3 e 13 geradores foram estudados, com função objetivo quadrática.

No ano seguinte, Coelho & Mariani (2008) apresentaram um algoritmo de otimização por enxame de partículas baseado em mecânica quântica e em poços de potencial de osciladores harmônicos. Os métodos de enxame de partículas baseados em mecânica quântica se inspiram em conceitos como a equação de Schrödinger e a distribuição de campo potencial. De acordo com os autores, o oscilador harmônico é um tipo bastante

comum de distribuição de potencial, e um dos modelos mais importantes em mecânica quântica. Resultados foram obtidos para o caso de 13 geradores do PDE-PV, utilizando o algoritmo com cinco configurações distintas de parâmetros.

Oliveira et al. (2008) resolveram o PDE-PPV multiperíodo, com restrições de limite de rampa, zonas de operação proibidas e limites de fluxo de potência ativa nas linhas de transmissão. Para tanto, utilizaram um algoritmo baseado em estratégias evolutivas, o qual empregava os operadores evolutivos de recombinação, mutação e seleção. O *fitness* é calculado como a soma do custo de produção e um termo de penalidade quadrática para a restrição de demanda. Os testes foram realizados para sistemas com 6, 13, 15 e 40 geradores.

No trabalho de Sousa (2008), o problema de FPO foi resolvido através da associação de métodos de pontos interiores e métodos de região de confiança. A autora utilizou os métodos de pontos interiores, do tipo primal-dual previsor-corretor para programação quadrática, na resolução dos subproblemas derivados do método de região de confiança. Os testes foram conduzidos para diversos sistemas, sendo que se concluiu que o número de iterações externas, do método de região de confiança, é relativamente baixo. Por outro lado, cada iteração de região de confiança requer a solução de dois subproblemas de programação quadrática, cuja solução pelo método de pontos interiores requer praticamente o mesmo número de iterações utilizadas pelo método de pontos interiores para programação não-linear na resolução direta do FPO. Entretanto, a autora enfatiza que o algoritmo proposto não objetiva competir em termos de tempo computacional com o método primal-dual de pontos interiores clássico, mas sim mostrar-se robusto para a resolução de problemas de FPO de convergência difícil. Este fato também foi verificado, uma vez que o algoritmo da autora, que combina regiões de confiança e pontos interiores, conseguiu convergir para a solução de sistemas críticos, enquanto os algoritmos de pontos interiores puros não o fizeram.

Selvakumar & Thanushkodi (2009) propuseram um algoritmo heurístico denominado Otimização por Enxame Civilizado (*Civilized Swarm Optimization – CSO*), que combina o Algoritmo de Civilização da Sociedade (*Society-Civilization Algorithm – SCA*) com Otimização por Enxame de Partículas. No SCA, os indivíduos são agrupados em pequenas sociedades, sendo que os indivíduos com melhores resultados em cada grupo atuam como líderes. Este algoritmo se baseia em dois fatos: uma civilização progride devido à cooperação entre suas sociedades e os indivíduos das sociedades melhoram devido à interação entre membros da sociedade. Os testes foram realizados em diversas instâncias do PDE, incluindo casos de grande porte com 200 geradores.

Yuan et al. (2009) utilizaram um método híbrido baseado em evolução diferencial para resolução do despacho econômico dinâmico com ponto de carregamento de válvula e res-

trições de limite de rampa. O método foi hibridizado com o r -algoritmo de Shor, que é utilizado ao final do processo para refinar a solução obtida pelo método de evolução diferencial. O r -algoritmo de Shor é um método baseado em subgradientes utilizado na minimização de funções não-diferenciáveis. Para tanto, utiliza-se um operador de dilatação do espaço, definido a partir de um vetor unitário. Os testes foram conduzidos para um sistema com 10 unidades geradoras, com um horizonte de planejamento de 24 horas, mostrando que o método permite a obtenção de soluções de boa qualidade.

No ano seguinte, Bhattacharya & Chattopadhyay (2010) apresentaram um método híbrido de evolução diferencial com otimização baseada em biogeografia. De acordo com os autores, a biogeografia é modelada a partir do processo de imigração e emigração de espécies entre ilhas, em busca de habitats mais convenientes. Um habitat é uma ilha (área) que é geograficamente isolada de outras ilhas. A otimização baseada em biogeografia utiliza dois mecanismos: migração e mutação. O método foi testado em quatro variantes de problemas de despacho, incluindo o PDE-PV e o PDE clássico.

Hemamalini & Simon (2010) apresentaram um método lagrangiano baseado na série de Maclaurin para resolução do PDE-PV multiperíodo com restrições de limite de rampa. O autor utiliza a série de Maclaurin para tratar o termo modular senoidal presente na função objetivo de custos com ponto de carregamento de válvula. Em seguida, o método lagrangiano é aplicado para a resolução do problema. A validação do método proposto foi realizado em um caso monoperíodo de 40 geradores, e em um multiperíodo com 5 geradores.

Khamsawang & Jiriwibhakorn (2010) utilizaram uma abordagem híbrida, combinando sequências de Sobol, enxame de partículas e busca tabu para a resolução de problemas de despacho, incluindo o PDE-PV. As sequências de Sobol produzem uma sequência de números entre 0 e 1, em que os sucessivos pontos gerados preenchem os espaços entre os pontos gerados previamente.

Malik et al. (2010) resolveram o PDE-PPV, bem como uma variante do PDE-PV considerando o sistema de transmissão, combinando Algoritmo Genético (AG) e o método de Newton. O objetivo dessa abordagem é utilizar os melhores aspectos dos dois métodos: a capacidade de busca pelo ótimo global e a convergência rápida quando o ponto está próximo do ótimo, respectivamente. Assim, o AG funciona como otimizador global, encontrando soluções quase ótimas. A geração determinada pelo AG, então, é utilizada como entrada para o algoritmo que utiliza o método de Newton.

Meng et al. (2010) introduziram um algoritmo de otimização por enxame de partículas com inspiração quântica. Neste método heurístico, o estado de uma partícula é determinado por bits quânticos e ângulos, ao invés da posição e velocidade das partículas, como nos

algoritmos por enxame de partículas tradicionais. O mecanismo de atualização básico do algoritmo é a evolução dos bits quânticos e ângulos, de modo que os bits quânticos preservem a condição de normalização sobre eles imposta. O método proposto foi testado na resolução de problemas-teste matemáticos e na resolução do PDE-PV.

No trabalho desenvolvido por Niknam (2010), o autor empregou um método de otimização por enxame de partículas adaptativo, que utiliza sistemas *fuzzy*, bem como o método de Nelder-Mead. O sistema *fuzzy* é utilizado para auxiliar no controle dos parâmetros da otimização por enxame de partículas. Já o método de Nelder-Mead foi utilizado como forma de busca local. Este método é usualmente empregado em otimização não-linear, e opera de modo a rescalar a região viável, baseado no comportamento local da função. A metodologia proposta foi empregada na resolução do PDE-PV.

Em Arul et al. (2011), os autores propõem a utilização de um algoritmo de busca harmônica para a solução do PDE-PV. A busca harmônica é uma metaheurística baseada no processo de desempenho musical que acontece quando um músico busca atingir um melhor estado de harmonia. A harmonia na música é o análogo do vetor de soluções, enquanto o comportamento de improvisação do músico é análogo aos esquemas de busca local e global. Neste algoritmo, os operadores básicos de otimização são a memória harmônica, o tamanho da memória harmônica, o número de improvisações e a taxa de ajuste do som. O método foi validado utilizando sistemas de 6, 14 e 30 barras.

Hemamalini & Simon (2011) utilizaram um algoritmo de colônia artificial de abelhas para resolução do problema de despacho econômico dinâmico com ponto de carregamento de válvula. Neste método heurístico, abelhas artificiais voam pelo espaço de busca multidimensional. Determinados tipos de abelha escolhem fontes de alimento baseadas na experiência própria, enquanto outras utilizam a experiência das companheiras de colméia, e ajustam suas posições. Outros tipos de abelha (exploradoras) voam e escolhem aleatoriamente as fontes de alimento, sem se basear na experiência. A quantidade de néctar encontrada na fonte de alimento representa o *fitness* da solução. Quando uma solução que representa uma fonte de alimento não apresenta melhora após determinado número de tentativas, então a fonte de alimento é abandonada e a abelha se torna exploradora. O método foi validado realizando o planejamento de 5 e 10 unidades geradoras, ao longo de 24 horas.

Selvakumar (2011) apresentou um método aperfeiçoado de entropia cruzada para a resolução do problema de despacho econômico dinâmico com ponto de carregamento de válvula. De acordo com o autor, o método de entropia cruzada se originou como um algoritmo para minimização de variância para estimar probabilidades de eventos raros. Posteriormente, ele foi aplicado em problemas combinatórios complexos. No artigo, sequências caóticas são utilizadas para melhorar o desempenho do método de entropia cruzada. Sem esta estraté-

gia, de acordo com o autor, o método de entropia cruzada pode sofrer significativamente com a parada prematura devido à perda de diversidade de soluções.

Nos trabalhos de Soler (2011) e Soler et al. (2013) é apresentada uma metodologia determinística de resolução do problema de FPO com variáveis de controle discretas. Para tanto, desenvolveu-se uma função que penaliza a função objetivo, quando as variáveis discretas não assumem os valores discretos possíveis. A inclusão da função de penalidade na função objetivo resulta em um problema de programação não-linear que possui apenas variáveis contínuas. A solução do problema modificado é equivalente à do problema original. Para a resolução do problema de programação não-linear, utilizou-se um método de pontos interiores com filtro. A abordagem mostrou-se eficiente para a resolução de problemas de FPO, ao ser testada nos sistemas IEEE 14, 30, 118 e 300 barras.

Yasar & Özyön (2011) empregaram um método híbrido envolvendo busca harmônica e o método de subgradientes modificado, desenvolvido por Gasimov, para resolução do PDE-PPV. O método de subgradientes modificado constitui a principal estrutura da abordagem híbrida, e funciona adicionando as restrições do sistema em uma função lagrangiana aumentada não-diferenciável, convertendo o problema restrito em um problema irrestrito. Este problema irrestrito é, então, resolvido pelo método de busca harmônica. De acordo com o autor, dentre as principais características do método de subgradientes modificado está a capacidade de atingir o gap de dualidade nulo para uma ampla quantidade de problemas não-convexos e a criação de uma sequência crescente de valores duais, assegurando a convergência.

Balbo et al. (2012) apresentaram um método primal-dual previsor-corretor de pontos interiores para resolução do PDE clássico. O método desenvolvido possui dois tipos de busca unidimensional: a busca de Fibonacci no passo previsor e a busca de Armijo no passo corretor. A ideia básica, segundo os autores, é utilizar uma busca unidimensional mais precisa no passo previsor (que utiliza aproximação de primeira ordem para o cálculo dos resíduos), e uma busca mais simples, porém robusta, no passo corretor (que utiliza uma aproximação de segunda ordem para o cálculo dos resíduos). Os autores utilizam o método apresentado para resolver instâncias de 3, 6 e 13 geradores do PDE, e incluem uma análise pós-otimização a respeito do preço da energia para os sistemas investigados.

Moradi-Dalvand et al. (2012) propôs um método denominado otimizador por busca em grupo rápida contínua (*Continuous Quick Group Search Optimizer – CQGSO*). De acordo com o autor, dois modelos gerais podem ser identificados na vida em grupo de animais e na busca por alimentos: compartilhamento de informações e um esquema produtor-usurpador. A primeira dessas características é comum em heurísticas do tipo enxame de partículas, em que cada indivíduo busca por alimento separadamente e então compartilha informações com os outros. No modelo produtor-usurpador, um membro que encontra

mais alimento que outros membros é chamado produtor. Outros membros do grupo são chamados usurpadores e seguem o produtor. O CQGSO é uma heurística inspirada no modelo produtor-usurpador, em que os membros do grupo são divididos em três tipos: produtor (aquele com melhor *fitness* no grupo), usurpador (membro que segue o produtor) e o dispersor (um membro que vaga pelo espaço do problema). A função do dispersor é aumentar as chances do grupo escapar de mínimos locais. Os testes foram conduzidos em várias instâncias do PDE, com características diversificadas: ponto de carregamento de válvula, restrições de limite de rampa, zonas de operação proibidas e perdas na transmissão.

Pinheiro (2012) resolveu o problema de FPO reativo, visando a minimização de perdas de potência ativa. Para tanto, utilizou um método primal-dual previsor-corretor baseado na função lagrangiana barreira logarítmica modificada. O autor ainda incluiu uma estratégia de ajuste cúbico para a função barreira modificada, bem como uma estratégia de convergência global, variante de Levenberg-Marquardt, a fim de auxiliar o método a gerar direções de descida. O sistema de KKT obtido ao aplicar as condições necessárias de otimalidade à função lagrangiana foi resolvido explorando a esparsidade da matriz, de modo semelhante à proposta desenvolvida no trabalho de Karmarkar. O método proposto mostrou-se eficiente na resolução do FPO.

No trabalho de Sajjadi et al. (2012), um método baseado em bom-senso para resolução do PDE-PV é apresentado. De acordo com o autor, este não é um método heurístico, uma vez que fornece uma única solução toda vez que o algoritmo é executado. O método consiste basicamente em utilizar os limites mínimo e máximo de geração, bem como as potências ativas geradas em que ocorrem os pontos de carregamento de válvula, para calcular a soma de todas as possíveis combinações de potência ativa envolvendo limitantes de geração ou pontos de carregamento de válvula de todos os geradores. A partir dessa lista de combinações, escolhe-se a de menor custo que satisfaça a demanda ou, caso isso não seja possível, as que estão mais próximas da potência demandada. Neste último caso, um reajuste de potência é realizado no gerador com menor custo incremental, a fim de satisfazer a demanda. As perdas são, então, calculadas, e a nova demanda passa a ser dada pela demanda anterior acrescida das perdas. O procedimento é, então, realizado novamente, até que a convergência seja atingida. Testes realizados com o sistema IEEE-30 barras mostram que a estratégia é eficiente quando comparada à solução obtida por diversas heurísticas, porém uma demonstração de que a solução encontrada é um mínimo não é delineada.

Em Stanzani (2012) e Stanzani et al. (2014) é apresentado um método primal-dual previsor-corretor de pontos interiores, a fim de resolver o problema multiobjetivo de despacho econômico e ambiental. A autora determinou soluções eficientes para o problema

utilizando os métodos da soma ponderada e ϵ -restrito. No caso do método ϵ -restrito, o problema de despacho econômico foi condicionado à restrição ambiental, limitada superiormente para níveis permissíveis de emissão. Os resultados obtidos demonstraram a eficiência do método, quando comparado a algoritmos genéticos, atávico híbrido e cultural.

Vishwakarma et al. (2012) empregou a heurística de *simulated annealing*, inspirada no processo natural de cristalização, isto é, resfriamento de metais, na resolução do PDE-PV. Este processo consiste no aumento da temperatura de um metal para um valor em que ele se funde e, em seguida, um resfriamento controlado é realizado, até a solidificação do material. O *simulated annealing* tem a capacidade de escapar de regiões de mínimo local através da incorporação de uma função probabilidade para aceitação ou rejeição de novas soluções. Para o cálculo dessa probabilidade, utiliza-se um parâmetro denominado temperatura, que é elevado no início do processo e reduzido gradativamente. De acordo com o autor, existem quatro parâmetros críticos na utilização do *simulated annealing*: a temperatura inicial, a temperatura final, a taxa de decremento da temperatura e o número de iterações em cada temperatura.

Yang et al. (2012) utilizou um algoritmo de vagalumes para resolução do PDE-PV. Este método se baseia no comportamento idealizado das características de brilho em vagalumes. A primeira delas é a de que não há distinção entre os vagalumes, isto é, um vagalume é atraído por outros vagalumes, independentemente do sexo. A segunda característica é que a atratividade é proporcional ao brilho, de modo que para quaisquer dois vagalumes, o menos brilhante se moverá na direção do mais brilhante. A atratividade é proporcional ao brilho e diminui conforme a distância aumenta. Se nenhum vagalume é mais brilhante que um vagalume em particular, o movimento é aleatório. A terceira característica é o fato de que o brilho ou intensidade de luz de um vagalume é afetado ou determinado pelo ambiente da função objetivo a ser otimizada. Para tratar das restrições do problema, o método utiliza uma função penalidade, semelhante às empregadas em métodos determinísticos.

No ano seguinte, Basu & Chowdhury (2013) utilizaram um algoritmo denominado *Cuckoo Search* na resolução de problemas de despacho econômico, incluindo geração eólica e compra de energia. A ideia básica deste algoritmo é baseada no comportamento parasitário obrigatório de algumas espécies de cuco em combinação com os padrões de comportamento de voo de Lévy apresentado por determinados tipos de pássaros e moscas de fruta. Vãos de Lévy podem ser entendidos como cadeias de Markov, com aplicações no estudo de movimentos animais em busca de alimento e outros fenômenos naturais. Muitas espécies de cucos adotam estratégias de reprodução agressivas, deixando seus ovos em ninhos de outros pássaros. Caso o pássaro descubra o ovo do cuco intruso, ele pode jogar os ovos fora ou abandonar o ninho para construir um novo em outro lugar. Desse modo, o algoritmo *Cuckoo Search* se baseia em três regras. A primeira é de que cada cuco bota um ovo de

cada vez, que é deixado aleatoriamente em algum dos ninhos disponíveis. A segunda é de que os melhores ninhos com ovos de alta qualidade (melhor solução) serão transferidos para a próxima geração. A terceira hipótese é de que o número de ninhos é fixo, e que cada ovo deixado por um cuco é descoberto com uma probabilidade no intervalo $[0, 1]$. Neste caso, o pássaro que construiu o ninho pode jogar o ovo fora ou abandonar o ninho e encontrar um novo lugar para construir um completamente novo. Os resultados da abordagem mostraram sua eficiência, quando comparada a outras heurísticas disponíveis na literatura.

Gonçalves et al. (2013) utilizaram um método híbrido, combinando algoritmos culturais e sistemas imunológicos artificiais para resolução do PDE-PPV. Os sistemas imunológicos artificiais herdaram determinadas características inspiradas no sistema imune biológico. O sistema imune biológico envolve agentes denominados anticorpos, que interagem com os outros visando impedir danos ao organismo causados por agentes invasores, ditos antígenos. No sistema imunológico artificial desenvolvido pelo autor, o principal mecanismo utilizado foi a seleção clonal, que consiste em selecionar para proliferar aquelas células que são capazes de reconhecer os antígenos. O algoritmo híbrido elaborado utiliza o espaço de crenças do algoritmo cultural para extrair informações dos anticorpos e utiliza este conhecimento como guia para gerar novos anticorpos e melhorar o desempenho do sistema imunológico artificial.

Benasla et al. (2014) propuseram um algoritmo de otimização espiral para resolução do PDE-PV, bem como do despacho econômico e ambiental. Esta metaheurística é inspirada em fenômenos espirais que ocorrem na natureza, em especial a espiral logarítmica que aparece em fenômenos como ciclones tropicais, tornados e galáxias. Segundo os autores, a abordagem apresenta as seguintes vantagens: poucas variáveis de controle, capacidade de busca local, resultados rápidos, uso fácil, estrutura simples e as fases de diversificação e intensificação ocorrem no mesmo processo. O método explora as possíveis combinações de matrizes de rotação envolvendo dois eixos. De acordo com os autores, o método apresenta diversificação nas fases iniciais e intensificação no centro da espiral nas fases finais.

Binetti et al. (2014) apresentaram um algoritmo determinístico baseado em leilão distribuído para resolução de problemas não-convexos de despacho econômico. O método utiliza técnicas de leilão e protocolos de consenso para resolução do problema. De acordo com o autor, a resolução da maioria dos problemas de despacho econômico não-convexo é realizada através de heurísticas de modo centralizado, o que não é viável no ambiente dos *smart-grids*. Daí a necessidade de utilização de algoritmos distribuídos que sejam imunes às variações topológicas. Cada unidade geradora calcula localmente valores utilizados como lances no mecanismo de leilão. Estas unidades enviam seus lances para suas vizinhas em um grafo de comunicação do sistema e que fornece o fluxo de informações

necessário. Um procedimento de consenso é utilizado para compartilhar os lances entre os agentes da rede e fechar o leilão. Como resultado, a distribuição de potência gerada entre as unidades é atualizada e o custo de geração minimizado. Os testes mostraram que a estratégia é viável, oferecendo boas soluções.

Em Borckmans et al. (2014), um algoritmo baseado em subgradiente Riemanniano é apresentado para resolução do PDE-PPV. Para tanto, o autor utiliza o fato de que a restrição de satisfação de demanda do PDE-PPV determina a equação de um elipsóide, que admite uma estrutura natural de variedade Riemanniana. Devido à não-diferenciabilidade da função objetivo, o autor propõe-se a utilizar os conceitos do cálculo generalizado de Clarke para determinar um método de resolução para o problema. A técnica desenvolvida explora a geometria da região factível do PDE-PPV, que é obtida pela porção do elipsóide delimitada pelas restrições de canalização de geração. Um subgradiente projetado é computado sempre que os iterandos se aproximam de um ponto de não-diferenciabilidade, visando gerar uma sequência de pontos factíveis. Para tanto, é adotado um parâmetro de proximidade que, de acordo com os autores, pode influenciar no funcionamento do método. Outro fato que o autor destaca é que o conjunto factível do PDE-PPV pode apresentar mais de uma componente conexa, porém isto não ocorreu em nenhuma das instâncias testadas.

Como se pode notar, diversos trabalhos tratam da resolução de problemas de despacho. Entretanto, quando a não-diferenciabilidade causada pelos pontos de carregamento de válvula se faz presente, a maioria das abordagens são heurísticas ou híbridas, com poucos enfoques determinísticos de resolução. Neste sentido, o objetivo deste trabalho é apresentar um método de otimização determinístico para resolução do PDE-PV e suas variantes. O método proposto, assim como a maioria dos métodos determinísticos, é essencialmente um minimizador local. Entretanto ele pode ser utilizado para a resolução direta do PDE-PV visando determinar mínimos locais, ou associado a algum outro tipo de método de busca global, como métodos heurísticos.

Capítulo 4

Métodos de Otimização Restrita

Neste capítulo são apresentados alguns métodos básicos utilizados na solução de problemas de otimização não-linear e que estão relacionados ao método de solução central deste trabalho. As principais definições e resultados da teoria de otimização não-linear podem ser encontradas no Apêndice B.

Abordam-se os problemas restritos e sua conversão em problemas irrestritos através da função Lagrangiana, bem como os métodos de penalidade e barreira. A parte final do capítulo é coroada com as ideias essenciais do método de rescalamento não-linear, e sua versão para a função barreira modificada.

4.1 Introdução à otimização

Problemas de otimização podem ser entendidos como a busca pela melhor solução possível, para determinado critério expresso por uma função objetivo, dentro de um determinado conjunto de possibilidades de solução (denominado conjunto factível). Estes problemas são encontrados em diversas áreas do conhecimento – como engenharia, economia, física, biologia e agronomia – e usualmente modelados como minimização ou maximização do valor da função objetivo.

Os primeiros rudimentos de otimização já podem ser notados na antiguidade, quando Euclides verificou que o retângulo de máxima área, entre todos aqueles que possuem o mesmo perímetro, é um quadrado. Entretanto, passos significativos para o desenvolvimento da teoria de otimização somente foram dados a partir da introdução do cálculo diferencial, por Newton e Leibniz.

Até a década de 1940, a teoria de otimização para várias variáveis ainda era pouco desen-

volvida. Com o advento do computador e a necessidade de alocação dos recursos durante a segunda guerra mundial, a otimização ganhou grande impulso. Nesse cenário, o desenvolvimento do método simplex, por Dantzig, em 1947, foi de fundamental importância. Desde então, a área de otimização passou a se desenvolver a um ritmo acelerado, impulsionada pelo aperfeiçoamento dos equipamentos computacionais e do conhecimento cada vez mais profundo de suas bases teóricas.

Nos anos 60 e 70, um campo de investigação cuja importância aumentou foi o do estudo da complexidade computacional. Com estes estudos, o interesse por algoritmos de tempo polinomial aumentou, e os algoritmos ditos “rápidos” passaram a ser aqueles para os quais havia garantia de resolução em tempo limitado superiormente por um polinômio relacionado ao tamanho do problema. O método simplex era, na prática, bastante eficiente e capaz de resolver problemas lineares de grande porte. Entretanto, Klee & Minty (1972) mostraram que o método simplex não possui complexidade de tempo polinomial, ao apresentar uma classe de problemas em que o método teria que visitar todos os vértices da região factível para atingir a solução ótima. Este fato incentivou a busca por métodos alternativos ao simplex.

O final da década de 50 e o início dos anos 60 também foram marcados pelo aparecimento dos primeiros métodos baseados em penalidade e barreira para resolução de problemas não-lineares restritos, fundamentados nos trabalhos de Frisch (1955) e Carroll & Fiacco (1961). De acordo com Wright (2004), os métodos de penalidade e barreira foram especialmente populares, ambos motivados pela minimização de uma função composta que reflete a função objetivo original bem como a influência das restrições. Os métodos modernos de pontos interiores estão intimamente associados aos métodos de barreira desenvolvidos na década de 60. Apesar de terem sido amplamente utilizados na década de 60, os métodos de barreira sofreram uma queda de popularidade a partir da década de 70 por vários motivos, em particular devido às preocupações com mau-condicionamento. À medida em que métodos de penalidade e barreira começaram a sair de cena, as abordagens dominantes passaram a ser baseadas diretamente nas condições de otimalidade para otimização restrita, especialmente nas propriedades da função Lagrangiana. Métodos de Lagrangiana aumentada se tornaram populares, e continuam até os dias de hoje.

No ano de 1979, Leonid Khachian propôs o primeiro algoritmo de tempo polinomial para programação linear, e que ficou conhecido como método dos elipsoides. Contudo, na prática, verificou-se que o desempenho do método dos elipsoides foi inferior ao do método simplex, fazendo com que ele fosse pouco utilizado. Em 1984, Karmarkar (1984) anunciou um algoritmo de tempo polinomial para programação linear, para o qual ele havia noticiado soluções com tempo 50 vezes mais rápidas que o método simplex. Essa nova metodologia se tornou o que denominamos métodos de pontos interiores.

A transformação sugerida por Karmarkar, contudo, era relativamente complexa, uma vez que requeria uma mudança no problema original. Este aspecto foi aperfeiçoado através dos trabalhos de Barnes (1986) e Vanderbei et al. (1986) que redescobriram, independentemente, a transformação projetiva desenvolvida por Dikin (1967) anos antes. Com esta transformação, o problema poderia ser resolvido utilizando sua forma padrão. O algoritmo de Dikin ficou conhecido como algoritmo afim, e sua convergência foi demonstrada nos trabalhos de Barnes (1986) e Vanderbei et al. (1986). Uma variante do algoritmo de Karmarkar, para resolução de problemas de programação linear com restrições de desigualdade, foi proposta por Adler et al. (1989), o qual ficou conhecido como Algoritmo Dual-Afim. Apesar do bom desempenho que estes métodos demonstraram e das transformações bem mais intuitivas e simples que empregavam quando comparadas à proposta originalmente descrita por Karmarkar, a prova da complexidade de tempo polinomial para os mesmos não foi obtida.

A complexidade de tempo polinomial somente veio a ser demonstrada quando foi proposta a incorporação de uma função barreira logarítmica ao problema de programação linear original, a fim de evitar que uma solução interior se aproximasse demais da fronteira, causando mau condicionamento, parada prematura do método ou obtenção de soluções insatisfatórias. Nesta classe de métodos, podem-se citar os métodos de pontos interiores que utilizam a função objetivo penalizada pela função barreira logarítmica, desenvolvida por diversos autores, como Renegar (1988), Megiddo & Shub (1989), Vaidya (1990) e Ye (1991), e os métodos de trajetória central propostos, por exemplo, em Monteiro & Adler (1989) e Gonzaga (1990). Ye (1991) aponta que, em todos estes métodos, é utilizada a técnica de escalamento e a resolução de um problema de mínimos quadrados ocorre a cada iteração. Em 1992, Mehrotra (1992) apresentou outro variante, o método da barreira logarítmica primal-dual previsor-corretor, em que a cada iteração é dado um passo previsor e um passo corretor, resultando em direções de busca melhores.

Os métodos de pontos interiores supracitados apresentaram bons resultados computacionais em problemas de grande porte de otimização linear. Entretanto, conforme Murray (1971) e Wright (1994), em problemas de otimização linear e, principalmente, em problemas não-lineares, de grande porte, os métodos de penalidade através da função barreira logarítmica apresentam o problema de mau condicionamento das matrizes hessianas à medida em que o método se aproxima da solução do problema.

Visando amenizar este problema, em 1992, Polyak (1992) propôs o método da função barreira modificada, que era a junção da função lagrangiana clássica e da barreira clássica. O objetivo deste método era evitar algumas dificuldades até então presentes nos métodos de pontos interiores que utilizavam a estratégia de barreira. As funções de barreira modificada apresentadas por Polyak visavam transformar um problema restrito em irrestrito,

resolvendo uma sequência de problemas irrestritos até atingir a solução ótima. Em 1996, uma variante do método de barreira modificada de Polyak foi proposto por Breitfeld & Shanno (1996), introduzindo-se uma estratégia de extrapolação quadrática para os termos logarítmicos. Os resultados mostraram que esta estratégia permitiu a convergência do método em situações que apenas a função barreira modificada não obtinha sucesso.

O trabalho de Polyak veio a ser estendido para uma teoria mais ampla, denominada rescalamento não-linear, através de uma série de trabalhos (Polyak & Teboulle, 1997; Polyak, 2002). Utilizando os conceitos desenvolvidos por Polyak e a teoria de rescalamento não-linear, que pode ser entendida como um método de pontos exteriores, Griva (2004) combinou os métodos de pontos interiores tradicionais, que são robustos e possuem resultados de convergência global, com os métodos de pontos exteriores, que possuem boas propriedades de convergência local. Os resultados mostraram que a combinação dos métodos levou a resultados superiores que apenas utilizar um método individualmente. Posteriormente, Griva & Polyak (2006) apresentaram um método primal-dual de rescalamento não-linear globalmente convergente.

No trabalho de Polyak (2008), é introduzido um método primal-dual de pontos exteriores para problemas de otimização convexa, baseado no método de rescalamento não-linear. No método de rescalamento não-linear existe uma alternância no cálculo variáveis primais, através da minimização irrestrita da função Lagrangiana, com a atualização dos multiplicadores de Lagrange e parâmetro de rescalamento. No referido artigo, o autor substitui esta forma de atualização por um sistema primal-dual de equações, que é resolvido através do método de Newton, originando o método primal-dual de pontos exteriores.

Em 2014, Polyak (2014) mostrou que existem conexões entre os métodos primais exteriores e os métodos duais de pontos interiores. Para tanto, utilizou-se de uma transformação Lagrangiana, e mostrou que o método primal exterior com transformação Lagrangiana é equivalente ao método dual de elipsoide interior. Mostrou-se, ainda, que a aplicação do método da transformação Lagrangiana associado à transformação de barreira modificada para programação linear, conduz ao método afim escala de Dikin.

Atualmente, após passarem por várias modificações e desenvolvimentos, os métodos de pontos interiores e exteriores têm sido largamente utilizados na solução de problemas de otimização linear, não-linear e inteira. Na engenharia elétrica, os métodos de pontos interiores e exteriores podem auxiliar na tomada de decisão, como no trabalho de Pinheiro et al. (2014), em que a função barreira modificada, associada a estratégias de convergência global e ajuste cúbico, foi aplicada à resolução do problema de FPO reativo.

4.2 Métodos de otimização

Os principais aspectos teóricos da teoria de otimização podem ser encontrados no Apêndice B. Agora, são apresentados alguns métodos que podem ser utilizados em problemas de otimização restrita. De particular interesse para este trabalho são os métodos de penalidade e barreira, que constituem a base do método que será apresentado no próximo capítulo. Os conceitos teóricos abordados no Apêndice B, especialmente as condições de otimalidade e a teoria de dualidade, permitem delinear diversos métodos que podem ser úteis à resolução de problemas restritos. Nesta seção, mostra-se como estes métodos podem ser utilizados de modo a converter um problema restrito em um problema irrestrito equivalente.

4.2.1 Método Dual-Lagrangiano

De acordo com Pinheiro (2012), o método dual-Lagrangiano é pouco utilizado na prática, mas de grande importância devido aos conceitos teóricos que o mesmo incorpora e que estão relacionados à teoria de dualidade desenvolvida no Apêndice B. Para tanto, considera-se o seguinte problema de programação não-linear:

$$\begin{aligned} &\text{minimizar} && f(\mathbf{x}) \\ &\text{sujeito a} && \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ &&& \mathbf{h}(\mathbf{x}) \leq \mathbf{0} \end{aligned} \tag{4.1}$$

em que $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ são funções de classe C^2 , $\mathbf{x} \in \Omega \subseteq \mathbb{R}^n$, com Ω um conjunto aberto.

Associado a ele, temos a função Lagrangiana:

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \nu_j h_j(\mathbf{x}). \tag{4.2}$$

em que λ_i , $i = 1, \dots, m$ são os multiplicadores de Lagrange associados às restrições de igualdade e $\nu_j \geq 0$, $j = 1, \dots, p$ são os multiplicadores de Lagrange associados às restrições de desigualdade.

O método dual-Lagrangiano segue o forte indicativo apresentado pelo teorema do ponto de sela: um ponto $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ que seja ótimo, com *gap* de dualidade nulo, deve satisfazer a condição de ponto de sela. Isto é, quando os multiplicadores de Lagrange estão fixos em $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, $\mathbf{x}^*$ deve ser um ponto de mínimo para a função $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$. Por outro lado, quando as variáveis primais são fixadas em $\mathbf{x}^*$, então $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ deve ser um ponto de

máximo para $\mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}, \boldsymbol{\nu})$. Com base nesse teorema, podemos esboçar um algoritmo para o método dual-Lagrangiano (Algoritmo 1).

Algoritmo 1: Método Dual-Lagrangiano

- 1 A partir do problema (4.1), determine a função Lagrangiana associada (4.2);
 - 2 Faça $k = 0$ e entre com os dados iniciais: ponto inicial $\mathbf{x}^0 \in \Omega$ e multiplicadores de Lagrange $\boldsymbol{\lambda}^0$ e $\boldsymbol{\nu}^0 \geq \mathbf{0}$;
 - 3 **enquanto** as condições de KKT não forem satisfeitas em $\mathbf{x}^k$, $\boldsymbol{\lambda}^k$ e $\boldsymbol{\nu}^k$ **faça**
 - 4 Resolva o problema Lagrangiano abaixo, com $\boldsymbol{\lambda}^k$ e $\boldsymbol{\nu}^k$ fixos:

$$\min_{\mathbf{x} \in \Omega} f(\mathbf{x}) + \sum_{i=1}^m \lambda_i^k g_i(\mathbf{x}) + \sum_{j=1}^p \nu_j^k h_j(\mathbf{x})$$

e obtenha uma solução $\mathbf{x}^{k+1}$;
 - 5 A partir da solução $\mathbf{x}^{k+1}$ determinada, atualize os multiplicadores de Lagrange, utilizando alguma heurística: $\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \Delta \boldsymbol{\lambda}^k$ e $\boldsymbol{\nu}^{k+1} = \boldsymbol{\nu}^k + \Delta \boldsymbol{\nu}^k$;
 - 6 Faça $k = k + 1$.
-

A atualização dos multiplicadores de Lagrange pode ser feita de diversas maneiras. De acordo com Campanha (2011), como se deseja maximizar nas variáveis duais, pode-se adotar a direção do gradiente para realizar a atualização, de acordo com a seguinte expressão:

$$\begin{pmatrix} \boldsymbol{\lambda}^{k+1} \\ \boldsymbol{\nu}^{k+1} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\lambda}^k \\ \boldsymbol{\nu}^k \end{pmatrix} + \alpha \begin{pmatrix} \mathbf{g}(\mathbf{x}^{k+1}) \\ \mathbf{s} \end{pmatrix} \quad (4.3)$$

na qual α é um passo a ser determinado por uma busca unidimensional e o vetor $\mathbf{s}$ é definido por:

$$s_j = \begin{cases} h_j(\mathbf{x}^{k+1}), & \text{se } \nu_j^k > 0 \\ \max\{0, h_j(\mathbf{x}^{k+1})\}, & \text{se } \nu_j^k = 0 \end{cases} \quad (4.4)$$

para $j = 1, \dots, p$.

Na prática, a utilização do método dual-Lagrangiano é limitada, devido à possibilidade de ocorrência do *gap* de dualidade, levando à divergência ou à soluções diferentes da ótima. Além disso, o emprego dos multiplicadores de Lagrange implica em um aumento no número de variáveis do problema. Por outro lado, este método é comumente empregado com outros métodos, inclusive os que serão apresentados nas seções posteriores. Cabe lembrar que no problema de FPO, sob determinadas hipóteses sobre a resistência dos transformadores, pode-se assumir que o problema de otimização apresenta *gap* de dualidade nulo, quando a função objetivo é de perdas ou de custo quadrático de geração. Esta condição é válida, inclusive, em sistemas comumente testados para verificar a eficiência de métodos computacionais, como os disponíveis no banco de dados da IEEE (Lavai & Low, 2012).

4.2.2 Método de Penalidade

Seja o problema de programação não-linear (4.1). A ideia fundamental do método de penalidade para a resolução do problema restrito (4.1) é a utilização de uma sequência de problemas irrestritos baseados em um parâmetro de penalidade, cujas soluções, no limite, convergem para a solução do problema restrito. Os problemas irrestritos a serem resolvidos utilizam uma função objetivo auxiliar que envolve a função objetivo original ou a função Lagrangiana, associada a uma função penalidade, que impõe um custo quando restrições são violadas, isto é, penaliza os pontos infactíveis.

No método de penalidade, o processo iterativo pode se dar através de pontos que são infactíveis. Um *parâmetro de penalidade* crescente, $\sigma \in \mathbb{R}_+^*$, é imposto à função penalidade e determina o grau de severidade da penalização realizada. A função auxiliar do método de penalidade é da forma:

$$f_\sigma(\mathbf{x}) = f(\mathbf{x}) + \sigma \mathcal{P}(\mathbf{x}). \quad (4.5)$$

de modo que o problema irrestrito a ser resolvido, denominado problema de penalidade, é expresso por:

$$\begin{aligned} &\text{minimizar} && f_\sigma(\mathbf{x}) \\ &\text{sujeito a} && \mathbf{x} \in \Omega \end{aligned} \quad (4.6)$$

Em geral, quanto maior o valor de σ , maior é a ênfase sobre a factibilidade. Isto é, ao resolver o problema irrestrito com a função auxiliar $f_\sigma(\mathbf{x})$, conforme $\sigma \rightarrow \infty$ e $\mathcal{P}(\mathbf{x}) \rightarrow 0$, espera-se que a solução do problema de penalidade convirja para a solução ótima $\mathbf{x}^*$ do problema original. A *função penalidade*, $\mathcal{P}$, associada ao problema (4.1) é dada por:

$$\mathcal{P}(\mathbf{x}) = \sum_{i=1}^m \zeta(g_i(\mathbf{x})) + \sum_{j=1}^p \varphi(h_j(\mathbf{x})), \quad (4.7)$$

em que $\zeta : \mathbb{R} \rightarrow \mathbb{R}$ e $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ são funções contínuas, satisfazendo as seguintes propriedades:

$$\begin{cases} \zeta(y) = 0, & \text{se } y = 0 \\ \zeta(y) > 0, & \text{se } y \neq 0 \end{cases} \quad \text{e} \quad \begin{cases} \varphi(y) = 0, & \text{se } y \leq 0 \\ \varphi(y) > 0, & \text{se } y > 0 \end{cases} \quad (4.8)$$

A função ζ penaliza a violação das restrições de igualdade, enquanto φ penaliza a violação das restrições de desigualdade. As funções penalidade tradicionalmente utilizadas são:

$$\zeta(y) = |y|^\beta \quad \text{e} \quad \varphi(y) = (\max\{0, y\})^\beta, \quad (4.9)$$

em que β é um número inteiro positivo. Em particular, se $\beta = 2$, dizemos que $\mathcal{P}$ é

uma função penalidade quadrática. Uma vez que se deseja que $f_\sigma(\mathbf{x}) \rightarrow f(\mathbf{x})$ quando $\sigma \rightarrow \infty$, então o produto $\sigma\mathcal{P}(\mathbf{x}) \rightarrow 0$, ou seja, o parâmetro σ deve tender ao infinito de modo mais lento que $\mathcal{P}(\mathbf{x})$ se aproxima de 0. Por conseguinte, o valor $\sigma\mathcal{P}(\mathbf{x})$ também pode ser utilizado como critério de parada para o método, quando seu valor absoluto for suficientemente pequeno. Um algoritmo para o método de penalidade é apresentado no Algoritmo 2.

Algoritmo 2: Método de Penalidade

- 1 A partir do problema (4.1), determine a função penalidade associada;
 - 2 Faça $k = 0$ e entre com os dados iniciais: ponto inicial $\mathbf{x}^0$, parâmetro de penalidade $\sigma^0 > 0$, parâmetro de atualização da penalidade $\kappa > 1$ e precisão ε ;
 - 3 **enquanto** $|\sigma^k\mathcal{P}(\mathbf{x}^k)| > \varepsilon$ e $\|\nabla f_{\sigma^k}(\mathbf{x}^k)\| > \varepsilon$ **faça**
 - 4 Resolva o problema de penalidade (4.6) com σ^k fixo:

$$\min_{\mathbf{x} \in \Omega} f(\mathbf{x}) + \sigma^k\mathcal{P}(\mathbf{x})$$
 - e obtenha uma solução $\mathbf{x}^{k+1}$;
 - 5 Atualize o parâmetro de penalidade por $\sigma^{k+1} = \kappa\sigma^k$;
 - 6 Faça $k = k + 1$.
-

De acordo com Griva et al. (2009), os métodos de penalidade são mais indicados para tratamento das restrições de igualdade. Apesar de também serem capazes de tratar desigualdades, os problemas irrestritos obtidos normalmente não possuem derivadas de segunda ordem contínuas, de modo que métodos de barreira são preferíveis. Quando os métodos de penalidade são associados à função Lagrangiana, dão origem a métodos conhecidos como Lagrangiana aumentada.

Do ponto de vista prático, a utilização de métodos de penalidade pode encontrar dificuldade quando os valores do parâmetro de penalidade σ é elevado, devido ao mau-condicionamento da matriz Hessiana. Além disso, a escolha do parâmetro de penalidade inicial e do fator de atualização do mesmo deve ser feita de maneira criteriosa, a fim de evitar a parada prematura do método.

4.2.3 Método de Barreira

Os métodos de barreira são desenvolvidos para tratar de problemas com restrições de desigualdade, cujo interior da região viável é não vazio, isto é, problemas no formato:

$$\begin{aligned} &\text{minimizar} && f(\mathbf{x}) \\ &\text{sujeito a} && \mathbf{h}(\mathbf{x}) \geq \mathbf{0} \end{aligned} \tag{4.10}$$

no qual assumiremos que $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ são funções de classe C^2 , $\mathbf{x} \in \Omega \subseteq \mathbb{R}^n$, com Ω um conjunto aberto.

Os métodos de barreira se assemelham ao método de penalidade, uma vez que convertem o problema (4.10) em um problema irrestrito, o qual deve ser resolvido para uma sequência de parâmetros de barreira que se aproxima de 0. Neste caso, cada uma das restrições de desigualdade é tratada através de uma função barreira, que é incorporada à função objetivo original, dando origem a uma função auxiliar.

A diferença básica entre os métodos de penalidade e barreira é que este último é um método estritamente factível: os pontos estão sempre no interior da região factível, ao longo do processo iterativo. Daí a necessidade de assumir que, neste caso, o interior da região factível determinada pelas restrições é não-vazio, isto é,

$$\overset{\circ}{\mathcal{X}} = \{\mathbf{x} \in \Omega : \mathbf{h}(\mathbf{x}) > \mathbf{0}\} \neq \emptyset.$$

Além disso, assume-se que todo ponto factível que não esteja em $\overset{\circ}{\mathcal{X}}$, isto é, situado na fronteira, pode ser arbitrariamente aproximado por um elemento de $\overset{\circ}{\mathcal{X}}$, ou seja, é um ponto de acumulação do conjunto $\overset{\circ}{\mathcal{X}}$.

Os métodos de barreira preservam a factibilidade criando uma barreira que mantém os pontos iterandos no interior da região factível, sendo por isso também conhecidos como métodos de pontos interiores. Neste método, a função auxiliar assume a forma:

$$f_\mu(\mathbf{x}) = f(\mathbf{x}) + \mu \mathcal{B}(\mathbf{x}). \quad (4.11)$$

em que $\mu \in \mathbb{R}_+^*$ é denominado *parâmetro de barreira* e $\mathcal{B}$ é a função barreira. Desta forma, o problema irrestrito a ser resolvido, conhecido como problema barreira, é dado por:

$$\begin{aligned} &\text{minimizar} && f_\mu(\mathbf{x}) \\ &\text{sujeito a} && \mathbf{x} \in \overset{\circ}{\mathcal{X}} \end{aligned} \quad (4.12)$$

A função barreira $\mathcal{B}$ associada ao problema (4.10) é da forma:

$$\mathcal{B}(\mathbf{x}) = \sum_{j=1}^p \phi(h_j(\mathbf{x})), \quad (4.13)$$

em que $\phi : \mathbb{R}_+^* \rightarrow \mathbb{R}$ é uma função contínua com as seguintes propriedades:

$$\begin{cases} \phi(y) \geq 0, & \text{se } y > 0 \\ \lim_{y \rightarrow 0^+} \phi(y) = \infty \end{cases} \quad (4.14)$$

As funções barreira mais comumente utilizadas são aquelas em que:

$$\phi(y) = \frac{1}{y} \quad (4.15)$$

$$\phi(y) = -\ln(\min\{1, y\}) \quad (4.16)$$

Observa-se que a função em (4.16) é não-diferenciável. De acordo com Lage (2013) as propriedades em (4.14) apenas precisam ser satisfeitas na vizinhança de $y = 0$, de modo que podemos tomar:

$$\phi(y) = -\ln(y). \quad (4.17)$$

A função barreira em (4.15) é denominada *barreira clássica ou inversa* e foi apresentada por Carroll & Fiacco (1961), enquanto que a apresentada em (4.17) é denominada *barreira logarítmica* e foi introduzida por Frisch (1955).

Naturalmente, à medida que a resolução dos problemas irrestritos é realizada, deseja-se que $\mu\mathcal{B}(\mathbf{x}) \rightarrow 0$, conforme $\mathbf{x} \rightarrow \mathbf{x}^*$. Se a solução ótima $\mathbf{x}^*$ do problema original está na fronteira da região factível, então quando $\mathbf{x} \rightarrow \mathbf{x}^*$ tem-se que $\mathcal{B}(\mathbf{x}) \rightarrow \infty$, de modo que é necessário que o parâmetro de barreira $\mu \rightarrow 0$ mais rapidamente que a função barreira $\mathcal{B}(\mathbf{x}) \rightarrow \infty$, a fim de que ocorra $\mu\mathcal{B}(\mathbf{x}) \rightarrow 0$. Esta redução no parâmetro de barreira visa reduzir o efeito do termo de barreira, a fim de que os pontos iterandos possam, gradualmente, se aproximar da fronteira da região factível.

Portanto, da mesma maneira que no método de penalidade, o produto $\mu\mathcal{B}(\mathbf{x})$ pode ser utilizado no critério de parada do método. Caso o valor absoluto de $\mu\mathcal{B}(\mathbf{x})$ não seja menor que a precisão desejada, o parâmetro de barreira deve ser reduzido utilizando alguma heurística, e o problema barreira deve ser resolvido novamente. Um algoritmo para o método de barreira é delineado a seguir (Algoritmo 3).

Algoritmo 3: Método de Barreira

- 1 A partir do problema (4.10), determine a função barreira associada;
 - 2 Faça $k = 0$ e entre com os dados iniciais: ponto inicial $\mathbf{x}^0 \in \overset{\circ}{\mathcal{X}}$, parâmetro de barreira $\mu^0 > 0$, parâmetro de atualização da barreira $0 < \tau < 1$ e precisão $\varepsilon > 0$;
 - 3 **enquanto** $|\mu^k \mathcal{B}(\mathbf{x}^k)| > \varepsilon$ **faça**
 - 4 Resolva o problema barreira (4.12) com μ^k fixo:

$$\min_{\mathbf{x} \in \overset{\circ}{\mathcal{X}}} f(\mathbf{x}) + \mu^k \mathcal{B}(\mathbf{x})$$
 e obtenha uma solução $\mathbf{x}^{k+1}$;
 - 5 Atualize o parâmetro de barreira por $\mu^{k+1} = \tau \mu^k$;
 - 6 Faça $k = k + 1$.
-

Do ponto de vista prático, uma das dificuldades do método de barreira é a escolha de

um ponto inicial que pertença ao interior da região factível ($\overset{\circ}{\mathcal{X}}$). Além disso, conforme o parâmetro de barreira é reduzido, os problemas podem se tornar mais difíceis de serem resolvidos, principalmente devido ao mau-condicionamento da matriz hessiana. Assim como no método de penalidade, a escolha do parâmetro de barreira inicial e seu fator de atualização deve ser feita de modo cuidadoso para evitar instabilidades numéricas.

4.2.4 Método de Rescalamento Não-Linear e de Barreira Modificada

Em 1992, Polyak (1992) desenvolveu e apresentou uma teoria de métodos da função barreira modificada, como forma de tratar as restrições de desigualdade em problemas de otimização. A função barreira modificada pode ser vista, de acordo com Pinheiro (2012), como uma função Lagrangiana aumentada interior. Da mesma forma que nos outros métodos apresentados, a ideia do método da função barreira modificada é a resolução de uma sequência de problemas irrestritos, até a convergência com determinada precisão para a solução ótima do problema original.

A fim de tratar deste método, considere um problema no formato:

$$\begin{aligned} &\text{minimizar} && f(\mathbf{x}) \\ &\text{sujeito a} && \mathbf{h}(\mathbf{x}) \geq \mathbf{0} \end{aligned} \tag{4.18}$$

no qual assumiremos que $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ são funções de classe C^2 , $\mathbf{x} \in \Omega \subseteq \mathbb{R}^n$, com Ω um conjunto aberto.

Posteriormente, Polyak estendeu a teoria da função barreira modificada a uma teoria mais geral, denominada *rescalamento não-linear*, que incorpora a função barreira modificada, bem como outras funções que satisfazem determinadas propriedades. No método de rescalamento não-linear, é feita uma transformação na função objetivo e/ou restrições obtendo um problema equivalente com a mesma região factível. A transformação não-linear utilizada é uma função suave e depende de um parâmetro positivo, e usualmente é denominada *função de rescalamento*. Os multiplicadores de Lagrange das restrições modificadas são, normalmente, chamados de *estimadores*.

De acordo com Polyak & Teboulle (1997), uma função de rescalamento é uma função ψ de classe C^2 , definida sobre um intervalo $(a, +\infty)$, com $-\infty \leq a < 0$, $\psi(a) = -\infty$ e $\psi'(a) = +\infty$ satisfazendo as seguintes propriedades:

- (i) $\psi(0) = 0$ e $\psi'(0) = 1$,
- (ii) ψ é estritamente crescente, isto é, $\psi'(t) > 0$,

(iii) ψ é côncava, isto é, $\psi''(t) < 0$,

(iv) $\lim_{t \rightarrow +\infty} \psi'(t) = 0$.

Denota-se a classe das funções de rescalamento por Ψ . Dada $\psi \in \Psi$, o problema (4.18) é convertido no seguinte problema equivalente devido às propriedades da função ψ , utilizando um parâmetro $\mu > 0$:

$$\begin{aligned} &\text{minimizar} && f(\mathbf{x}) \\ &\text{sujeito a} && -\mu\psi(\mu^{-1}h_j(\mathbf{x})) \leq 0, \quad j = 1, \dots, p \end{aligned} \tag{4.19}$$

A transformação acima não altera o conjunto de soluções ótimas (Polyak & Teboulle, 1997). Por outro lado, o problema transformado (4.19) possui a vantagem de que, sendo a função ψ definida em $(a, +\infty)$, os valores de $\mathbf{x}$ podem ser tais que $\mu^{-1}h_j(\mathbf{x}) > a$ se, e somente se, $h_j(\mathbf{x}) > a\mu$. Como $a < 0$ implica que $a\mu < 0$, isto corresponde a uma relaxação da região factível permitida para o processo iterativo, já que pode-se ter $a\mu < h_j(\mathbf{x}) < 0$ ou $h_j(\mathbf{x}) \geq 0$, para todo $j = 1, \dots, p$.

A região factível relaxada, $\hat{\mathcal{X}}$, portanto, pode ser vista como a união da região factível original $\mathcal{X} = \{\mathbf{x} \in \Omega : \mathbf{h}(\mathbf{x}) \geq \mathbf{0}\}$, com uma região adicional que depende do parâmetro μ , dada por $\mathcal{X}_\mu = \{\mathbf{x} \in \Omega : a\mu < \mathbf{h}(\mathbf{x}) < \mathbf{0}\}$. À medida em que o parâmetro de barreira/relaxação $\mu \rightarrow 0$, tem-se que $a\mu \rightarrow 0$, de modo que a região relaxada, no limite, torna-se a região factível original.

A função Lagrangiana associada ao problema original, $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}_+^p \rightarrow \mathbb{R}$, é expressa por:

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\nu}) = f(\mathbf{x}) - \sum_{j=1}^p \nu_j h_j(\mathbf{x}), \tag{4.20}$$

em que $\nu_j \geq 0$, $j = 1, \dots, p$ são os multiplicadores de Lagrange associados às restrições de desigualdade. Por outro lado, a função Lagrangiana $\mathcal{L}_\mu : \mathbb{R}^n \times \mathbb{R}_+^p \rightarrow \mathbb{R}$, associada ao problema transformado, é dada por:

$$\mathcal{L}_\mu(\mathbf{x}, \boldsymbol{\delta}) = f(\mathbf{x}) - \mu \sum_{j=1}^p \delta_j \psi(\mu^{-1}h_j(\mathbf{x})). \tag{4.21}$$

em que δ_j , $j = 1, \dots, p$ são os multiplicadores de Lagrange associados às restrições de desigualdade transformadas ou estimadores dos multiplicadores de Lagrange. A relaxação da região factível garante que, diferentemente dos métodos de barreira, nos métodos baseados em rescalamento não-linear a função Lagrangiana $\mathcal{L}_\mu(\mathbf{x}, \boldsymbol{\delta})$ é definida na fronteira da região factível do problema original, para qualquer valor $\mu > 0$.

Para resolver o problema pelo método do rescalamento não-linear, as ideias empregadas

nos métodos dual-Lagrangiano e de barreira são utilizadas: os valores dos estimadores dos multiplicadores de Lagrange em uma iteração k , $\boldsymbol{\delta}^k$, e do parâmetro de barreira, μ^k , são fixados, e a função Lagrangiana é minimizada em $\mathbf{x}$. Caso a convergência não tenha sido atingida, os estimadores dos multiplicadores de Lagrange e o parâmetro barreira são atualizados, e o procedimento de minimização da Lagrangiana é repetido. Assim, o problema a ser resolvido é da forma:

$$\begin{aligned} \text{minimizar} \quad & \mathcal{L}_\mu(\mathbf{x}, \boldsymbol{\delta}) = f(\mathbf{x}) - \mu \sum_{j=1}^p \delta_j \psi(\mu^{-1} h_j(\mathbf{x})) \\ \text{sujeito a} \quad & \mathbf{x} \in \widehat{\mathcal{X}} \end{aligned} \quad (4.22)$$

para δ_j , $j = 1, \dots, p$ e μ fixos.

A seguir, investiga-se a questão de como atualizar os estimadores dos multiplicadores de Lagrange. Aplicando as condições de otimalidade sobre a função Lagrangiana do problema transformado em (4.21), e assumindo que uma nova solução $\mathbf{x}^{k+1}$ foi determinada, tem-se que:

$$\begin{aligned} \nabla_{\mathbf{x}} \mathcal{L}_\mu(\mathbf{x}^{k+1}, \boldsymbol{\delta}^k) &= \nabla f(\mathbf{x}^{k+1}) - \mu^k \sum_{j=1}^p \delta_j^k \psi'((\mu^k)^{-1} h_j(\mathbf{x}^{k+1})) (\mu^k)^{-1} \nabla h_j(\mathbf{x}^{k+1}) = \mathbf{0} \\ \text{se, e somente se, } \nabla f(\mathbf{x}^{k+1}) &- \sum_{j=1}^p \delta_j^k \psi'((\mu^k)^{-1} h_j(\mathbf{x}^{k+1})) \nabla h_j(\mathbf{x}^{k+1}) = \mathbf{0}. \end{aligned}$$

Como se deseja que a condição de estacionariedade também seja verificada para o problema original, Polyak & Teboulle (1997) sugerem que o estimador do multiplicador de Lagrange seja atualizado por:

$$\delta_j^{k+1} = \delta_j^k \psi'((\mu^k)^{-1} h_j(\mathbf{x}^{k+1})), \quad j = 1, \dots, p. \quad (4.23)$$

Dessa forma, considerando a Lagrangiana do problema original (4.20), tem-se que:

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^{k+1}, \boldsymbol{\delta}^{k+1}) = \nabla f(\mathbf{x}^{k+1}) - \sum_{j=1}^p \delta_j^{k+1} \nabla h_j(\mathbf{x}^{k+1}) = \mathbf{0}. \quad (4.24)$$

Um ponto importante a se notar é que, sendo $\boldsymbol{\delta}^0 > \mathbf{0}$, a propriedade (ii) da função de rescalamento garante que os estimadores das iterações seguintes também serão tais que $\boldsymbol{\delta}^k > \mathbf{0}$. O procedimento abordado nesta seção é esboçado no Algoritmo 4.

Em seu trabalho, Polyak & Teboulle (1997) apresentam algumas funções de rescalamento, a saber:

1. Rescalamento exponencial: $\psi(t) = 1 - e^{-t}$, $t \in \mathbb{R}$;

Algoritmo 4: Método de Rescalamento Não-Linear

- 1 A partir do problema (4.18) e de uma função de rescalamento $\psi \in \Psi$, determine a função Lagrangiana (4.21) do problema transformado;
 - 2 Faça $k = 0$ e entre com os dados iniciais: ponto inicial $\mathbf{x}^0 \in \hat{\mathcal{X}}$, estimador do multiplicador de Lagrange $\boldsymbol{\delta}^0 > \mathbf{0}$, parâmetro de barreira $\mu^0 > 0$, parâmetro de atualização da barreira $0 < \tau < 1$ e precisão $\varepsilon > 0$;
 - 3 **enquanto** *as condições de KKT não forem satisfeitas para $\mathbf{x}^k$ e $\boldsymbol{\delta}^k$, ou $h_j(\mathbf{x}^k) < 0$ para algum $j = 1, \dots, p$* **faça**
 - 4 Resolva o problema de rescalamento não-linear (4.22) com μ^k e $\boldsymbol{\delta}^k$ fixos:
$$\min_{\mathbf{x} \in \hat{\mathcal{X}}} f(\mathbf{x}) - \mu^k \sum_{j=1}^p \delta_j^k \psi((\mu^k)^{-1} h_j(\mathbf{x}))$$

e obtenha uma solução $\mathbf{x}^{k+1}$;
 - 5 Atualize os estimadores dos multiplicadores de Lagrange através de (4.23), determinando $\boldsymbol{\delta}^{k+1}$;
 - 6 Atualize o parâmetro de barreira por $\mu^{k+1} = \tau \mu^k$;
 - 7 Faça $k = k + 1$.
-

2. Função barreira logarítmica modificada: $\psi(t) = \ln(t + 1)$, $t \in (-1, \infty)$;

3. Função barreira inversa modificada: $\psi(t) = \frac{t}{t+1}$, $t \in (-1, \infty)$.

Neste trabalho, o interesse está voltado para a função barreira logarítmica modificada, a qual já foi utilizada com sucesso em problemas de despacho econômico e de FPO (Sousa et al., 2012; Pinheiro, 2012; Lage, 2013).

4.2.4.1 Função Barreira Logarítmica Modificada

O método da função barreira logarítmica modificada se origina quando a função de rescalamento utilizada é expressa por $\psi(t) = \ln(t + 1)$, com $t \in (-1, \infty)$. Neste caso, a função Lagrangiana do problema transformado (4.21) passa a ser chamada de *função barreira modificada*, e assume a forma:

$$\mathcal{M}_\mu(\mathbf{x}, \boldsymbol{\delta}) = f(\mathbf{x}) - \mu \sum_{j=1}^p \delta_j \ln(\mu^{-1} h_j(\mathbf{x}) + 1). \quad (4.25)$$

em que δ_j , $j = 1, \dots, p$ são os multiplicadores de Lagrange associados às restrições de desigualdade transformadas ou estimadores dos multiplicadores de Lagrange.

De acordo com Nash et al. (1994), o uso dos termos de barreira $\ln(\mu^{-1} h_j(\mathbf{x}) + 1)$ corresponde a uma perturbação nas restrições, de modo que $h_j(\mathbf{x}) > -\mu$, para $j = 1, \dots, p$. Esta perturbação nada mais é do que a relaxação citada anteriormente. Neste caso, $a = -1$,

pois o domínio de ψ é $(-1, \infty)$.

O problema de barreira modificada a ser resolvido é:

$$\begin{aligned} & \text{minimizar} \quad \mathcal{M}_\mu(\mathbf{x}, \boldsymbol{\delta}) = f(\mathbf{x}) - \mu \sum_{j=1}^p \delta_j \ln(\mu^{-1} h_j(\mathbf{x}) + 1) \\ & \text{sujeito a} \quad \mathbf{x} \in \widehat{\mathcal{X}} \end{aligned} \tag{4.26}$$

Uma vez que advém de uma função de rescalamento não-linear, a função barreira modificada existe em uma solução $\mathbf{x}^*$ para qualquer valor positivo do parâmetro de barreira μ , o que não acontece no caso da função barreira logarítmica clássica. Além disso, se os valores ótimos $\mathbf{x}^*$ e $\boldsymbol{\delta}^*$ forem conhecidos, a função $\mathcal{M}_\mu$ goza das seguintes propriedades:

- (P1) $\mathcal{M}_\mu(\mathbf{x}^*, \boldsymbol{\delta}^*) = f(\mathbf{x}^*)$, para todo $\mu > 0$;
- (P2) $\nabla_{\mathbf{x}} \mathcal{M}_\mu(\mathbf{x}^*, \boldsymbol{\delta}^*) = \nabla f(\mathbf{x}^*) - \sum_{j=1}^p \delta_j^* \nabla h_j(\mathbf{x}^*) = \mathbf{0}$, para todo $\mu > 0$.

Segundo Nash et al. (1994), Polyak mostrou que se os estimadores dos multiplicadores de Lagrange iniciais são positivos e os parâmetros de barreira estão abaixo de um limitante $\bar{\mu}$, então o método converge. Entretanto, enfatiza que em problemas práticos não é possível saber *a priori* se a escolha inicial do parâmetro está ou não abaixo do limitante $\bar{\mu}$. Além disso, se o problema for convexo, vale a propriedade:

- (P3) $\mathbf{x}^* = \arg \min \{\mathcal{M}_\mu(\mathbf{x}, \boldsymbol{\delta}^*)\}$, para todo $\mu > 0$.

Isto significa que, se os multiplicadores de Lagrange ótimos $\boldsymbol{\delta}^*$ forem conhecidos de antemão, o problema de otimização restrito pode ser resolvido utilizando um único problema de otimização irrestrito, que independe do valor do parâmetro de barreira. Portanto, quando uma boa estimativa dos multiplicadores de Lagrange é conhecida, o parâmetro de barreira μ não precisa tender a zero com a mesma intensidade dos métodos de barreira clássicos, contribuindo para a melhoria do condicionamento da matriz hessiana e estabilidade numérica do método.

No caso da função barreira modificada, aplicando a expressão (4.23), conclui-se que os estimadores dos multiplicadores de Lagrange são atualizados por:

$$\delta_j^{k+1} = \frac{\delta_j^k}{(\mu^k)^{-1} h_j(\mathbf{x}^{k+1}) + 1}. \tag{4.27}$$

4.2.4.2 Sobre a atualização dos estimadores dos multiplicadores de Lagrange

A expressão (4.27) para atualização dos estimadores dos multiplicadores de Lagrange, a qual foi apresentada no trabalho de Polyak (1992), se mostrou, em diversos trabalhos, bastante eficiente (Sousa, 2006; Stanzani, 2012; Lage, 2013). Entretanto, Pinheiro (2012) sugere uma maneira mais barata computacionalmente e que foi empregada com sucesso na resolução de problemas de FPO, e que também já havia sido proposta previamente por Griva & Polyak (2006).

Para tanto, consideremos o problema da forma:

$$\begin{aligned} & \text{minimizar} && f(\mathbf{x}) \\ & \text{sujeito a} && \mathbf{h}(\mathbf{x}) \leq \mathbf{0} \end{aligned} \tag{4.28}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ são funções de classe C^2 , $\mathbf{x} \in \Omega \subseteq \mathbb{R}^n$, com Ω um conjunto aberto. Este problema, então, é convertido no seguinte problema equivalente, acrescentando variáveis de folga não-negativas às restrições de desigualdade:

$$\begin{aligned} & \text{minimizar} && f(\mathbf{x}) \\ & \text{sujeito a} && \mathbf{h}(\mathbf{x}) + \mathbf{s} = \mathbf{0} \\ & && \mathbf{s} \geq \mathbf{0}, \mathbf{s} \in \mathbb{R}^p \end{aligned} \tag{4.29}$$

Tratando as restrições de não-negatividade das folgas através da função barreira modificada e a igualdade pelo método Lagrangiano, obtém-se a seguinte *função Lagrangiana barreira logarítmica modificada* associada ao problema (4.29), e que foi utilizada no trabalho de Sousa (2006):

$$\mathcal{L}_\mu(\mathbf{x}, \mathbf{s}, \boldsymbol{\nu}) = f(\mathbf{x}) + \sum_{j=1}^p \nu_j (h_j(\mathbf{x}) + s_j) - \mu \sum_{j=1}^p \delta_j \ln(\mu^{-1} s_j + 1). \tag{4.30}$$

em que $\nu_j \geq 0$, $j = 1, \dots, p$ são os multiplicadores de Lagrange associados às restrições de desigualdade convertidas em igualdade, e $\delta_j \geq 0$, $j = 1, \dots, p$ são os estimadores dos multiplicadores de Lagrange, relacionados às restrições de não-negatividade das variáveis de folga.

Suponha que, em uma iteração, tenha sido determinado um novo ponto $(\mathbf{x}^{k+1}, \mathbf{s}^{k+1}, \boldsymbol{\nu}^{k+1})$ e que, a partir do mesmo, um novo estimador deve ser calculado. Como este ponto é calculado de modo a satisfazer as condições de otimalidade, deve-se ter:

$$\frac{\partial \mathcal{L}_{\mu^k}}{\partial s_j}(\mathbf{x}^{k+1}, \mathbf{s}^{k+1}, \boldsymbol{\nu}^{k+1}) = 0, \text{ ou seja, } \nu_j^{k+1} - \frac{\delta_j^k}{(\mu^k)^{-1} s_j^{k+1} + 1} = 0. \tag{4.31}$$

Disto, resulta que

$$\nu_j^{k+1} = \frac{\delta_j^k}{(\mu^k)^{-1} s_j^{k+1} + 1}. \quad (4.32)$$

Comparando a igualdade em (4.27) com (4.32), conclui-se que:

$$\delta^{k+1} = \nu^{k+1}. \quad (4.33)$$

Portanto, a atualização do estimador δ^{k+1} pode ser feita atribuindo a ele o valor do multiplicador de Lagrange da restrição de desigualdade que foi convertida em igualdade, ν^{k+1} . Pinheiro (2012) aponta que este resultado apresenta as seguintes vantagens:

- A expressão (4.27) não precisa ser utilizada para a atualização dos estimadores dos multiplicadores de Lagrange;
- Esta estratégia para o cálculo de δ^{k+1} será robusta, desde que o multiplicador de Lagrange ν^{k+1} tenha sido calculado através de um método robusto, como o método de Newton. Além disso, o estimador também possuirá a propriedade de não-negatividade proporcionada pela expressão (4.27), já que o multiplicador de Lagrange ν^{k+1} é calculado através de uma busca unidimensional que preserva a não-negatividade;
- O custo computacional dessa atualização é baixo, já que as componentes de δ^{k+1} recebem o respectivo valor das componentes de ν^{k+1} .

Seguindo a mesma proposta, este trabalho utiliza essa forma de atualização dos estimadores dos multiplicadores de Lagrange.

Capítulo 5

Metodologia de Resolução

Um dos objetivos deste capítulo é apresentar o método primal-dual previsor-corretor de pontos interiores, visando aplicá-lo à resolução de problemas de despacho econômico com ponto de carregamento de válvula. Tal método utiliza o rescalamento não-linear baseado na função barreira logarítmica modificada, bem como as ideias apresentadas no trabalho pioneiro de Polyak (1992), o qual foi um marco no desenvolvimento dos métodos de pontos interiores e exteriores. O procedimento previsor-corretor se baseia no trabalho de Mehrotra (1992) e Wu et al. (1994), sendo que este último o aplicou na resolução de problemas de FPO.

Os métodos Lagrangianos, em geral, são bastante eficazes na resolução de problemas com restrições de igualdade. Entretanto, na presença de restrições de desigualdade, este método tende a apresentar maiores dificuldades, pois requer o conhecimento do conjunto das restrições ativas na solução.

Em contrapartida, os métodos de barreira podem ser aplicados apenas a problemas cujas restrições são de desigualdade e, usualmente, apresentam bons resultados. A maior dificuldade dos métodos de barreira está relacionada ao mau condicionamento da matriz Hessiana da função auxiliar, em decorrência da redução do parâmetro de barreira. Além disso, as funções barreira clássicas (logarítmica e inversa) não estão definidas na fronteira da região factível e o processo de inicialização pode ser difícil se um ponto interior não for conhecido.

Neste sentido, os métodos baseados em rescalamento não-linear, em particular os que se baseiam na função barreira logarítmica modificada, se apresentam como uma alternativa vantajosa para o tratamento das restrições de desigualdade: os problemas típicos dos métodos de barreira clássica são amenizados devido à relaxação na região factível. Esta relaxação permite que pontos ótimos na fronteira do conjunto sejam atingidos, além de

melhorar o condicionamento da matriz hessiana e facilitar a inicialização.

Assim, visando aproveitar as melhores características dos métodos apresentados, Sousa (2006) desenvolveu uma abordagem denominada método da função Lagrangiana barreira modificada, que combina os métodos primal-dual barreira logarítmica e de barreira modificada, no qual a função barreira é aplicada sobre as folgas das restrições de desigualdade. As condições necessárias de primeira ordem foram aplicadas sobre a função Lagrangiana barreira modificada, originando um sistema não-linear. A resolução do sistema resultante, através do método de Newton, permitiu a atualização simultânea das variáveis primais e duais.

O trabalho de Sousa (2006) serviu como base para Pinheiro (2012) desenvolver um método primal-dual previsor corretor de pontos interiores, baseado na função Lagrangiana barreira logarítmica modificada. No referido trabalho, o autor ainda propôs um ajuste cúbico para o logaritmo, a fim de evitar problemas de inviabilidade de pontos em que a função barreira logarítmica modificada não está definida, isto é, fora da região relaxada, além de preservar as derivadas de primeira e segunda ordem no ponto de ajuste. Adicionalmente, apresentou-se uma estratégia de convergência global através de uma variante do método de Levenberg-Marquardt, na qual ajustes eram realizados na matriz dual normal da função Lagrangiana, caso esta não seja definida positiva. Posteriormente, em Pinheiro et al. (2014), esta metodologia foi comparada aos métodos de pontos interiores clássicos. Os resultados obtidos mostraram a eficiência e robustez do método.

O outro objetivo deste capítulo é apresentar a técnica de suavização hiperbólica, já empregada no trabalho de outros autores (Junior, 2010; Xavier, 2010), que foi utilizada para tratamento da função objetivo do PDE-PV e suas variantes. Esta forma de suavização pode ser utilizada para qualquer função objetivo que apresente termos modulares, e é controlada através de um parâmetro de aproximação positivo. À medida em que este parâmetro tende a zero, a função suavizante se aproxima cada vez mais da função objetivo não-diferenciável original.

No trabalho de Junior (2010), a técnica de suavização hiperbólica foi utilizada para a resolução do problema de empacotamento de circunferências sobre a superfície de uma esfera, articulando a técnica de suavização hiperbólica a uma técnica de penalização hiperbólica, sendo capaz de determinar ótimos locais melhores que os até então apresentados na literatura. Já no trabalho de Souza (2010), a técnica de suavização hiperbólica foi empregada para resolver problemas de geometria molecular.

5.1 O Método Primal-Dual Previsor-Corretor Barreira Logarítmica Modificada com Correção de Inércia

Fundamentado nos trabalhos de Polyak (1992), Griva & Polyak (2006), Sousa (2006) e Pinheiro (2012), nesta seção é proposto um método primal-dual previsor-corretor de pontos interiores, que utiliza a função barreira logarítmica modificada. Diferentemente dos trabalhos de Wu et al. (1994) e Torres (1998), em que o parâmetro de barreira é empregado apenas no passo corretor, mediante uma estratégia de escolha adaptativa, neste trabalho o parâmetro de barreira é utilizado tanto nos passos previsor quanto corretor. Além disso, é apresentada uma estratégia de convergência, por meio da correção da inércia da matriz do sistema de direções do método. Na abordagem proposta, as restrições de desigualdade são convertidas em restrições de igualdade através da adição de variáveis de folga. A condição de não-negatividade das folgas é, então, tratada através da função barreira modificada. Os estimadores dos multiplicadores de Lagrange foram atualizados pela regra apresentada em Pinheiro (2012).

5.1.1 O sistema de direções de busca

Nesta seção é apresentado o sistema de direções de busca do método Primal-Dual Previsor-Corretor Barreira Logarítmica Modificada (PDPCBLM). Para tanto, considera-se um problema de otimização com o seguinte formato:

$$\begin{aligned} &\text{minimizar} && f(\mathbf{x}) \\ &\text{sujeito a} && \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ &&& \mathbf{h}(\mathbf{x}) \leq \mathbf{0} \end{aligned} \tag{5.1}$$

em que $\mathbf{x} \in \mathbb{R}^n$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ são funções de classe C^2 .

As restrições de desigualdade são transformadas em restrições de igualdade pelo acréscimo de variáveis de folga $\mathbf{s} \in \mathbb{R}_+^p$. Assim, o problema (5.1) pode se reescrito como:

$$\begin{aligned} &\text{minimizar} && f(\mathbf{x}) \\ &\text{sujeito a} && \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ &&& \mathbf{h}(\mathbf{x}) + \mathbf{s} = \mathbf{0} \\ &&& \mathbf{s} \geq \mathbf{0} \end{aligned} \tag{5.2}$$

As restrições de não-negatividade das folgas são acrescentadas à função objetivo com auxílio da função de rescalamento baseada na barreira logarítmica modificada e de esti-

madores dos multiplicadores de Lagrange. Obtém-se, dessa forma, o seguinte problema transformado:

$$\begin{aligned}
& \text{minimizar} && f(\mathbf{x}) - \mu \sum_{j=1}^p \delta_j \ln(\mu^{-1} s_j + 1) \\
& \text{sujeito a} && \mathbf{g}(\mathbf{x}) = \mathbf{0} \\
& && \mathbf{h}(\mathbf{x}) + \mathbf{s} = \mathbf{0}
\end{aligned} \tag{5.3}$$

em que $\mu > 0$ é o parâmetro de barreira e $\boldsymbol{\delta} \in \mathbb{R}_+^p$ é o vetor de estimadores dos multiplicadores de Lagrange relativos às restrições de não-negatividade das folgas.

As restrições de igualdade são, então, incorporadas à função objetivo utilizando multiplicadores de Lagrange $\boldsymbol{\lambda} \in \mathbb{R}^m$ e $\boldsymbol{\nu} \in \mathbb{R}_+^p$. Dessa forma, obtém-se a seguinte função Lagrangiana barreira logarítmica modificada:

$$\mathcal{L}_\mu(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f(\mathbf{x}) - \mu \sum_{j=1}^p \delta_j \ln(\mu^{-1} s_j + 1) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \nu_j [h_j(\mathbf{x}) + s_j]. \tag{5.4}$$

As condições necessárias de otimalidade de primeira ordem de KKT são, em seguida, aplicadas sobre a função Lagrangiana (5.4), isto é, $\nabla \mathcal{L}_\mu(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \mathbf{0}$. Definem-se:

$$\bar{S} = \text{diag}(s_1 + \mu, \dots, s_m + \mu) \tag{5.5}$$

$$\bar{S}^{-1} = \text{diag}\left(\frac{1}{s_1 + \mu}, \dots, \frac{1}{s_m + \mu}\right) \tag{5.6}$$

Com isto, obtém-se o sistema não-linear:

$$\begin{cases} \nabla_{\mathbf{x}} \mathcal{L}_\mu(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \nabla f(\mathbf{x}) + J\mathbf{g}(\mathbf{x})^t \boldsymbol{\lambda} + J\mathbf{h}(\mathbf{x})^t \boldsymbol{\nu} = \mathbf{0} \end{cases} \tag{5.7a}$$

$$\begin{cases} \nabla_{\mathbf{s}} \mathcal{L}_\mu(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = -\mu \bar{S}^{-1} \boldsymbol{\delta} + \boldsymbol{\nu} = \mathbf{0} \end{cases} \tag{5.7b}$$

$$\begin{cases} \nabla_{\boldsymbol{\lambda}} \mathcal{L}_\mu(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \mathbf{g}(\mathbf{x}) = \mathbf{0} \end{cases} \tag{5.7c}$$

$$\begin{cases} \nabla_{\boldsymbol{\nu}} \mathcal{L}_\mu(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \mathbf{h}(\mathbf{x}) + \mathbf{s} = \mathbf{0} \end{cases} \tag{5.7d}$$

no qual $J\mathbf{g}(\mathbf{x})$ e $J\mathbf{h}(\mathbf{x})$ são as matrizes jacobianas dos funcionais $\mathbf{g}$ e $\mathbf{h}$, respectivamente.

De acordo com Pinheiro (2012), a condição de folgas complementares entre as variáveis primais $\mathbf{s}$ e duais $\boldsymbol{\nu}$ está associada ao produto $\bar{S}\boldsymbol{\nu}$. Por este motivo, o sistema (5.7) usualmente é reescrito como:

$$\begin{cases} \nabla f(\mathbf{x}) + J\mathbf{g}(\mathbf{x})^t \boldsymbol{\lambda} + J\mathbf{h}(\mathbf{x})^t \boldsymbol{\nu} = \mathbf{0} \end{cases} \tag{5.8a}$$

$$\begin{cases} \bar{S}\boldsymbol{\nu} - \mu \boldsymbol{\delta} = \mathbf{0} \end{cases} \tag{5.8b}$$

$$\begin{cases} \mathbf{g}(\mathbf{x}) = \mathbf{0} \end{cases} \tag{5.8c}$$

$$\begin{cases} \mathbf{h}(\mathbf{x}) + \mathbf{s} = \mathbf{0} \end{cases} \tag{5.8d}$$

O método PDPCBLM é um método iterativo, o qual gera uma sequência de variáveis primais e duais a partir de um ponto inicial, até que a solução ótima seja atingida. Dada uma solução numa iteração k , $(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)$, a partir da mesma é calculada uma direção de busca, $(\Delta \mathbf{x}^k, \Delta \mathbf{s}^k, \Delta \boldsymbol{\lambda}^k, \Delta \boldsymbol{\nu}^k)$, a qual permite, após o cálculo do tamanho do passo, a obtenção de um novo ponto $(\mathbf{x}^{k+1}, \mathbf{s}^{k+1}, \boldsymbol{\lambda}^{k+1}, \boldsymbol{\nu}^{k+1})$. A fim de determinar as direções, o sistema (5.8) é linearizado através da fórmula de Taylor de primeira ordem. Nas próximas subseções é apresentado o processo de linearização de cada equação.

5.1.1.1 Linearização da equação (5.8a)

A fim de realizar a linearização, considera-se que:

$$(\mathbf{x}^{k+1}, \mathbf{s}^{k+1}, \boldsymbol{\lambda}^{k+1}, \boldsymbol{\nu}^{k+1}) = (\mathbf{x}^k + \Delta \mathbf{x}^k, \mathbf{s}^k + \Delta \mathbf{s}^k, \boldsymbol{\lambda}^k + \Delta \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k + \Delta \boldsymbol{\nu}^k).$$

A equação (5.8a) pode ser reescrita como:

$$\nabla f(\mathbf{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}) + \sum_{j=1}^p \nu_j \nabla h_j(\mathbf{x}) = \mathbf{0}. \quad (5.9)$$

Aplicando a fórmula de Taylor de primeira ordem na vizinhança de $\mathbf{x}^k$ e desconsiderando o resto:

$$\nabla f(\mathbf{x}^k + \Delta \mathbf{x}^k) + \sum_{i=1}^m (\lambda_i^k + \Delta \lambda_i^k) \nabla g_i(\mathbf{x}^k + \Delta \mathbf{x}^k) + \sum_{j=1}^p (\nu_j^k + \Delta \nu_j^k) \nabla h_j(\mathbf{x}^k + \Delta \mathbf{x}^k) = \mathbf{0}$$

ou seja,

$$\begin{aligned} \nabla f(\mathbf{x}^k) + \nabla^2 f(\mathbf{x}^k) \Delta \mathbf{x}^k + \sum_{i=1}^m (\lambda_i^k + \Delta \lambda_i^k) [\nabla g_i(\mathbf{x}^k) + \nabla^2 g_i(\mathbf{x}^k) \Delta \mathbf{x}^k] \\ + \sum_{j=1}^p (\nu_j^k + \Delta \nu_j^k) [\nabla h_j(\mathbf{x}^k) + \nabla^2 h_j(\mathbf{x}^k) \Delta \mathbf{x}^k] = \mathbf{0} \end{aligned}$$

Isto resulta em

$$\begin{aligned} \nabla f(\mathbf{x}^k) + \nabla^2 f(\mathbf{x}^k) \Delta \mathbf{x}^k \\ + \sum_{i=1}^m \lambda_i^k \nabla g_i(\mathbf{x}^k) + \sum_{i=1}^m \lambda_i^k \nabla^2 g_i(\mathbf{x}^k) \Delta \mathbf{x}^k + \sum_{i=1}^m \Delta \lambda_i^k \nabla g_i(\mathbf{x}^k) + \sum_{i=1}^m \Delta \lambda_i^k \nabla^2 g_i(\mathbf{x}^k) \Delta \mathbf{x}^k \\ + \sum_{j=1}^p \nu_j^k \nabla h_j(\mathbf{x}^k) + \sum_{j=1}^p \nu_j^k \nabla^2 h_j(\mathbf{x}^k) \Delta \mathbf{x}^k + \sum_{j=1}^p \Delta \nu_j^k \nabla h_j(\mathbf{x}^k) + \sum_{j=1}^p \Delta \nu_j^k \nabla^2 h_j(\mathbf{x}^k) \Delta \mathbf{x}^k = \mathbf{0} \end{aligned}$$

Desconsiderando os termos mistos que envolvem produtos de direções, $\sum_{i=1}^m \Delta \lambda_i^k \nabla^2 g_i(\mathbf{x}^k) \Delta \mathbf{x}^k$ e $\sum_{j=1}^p \Delta \nu_j^k \nabla^2 h_j(\mathbf{x}^k) \Delta \mathbf{x}^k$, obtém-se:

$$\begin{aligned} & \nabla f(\mathbf{x}^k) + \nabla^2 f(\mathbf{x}^k) \Delta \mathbf{x}^k + \sum_{i=1}^m \lambda_i^k \nabla g_i(\mathbf{x}^k) + \sum_{i=1}^m \lambda_i^k \nabla^2 g_i(\mathbf{x}^k) \Delta \mathbf{x}^k \\ & + \sum_{i=1}^m \Delta \lambda_i^k \nabla g_i(\mathbf{x}^k) + \sum_{j=1}^p \nu_j^k \nabla h_j(\mathbf{x}^k) + \sum_{j=1}^p \nu_j^k \nabla^2 h_j(\mathbf{x}^k) \Delta \mathbf{x}^k + \sum_{j=1}^p \Delta \nu_j^k \nabla h_j(\mathbf{x}^k) = \mathbf{0}. \end{aligned}$$

Agrupando os termos relacionados às direções, tem-se:

$$\begin{aligned} & \left[\nabla^2 f(\mathbf{x}^k) + \sum_{i=1}^m \lambda_i^k \nabla^2 g_i(\mathbf{x}^k) + \sum_{j=1}^p \nu_j^k \nabla^2 h_j(\mathbf{x}^k) \right] \Delta \mathbf{x}^k + \sum_{i=1}^m \Delta \lambda_i^k \nabla g_i(\mathbf{x}^k) \\ & + \sum_{j=1}^p \Delta \nu_j^k \nabla h_j(\mathbf{x}^k) + \nabla f(\mathbf{x}^k) + \sum_{i=1}^m \lambda_i^k \nabla g_i(\mathbf{x}^k) + \sum_{j=1}^p \nu_j^k \nabla h_j(\mathbf{x}^k) = \mathbf{0}. \end{aligned}$$

Passando para o segundo membro os termos que não dependem das direções:

$$\begin{aligned} & \left[\nabla^2 f(\mathbf{x}^k) + \sum_{i=1}^m \lambda_i^k \nabla^2 g_i(\mathbf{x}^k) + \sum_{j=1}^p \nu_j^k \nabla^2 h_j(\mathbf{x}^k) \right] \Delta \mathbf{x}^k + \sum_{i=1}^m \Delta \lambda_i^k \nabla g_i(\mathbf{x}^k) \\ & + \sum_{j=1}^p \Delta \nu_j^k \nabla h_j(\mathbf{x}^k) = -\nabla f(\mathbf{x}^k) - \sum_{i=1}^m \lambda_i^k \nabla g_i(\mathbf{x}^k) - \sum_{j=1}^p \nu_j^k \nabla h_j(\mathbf{x}^k). \end{aligned}$$

Observando que $\nabla_{\mathbf{x}\mathbf{x}}^2 \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k) = \nabla^2 f(\mathbf{x}^k) + \sum_{i=1}^m \lambda_i^k \nabla^2 g_i(\mathbf{x}^k) + \sum_{j=1}^p \nu_j^k \nabla^2 h_j(\mathbf{x}^k)$, a equação anterior pode ser reescrita da seguinte forma compacta:

$$\nabla_{\mathbf{x}\mathbf{x}}^2 \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k) \Delta \mathbf{x}^k + J\mathbf{g}(\mathbf{x}^k)^t \Delta \boldsymbol{\lambda}^k + J\mathbf{h}(\mathbf{x}^k)^t \Delta \boldsymbol{\nu}^k = \mathbf{m}^k, \quad (5.10)$$

em que:

$$\mathbf{m}^k = -\nabla f(\mathbf{x}^k) - J\mathbf{g}(\mathbf{x}^k)^t \boldsymbol{\lambda}^k - J\mathbf{h}(\mathbf{x}^k)^t \boldsymbol{\nu}^k. \quad (5.11)$$

Segundo Pinheiro (2012), o termo $\mathbf{m}^k$ é denominado *resíduo de viabilidade dual*.

5.1.1.2 Linearização da equação (5.8b)

Agora, efetua-se a linearização da equação (5.8b). Para tanto, considera-se uma linha j da equação $\bar{S}\boldsymbol{\nu} - \mu\boldsymbol{\delta} = \mathbf{0}$, a qual assume a forma:

$$(s_j + \mu)\nu_j - \mu\delta_j = 0. \quad (5.12)$$

Linearizando cada equação, tem-se que:

$$(s_j^{k+1} + \mu)\nu_j^{k+1} - \mu\delta_j^k = 0, \text{ ou seja, } (s_j^k + \Delta s_j^k + \mu)(\nu_j^k + \Delta \nu_j^k) - \mu\delta_j^k = 0.$$

Esta expressão pode ser reescrita como:

$$(s_j^k + \mu)\nu_j^k + (s_j^k + \mu)\Delta \nu_j^k + \nu_j^k \Delta s_j^k + \Delta s_j^k \Delta \nu_j^k - \mu\delta_j^k = 0. \quad (5.13)$$

Tendo em vista a igualdade (5.13), a linearização para todas as equações em (5.8b) pode ser representada por:

$$\overline{S}_k \boldsymbol{\nu}^k + \overline{S}_k \Delta \boldsymbol{\nu}^k + \mathbf{N}^k \Delta \mathbf{s}^k + \Delta \mathbf{s}^k \circ \Delta \boldsymbol{\nu}^k - \mu \boldsymbol{\delta}^k = \mathbf{0}, \quad (5.14)$$

em que $\mathbf{N}^k = \text{diag}(\boldsymbol{\nu}^k)$ e $\Delta \mathbf{s}^k \circ \Delta \boldsymbol{\nu}^k$ denota o produto de Hadamard, isto é, o produto componente-a-componente, entre os vetores $\Delta \mathbf{s}^k$ e $\Delta \boldsymbol{\nu}^k$. A igualdade em (5.14) pode ser escrita como:

$$\mathbf{N}^k \Delta \mathbf{s}^k + \overline{S}_k \Delta \boldsymbol{\nu}^k = \boldsymbol{\pi}^k, \quad (5.15)$$

onde

$$\boldsymbol{\pi}^k = -\overline{S}_k \boldsymbol{\nu}^k + \mu \boldsymbol{\delta}^k - \Delta \mathbf{s}^k \circ \Delta \boldsymbol{\nu}^k. \quad (5.16)$$

5.1.1.3 Linearização da equação (5.8c)

A linearização da equação (5.8c) pode ser feita da seguinte maneira:

$$\mathbf{g}(\mathbf{x}^{k+1}) = \mathbf{0}, \text{ ou seja, } \mathbf{g}(\mathbf{x}^k + \Delta \mathbf{x}^k) = \mathbf{0}.$$

Disto, segue que:

$$\mathbf{g}(\mathbf{x}^k) + J\mathbf{g}(\mathbf{x}^k)\Delta \mathbf{x}^k = \mathbf{0}. \quad (5.17)$$

A igualdade expressa por (5.17) pode ser reescrita como:

$$J\mathbf{g}(\mathbf{x}^k)\Delta \mathbf{x}^k = \mathbf{t}^k, \quad (5.18)$$

em que

$$\mathbf{t}^k = -\mathbf{g}(\mathbf{x}^k). \quad (5.19)$$

5.1.1.4 Linearização da equação (5.8d)

A linearização da equação (5.8d) é realizada de modo semelhante ao anterior:

$$\mathbf{h}(\mathbf{x}^{k+1}) + \mathbf{s}^{k+1} = \mathbf{0}, \text{ ou seja, } \mathbf{h}(\mathbf{x}^k + \Delta \mathbf{x}^k) + \mathbf{s}^k + \Delta \mathbf{s}^k = \mathbf{0}.$$

Disto, segue que:

$$\mathbf{h}(\mathbf{x}^k) + J\mathbf{h}(\mathbf{x}^k)\Delta \mathbf{x}^k + \mathbf{s}^k + \Delta \mathbf{s}^k = \mathbf{0}. \quad (5.20)$$

A igualdade apresentada na equação (5.20) pode ser reescrita como:

$$J\mathbf{h}(\mathbf{x}^k)\Delta \mathbf{x}^k + \Delta \mathbf{s}^k = \mathbf{u}^k, \quad (5.21)$$

em que

$$\mathbf{u}^k = -\mathbf{h}(\mathbf{x}^k) - \mathbf{s}^k. \quad (5.22)$$

5.1.1.5 O sistema de direções

Tendo em vista as equações (5.10), (5.15), (5.18) e (5.21), as direções de busca podem ser determinadas através da resolução do seguinte sistema:

$$\begin{pmatrix} K & 0 & J\mathbf{g}(\mathbf{x}^k)^t & J\mathbf{h}(\mathbf{x}^k)^t \\ 0 & \mathbf{N}^k & 0 & \bar{\mathbf{S}}_k \\ J\mathbf{g}(\mathbf{x}^k) & 0 & 0 & 0 \\ J\mathbf{h}(\mathbf{x}^k) & I_p & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta \mathbf{x}^k \\ \Delta \mathbf{s}^k \\ \Delta \boldsymbol{\lambda}^k \\ \Delta \boldsymbol{\nu}^k \end{pmatrix} = \begin{pmatrix} \mathbf{m}^k \\ \boldsymbol{\pi}^k \\ \mathbf{t}^k \\ \mathbf{u}^k \end{pmatrix} \quad (5.23)$$

em que I_p é a matriz identidade de ordem p e $K = \nabla_{\mathbf{x}\mathbf{x}}^2 \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)$. De acordo com Nocedal & Wright (2006), frequentemente o sistema (5.23) é reescrito em sua forma simétrica, através da premultiplicação da segunda linha por $\bar{\mathbf{S}}_k^{-1}$:

$$\begin{pmatrix} K & 0 & J\mathbf{g}(\mathbf{x}^k)^t & J\mathbf{h}(\mathbf{x}^k)^t \\ 0 & \bar{\mathbf{S}}_k^{-1} \mathbf{N}^k & 0 & I_p \\ J\mathbf{g}(\mathbf{x}^k) & 0 & 0 & 0 \\ J\mathbf{h}(\mathbf{x}^k) & I_p & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta \mathbf{x}^k \\ \Delta \mathbf{s}^k \\ \Delta \boldsymbol{\lambda}^k \\ \Delta \boldsymbol{\nu}^k \end{pmatrix} = \begin{pmatrix} \mathbf{m}^k \\ \bar{\mathbf{S}}_k^{-1} \boldsymbol{\pi}^k \\ \mathbf{t}^k \\ \mathbf{u}^k \end{pmatrix} \quad (5.24)$$

Esta formulação tem a vantagem de permitir o uso de um pacote computacional para a resolução de sistemas de equações lineares simétricos, o que reduz o esforço computacional em cada iteração do método.

5.1.2 Correção de Inércia

Quando o problema de otimização é não-convexo, as direções de busca determinadas através da resolução do sistema (5.23) ou (5.24) nem sempre são produtivas na busca de mínimos locais. Isto ocorre porque a resolução do sistema apenas visa a busca por pontos que satisfaçam as condições de KKT. Isto significa que não apenas pontos de mínimo local, mas também pontos de máximo e outros pontos estacionários podem ser determinados.

Os resultados expostos no Apêndice B, indicam que, em um problema com restrições de igualdade, um ponto de mínimo local é obtido quando a matriz hessiana da função Lagrangiana no ponto ótimo é definida positiva no espaço tangente das restrições. Um resultado semelhante a este, envolvendo um cone de direções viáveis, é válido para restrições de desigualdade.

Por outro lado, os resultados de dualidade local apresentados no mesmo apêndice indicam que, um ponto satisfazendo as condições de ponto de sela pode ser obtido se a matriz hessiana, em relação a $\mathbf{x}$, da função Lagrangiana no ponto ótimo, $\nabla_{\mathbf{xx}}^2 \mathcal{L}(\mathbf{x}^*, \mathbf{s}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, for definida positiva. Há, ainda, a forte relação entre o problema primal, de minimização, e o seu dual, que é um problema de maximização.

Neste contexto, a definição de inércia de uma matriz é de fundamental importância. Ela se baseia no fato de que, para matrizes simétricas, os autovalores são todos números reais.

Definição 5.1. A *inércia* de uma matriz simétrica A é a tripla ordenada $\mathcal{I}(A) = (i_+, i_-, i_0)$, em que i_+ é o número de autovalores positivos de A , i_- é o número de autovalores negativos, e i_0 é o número de autovalores nulos.

O teorema a seguir é de grande utilidade quando a inércia de uma matriz simétrica precisa ser determinada.

Teorema 5.1. (Teorema da Inércia de Sylvester) Seja A uma matriz simétrica e S uma matriz não singular qualquer de mesma dimensão. Então $\mathcal{I}(A) = \mathcal{I}(SAS^t)$.

Corolário 5.1. Qualquer permutação simétrica de linhas e colunas preserva a inércia de uma matriz.

Segundo Nocedal & Wright (2006), uma direção de descida para o sistema primal-dual (5.24) é obtida se a matriz

$$\begin{pmatrix} K & 0 \\ 0 & \Sigma \end{pmatrix} \quad (5.25)$$

em que $\Sigma = \bar{S}_k^{-1} \mathbf{N}^k$, for definida positiva no núcleo da matriz das restrições

$$\begin{pmatrix} J\mathbf{g}(\mathbf{x}^k) & 0 \\ J\mathbf{h}(\mathbf{x}^k) & I_p \end{pmatrix}. \quad (5.26)$$

Esta condição ocorre quando a matriz

$$A_k = \begin{pmatrix} K & 0 & J\mathbf{g}(\mathbf{x}^k)^t & J\mathbf{h}(\mathbf{x}^k)^t \\ 0 & \Sigma & 0 & 0 \\ J\mathbf{g}(\mathbf{x}^k) & 0 & 0 & 0 \\ J\mathbf{h}(\mathbf{x}^k) & I_p & 0 & 0 \end{pmatrix} \quad (5.27)$$

do sistema primal-dual (5.24) possui inércia

$$\mathcal{I}(A_k) = (n + p, m + p, 0). \quad (5.28)$$

Este resultado apontado por Nocedal & Wright (2006) está em concordância com os resultados de dualidade apresentados anteriormente: em um ponto de sela da função Lagrangiana, no qual o *gap* de dualidade é nulo entre as funções primal e dual, espera-se que a matriz hessiana possua $n + p$ autovalores positivos, relacionados às variáveis primais e ao problema de minimização, e $m + p$ autovalores negativos, relacionados às variáveis duais e ao problema de maximização.

Nocedal & Wright (2006) afirmam que, diante de situações em que a matriz A_k não possui a inércia desejada, uma modificação pode ser feita na mesma, visando favorecer a obtenção de direções de descida ao resolver o sistema. No trabalho de Pinheiro (2012), uma estratégia deste tipo é adotada. Para tanto, o autor realiza uma perturbação sobre a matriz dual normal do problema de otimização, θ_k , que será apresentada posteriormente no texto (vide a equação (5.49)), através da adição de um múltiplo da matriz identidade. Na prática, este tipo de modificação é bastante comum, sendo adotada em diversos pacotes comerciais de otimização, como IPOPT e LOQO (Benson et al., 2004). O objetivo desta modificação é garantir que a condição de positividade apresentada anteriormente seja verificada e, conseqüentemente, que um decréscimo para algum tipo de função mérito ou filtro seja obtido.

Neste trabalho, a modificação realizada visando corrigir a inércia da matriz A_k será baseada em Nocedal & Wright (2006). Primeiramente, nota-se que a matriz $\Sigma = \bar{S}_k^{-1} \mathbf{N}^k$ é sempre definida positiva. Por outro lado, para a matriz $K = \nabla_{\mathbf{x}\mathbf{x}}^2 \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)$ não há a mesma garantia. Portanto, a ideia é substituir esta matriz por $K + \beta I_n$, em que $\beta > 0$ é escolhido suficientemente grande para assegurar que a inércia de A_k é igual à apresentada em (5.28). O valor necessário para o parâmetro β não é conhecido *a priori*, de modo que

valores sucessivamente maiores devem ser testados, até que a inércia desejada seja obtida.

Além disso, a matriz A_k pode se tornar singular devido à deficiência de posto da matriz $J\mathbf{g}(\mathbf{x}^k)$. Este problema pode ser evitado através da utilização de um parâmetro $\gamma \geq 0$ além do termo βI_n apresentado anteriormente. Desta forma, temos a seguinte matriz primal-dual modificada:

$$\tilde{A}_k = \begin{pmatrix} K + \beta I_n & 0 & J\mathbf{g}(\mathbf{x}^k)^t & J\mathbf{h}(\mathbf{x}^k)^t \\ 0 & \Sigma & 0 & I_p \\ J\mathbf{g}(\mathbf{x}^k) & 0 & -\gamma I_m & 0 \\ J\mathbf{h}(\mathbf{x}^k) & I_p & 0 & 0 \end{pmatrix}. \quad (5.29)$$

O sinal “−” para a perturbação $-\gamma I_m$, além de evitar a singularidade pela deficiência de posto de $J\mathbf{g}(\mathbf{x}^k)$, também evita que um número excessivo de autovalores positivos sejam produzidos, o que prejudicaria a obtenção da inércia desejada.

5.1.2.1 Decomposição de Matrizes Simétricas Indefinidas

Até o momento, foram apresentadas as modificações realizadas sobre a matriz A_k vista em (5.27), de modo a obter $\tilde{A}_k$ apresentada em (5.29). Como dito anteriormente, esta modificação é realizada para valores crescentes de β , até que a inércia desejada seja atingida. Faz-se necessário, portanto, apresentar uma forma de computar a inércia de uma matriz simétrica.

No trabalho de Pinheiro (2012), a perturbação da matriz dual normal utilizando múltiplos da identidade é realizada até que a mesma se torne definida positiva. A fim de verificar quando isto ocorre, o autor empregou a decomposição de Cholesky, a qual é uma maneira bastante efetiva de verificar se uma matriz é definida positiva. De fato, uma matriz é definida positiva se, e somente se, a sua decomposição de Cholesky pode ser realizada.

Neste trabalho, o objetivo não é que a matriz A_k seja definida positiva, mas sim indefinida com um número específico de autovalores positivos e negativos. Em geral, uma matriz simétrica indefinida não possui a decomposição de Cholesky. Mesmo quando a decomposição é possível, ela normalmente apresenta instabilidades numéricas. Dessa forma, quando a matriz é simétrica e indefinida, a utilização de outros métodos para calcular a inércia é necessária.

A fim de determinar a inércia de A_k , a decomposição em valores singulares (*Singular Value Decomposition* – *SVD*) é, naturalmente, a forma mais óbvia, uma vez que permite a obtenção dos valores exatos dos autovalores. Entretanto, do ponto de vista computacional, a SVD é cara, em especial considerando que esta decomposição precisa ser realizada

sucessivas vezes, até que se obtenha a inércia planejada.

Quando matrizes simétricas indefinidas precisam ser fatoradas, as decomposições do tipo LDL^t são uma alternativa viável. Nocedal & Wright (2006) apontam que qualquer matriz simétrica $A \in \mathbb{R}^{n \times n}$, definida positiva ou não, pode ser escrita da forma

$$PAP^t = LDL^t \quad (5.30)$$

em que L é uma matriz triangular inferior, D é uma matriz diagonal por blocos, com blocos diagonais de dimensão 1×1 ou 2×2 , e P é uma matriz de permutação. Diferentemente da decomposição de Cholesky, na decomposição LDL^t , em que D é uma matriz diagonal por blocos 1×1 ou 2×2 , pode-se garantir que a fatoração (5.30) sempre existe e pode ser computada de maneira estável.

Pelo Teorema 5.1 (teorema da inércia de Sylvester) e seu corolário, a inércia da matriz A deve ser igual à da matriz diagonal por blocos D . Além disso, os blocos 2×2 da matriz D são sempre construídos de modo que possuam um autovalor positivo e um autovalor negativo. Assim, o número de autovalores positivos da matriz A é igual ao número de blocos positivos 1×1 acrescido do número de blocos 2×2 . Por outro lado, o número de autovalores negativos é igual ao número de blocos negativos 1×1 acrescido do número de blocos 2×2 .

Exemplo 5.1. Considere a matriz simétrica:

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & 1 \\ -1 & 1 & 1 \end{pmatrix}. \quad (5.31)$$

Esta matriz pode ser escrita na forma (5.30), tomando

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 & 0 \\ \frac{2}{3} & 1 & 0 \\ \frac{1}{3} & 0 & 1 \end{pmatrix} \quad \text{e} \quad D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -\frac{1}{3} & -\frac{5}{3} \\ 0 & -\frac{5}{3} & \frac{2}{3} \end{pmatrix} \quad (5.32)$$

Observe que a matriz D possui um bloco 1×1 positivo e um bloco 2×2 , ao qual devem estar associados um autovalor positivo e um negativo. Assim, espera-se que $\mathcal{I}(A) = (2, 1, 0)$. Computando os autovalores da matriz A , nota-se que valem aproximadamente $-1,1326$, $1,8595$ e $4,2730$, confirmando a inércia determinada ao contar os blocos da matriz D .

De acordo com Botta (2009), a decomposição LDL^t é iniciada através de uma matriz de permutação Π , que traz um elemento não-nulo ou um bloco 2×2 não-singular para o

canto superior esquerdo de $\Pi A \Pi^t$, isto é,

$$\Pi A \Pi^t = \begin{matrix} & \begin{matrix} s & n-s \end{matrix} \\ \begin{matrix} s \\ n-s \end{matrix} & \begin{pmatrix} E & C^t \\ C & B \end{pmatrix} \end{matrix}, \quad (5.33)$$

em que E é a matriz quadrada não-singular de dimensão $s = 1$ ou 2 . A matriz E é o elemento pivô utilizado na fatoração. Uma vez determinada a permutação Π , nota-se que, sendo E inversível,

$$\Pi A \Pi^t = \begin{pmatrix} I_s & 0 \\ C E^{-1} & I_{n-s} \end{pmatrix} \begin{pmatrix} E & 0 \\ 0 & B - C E^{-1} C^t \end{pmatrix} \begin{pmatrix} I_s & E^{-1} C^t \\ 0 & I_{n-s} \end{pmatrix}. \quad (5.34)$$

Segundo Nocedal & Wright (2006), o próximo passo na fatoração é aplicar o mesmo procedimento à matriz $B - C E^{-1} C^t$, que é o *complemento de Schur* da matriz E em $\Pi A \Pi^t$. A matriz $B - C E^{-1} C^t$ possui dimensão $(n-1) \times (n-1)$ ou $(n-2) \times (n-2)$, dependendo da dimensão s do bloco E . Assim, a aplicação consecutiva dessa forma de decomposição aos sucessivos complementos de Schur conduz à fatoração (5.30).

A fatoração simétrica indefinida delineada anteriormente apresenta o mesmo número de operações em ponto flutuante que a decomposição de Cholesky, o qual é de aproximadamente $\frac{n^3}{3}$ operações. Entretanto, a este número deve ser acrescentado o número de operações relacionadas à escolha dos elementos pivô e à realização das permutações. Na literatura, diversas estratégias são apresentadas para a escolha do elemento pivô e que influenciam no número de operações realizadas e nas propriedades numéricas do método.

Na estratégia proposta por Bunch & Parlett (1971), denominada Bunch-Parlett, o complemento de Schur é examinado por completo de modo a determinar os elementos diagonal e não-diagonal de maiores magnitudes, ξ_d e ξ_n , respectivamente. Quando a razão $\frac{\xi_d}{\xi_n}$ é aceitável, um pivô do tipo 1×1 é escolhido. Do contrário, um bloco 2×2 , relacionado aos índices da linha e coluna do elemento não-diagonal ξ_n , é escolhido como pivô. Devido à necessidade de determinar ξ_d e ξ_n , cada iteração requer várias comparações entre números, de ordem $O(n^3)$. Por outro lado, além de estável, esta forma de escolha dos pivôs garante que a matriz L possua elementos cujo valor absoluto é limitado por 2,781.

Em Bunch & Kaufman (1977), uma estratégia denominada Bunch-Kaufman, que visa ser mais econômica do ponto de vista do pivoteamento é apresentada. Nela, no máximo duas colunas da matriz são investigadas em cada etapa, reduzindo o número de comparações para $O(n^2)$. Apesar de reduzir o número de operações de comparação, nesta estratégia não existe um limitante para os elementos da matriz L , que podem se tornar arbitrariamente elevados.

Segundo Nocedal & Wright (2006), uma estratégia que visa encontrar um ponto de equilíbrio entre as estratégias de Bunch-Parlett e Bunch-Kaufman é a Bunch-Kaufman Limitada (*Bounded Bunch-Kaufman – BBK*). Nesta estratégia, os elementos da matriz L são monitorados, e a escolha dos blocos pivô de Bunch-Kaufman é feita quando o crescimento provocado nos elementos de L não é excessivo. Se isto não é possível, uma busca mais profunda é realizada para determinar um pivô aceitável. Com isto, o número de comparações realizadas é intermediário entre as estratégias de Bunch-Parlett e Bunch-Kaufman. Na prática, o custo desta estratégia é similar ao de Bunch-Kaufman, mas no pior caso pode se assemelhar ao de Bunch-Parlett.

Neste trabalho, a decomposição da matriz foi realizada utilizando o comando interno do Matlab, denominado `ldl`, que realiza a fatoração LDL^t de uma matriz simétrica. O comando `ldl` utiliza uma rotina da LAPACK (*Linear Algebra Package*), escrita em Fortran 90, para realizar a fatoração LDL^t , através da estratégia de Bunch-Kaufman. Segundo Druinsky & Toledo (2011), a estratégia de pivoteamento sintetizada no Algoritmo 5 é utilizada na matriz inicial e nos sucessivos complementos de Schur pela rotina da LAPACK.

Algoritmo 5: Estratégia de Pivoteamento de Bunch-Kaufman (LAPACK)

```

1  Determine a posição  $r$  e a magnitude  $\lambda$  do maior elemento não diagonal da
   coluna 1;
2  se  $|a_{11}| \geq \alpha\lambda$ , em que  $\alpha \in (0, 1)$  é uma constante então
3    | Escolha  $a_{11}$  como elemento pivô e retorne;
4  senão
5    | Determine a magnitude  $\sigma$  do maior elemento não diagonal da coluna  $r$ ;
6    | se  $|a_{11}| \sigma \geq \alpha\lambda^2$  então
7      | | Escolha  $a_{11}$  como elemento pivô e retorne;
8    | se  $|a_{rr}| \geq \alpha\sigma$  então
9      | | Escolha  $a_{rr}$  como elemento pivô e retorne;
10   senão
11   | | Utilize o pivô  $2 \times 2$ :  $\begin{pmatrix} a_{11} & a_{1r} \\ a_{r1} & a_{rr} \end{pmatrix}$ .
```

Nesta regra de pivoteamento, o parâmetro α geralmente é escolhido para valer $\frac{1+\sqrt{17}}{8}$, pois este valor garante um crescimento igual dos elementos, no pior cenário, em dois passos que utilizem pivô 1×1 ou um passo 2×2 .

5.1.2.2 Algoritmo da Correção de Inércia

Agora que já abordamos uma forma de calcular a inércia de uma matriz, podemos delinear um algoritmo para corrigir a inércia da matriz A_k , através da matriz modificada $\tilde{A}_k$. Neste

trabalho, foi adotada a seguinte heurística para aumento do parâmetro β em (5.29), em cada ciclo de correção de inércia:

$$\beta^{k+1} = \kappa_1 \beta^k, \quad (5.35)$$

em que $\kappa_1 > 1$ é o fator de aumento.

Uma vez que a inércia desejada é obtida, o valor de β é reduzido através de um parâmetro $\kappa_2 > 1$ para a iteração seguinte. Quanto ao parâmetro γ presente na matriz modificada $\tilde{A}_k$, adotou-se $\gamma = 10^{-6}$, sempre que este termo foi necessário, isto é, sempre que a matriz modificada se mostrava próxima da singularidade. Isto porque Nocedal & Wright (2006) sugerem que este termo seja acrescentado quando algum autovalor nulo é detectado na decomposição. Na prática, porém, dificilmente uma matriz é singular: o que ocorre é a “quase” singularidade. Para avaliar o quão próxima uma matriz estava da singularidade, o comando `cond` do Matlab foi utilizado, sendo que o termo γ passou a ser acrescentado quando o número de condicionamento era superior a 10^6 . Esta opção foi feita por praticidade, uma vez que a decomposição LDL^t permite saber o sinal dos autovalores, mas não avaliar o quão próximo da singularidade a matriz fatorada está.

O algoritmo para correção de inércia aqui apresentado e que foi utilizado na obtenção dos resultados deste trabalho, se baseia no trabalho de Nocedal & Wright (2006) que, por sua vez, se baseou em um procedimento mais elaborado descrito por Wächter & Biegler (2006). No Algoritmo 6, proposto a seguir, a principal diferença em relação àquele definido por Nocedal & Wright (2006) é: na correção de inércia proposta por Nocedal & Wright (2006), a inércia da matriz A_k original é sempre testada em cada iteração do método de ponto interior, sendo utilizada caso possua a inércia desejada. No algoritmo aqui proposto isto é feito inicialmente, mas a partir do momento em que o parâmetro β se faz necessário, ele sempre será utilizado nas iterações seguintes (porém é reduzido assintoticamente, através do parâmetro $\kappa_2 > 1$).

Algoritmo 6: Correção de Inércia

- 1 Inicialização: faça $\beta \leftarrow 0$ e escolha $\kappa_1 > 1$ e $\kappa_2 > 1$;

 - 2 Faça $\beta \leftarrow \beta / \kappa_2$ e $\gamma \leftarrow 0$;
 - 3 Realize a decomposição LDL^t da matriz $\tilde{A}_k$;
 - 4 **enquanto** $\mathcal{I}(\tilde{A}_k) \neq (n + p, m + p, 0)$ **faça**
 - 5 $\beta \leftarrow \kappa_1 \beta$;
 - 6 **se** $\tilde{A}_k$ possui autovalores nulos **então**
 - 7 $\gamma \leftarrow 10^{-6}$;
 - 8 **se** $\beta = 0$ **então**
 - 9 $\beta \leftarrow 10^{-4}$;
 - 10 Realize a decomposição LDL^t da matriz $\tilde{A}_k$.
-

A inicialização do Algoritmo 6 foi separada dos demais passos porque somente é realizada no início do código programado, isto é, o valor de β não é reiniciado em zero a cada iteração. O motivo para isto é que, nos testes realizados, verificou-se que se uma matriz necessita de correções, então ela provavelmente também precisará nas iterações subsequentes, ainda que com menor intensidade.

Outro ponto fundamental a se destacar sobre a correção de inércia é que, apesar da mesma em geral permitir que o método convirja apenas para mínimos locais, ela também pode ocasionar a divergência do método, caso os parâmetros não sejam escolhidos de forma apropriada. Por este motivo, pacotes computacionais comerciais que utilizam esta forma de correção também implementam um método de restauração como salvaguarda. Neste trabalho, a implementação realizada não apresenta este tipo de procedimento, porém, os resultados apresentados no Capítulo 6 mostram que ainda assim o método foi capaz de convergir. Agora que a estratégia de correção de inércia já foi abordada, calcular-se-ão as direções de busca utilizando a matriz modificada $\tilde{A}_k$.

5.1.3 Cálculo das Direções

Consideremos, agora, o seguinte sistema primal-dual modificado:

$$\begin{pmatrix} K + \beta I_n & 0 & J\mathbf{g}(\mathbf{x}^k)^t & J\mathbf{h}(\mathbf{x}^k)^t \\ 0 & \bar{S}_k^{-1} \mathbf{N}^k & 0 & I_p \\ J\mathbf{g}(\mathbf{x}^k) & 0 & -\gamma I_m & 0 \\ J\mathbf{h}(\mathbf{x}^k) & I_p & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta \mathbf{x}^k \\ \Delta \mathbf{s}^k \\ \Delta \boldsymbol{\lambda}^k \\ \Delta \boldsymbol{\nu}^k \end{pmatrix} = \begin{pmatrix} \mathbf{m}^k \\ \bar{S}_k^{-1} \boldsymbol{\pi}^k \\ \mathbf{t}^k \\ \mathbf{u}^k \end{pmatrix} \quad (5.36)$$

Este sistema pode, ser escrito da seguinte forma assimétrica:

$$\begin{pmatrix} K + \beta I_n & 0 & J\mathbf{g}(\mathbf{x}^k)^t & J\mathbf{h}(\mathbf{x}^k)^t \\ 0 & \mathbf{N}^k & 0 & \bar{S}_k \\ J\mathbf{g}(\mathbf{x}^k) & 0 & -\gamma I_m & 0 \\ J\mathbf{h}(\mathbf{x}^k) & I_p & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta \mathbf{x}^k \\ \Delta \mathbf{s}^k \\ \Delta \boldsymbol{\lambda}^k \\ \Delta \boldsymbol{\nu}^k \end{pmatrix} = \begin{pmatrix} \mathbf{m}^k \\ \boldsymbol{\pi}^k \\ \mathbf{t}^k \\ \mathbf{u}^k \end{pmatrix} \quad (5.37)$$

Agora, com base no trabalho de Mehrotra (1992), apresentamos o procedimento previsor-corretor do método PDPCBLM.

5.1.3.1 Procedimento Previsor

Um ponto importante a se notar é que, no resíduo de (5.40), dado por

$$\boldsymbol{\pi}^k = -\bar{S}_k \boldsymbol{\nu}^k + \mu \boldsymbol{\delta}^k - \Delta \mathbf{s}^k \circ \Delta \boldsymbol{\nu}^k, \quad (5.38)$$

existem termos não-lineares do tipo $\Delta s_j^k \Delta \nu_j^k$ que não são conhecidos *a priori*. Tendo em vista este fato, no procedimento previsor estes produtos são desconsiderados, isto é, considera-se que $\Delta s_j^k \Delta \nu_j^k = 0$, para todo $j = 1, \dots, p$. Desta forma, utiliza-se o seguinte resíduo no passo previsor:

$$\tilde{\pi}^k = -\bar{S}_k \boldsymbol{\nu}^k + \mu \boldsymbol{\delta}^k. \quad (5.39)$$

Na literatura, usualmente o passo previsor considera que $\mu = 0$ em (5.39), o qual é estimado de maneira adaptativa para o passo corretor. Neste trabalho, entretanto, seguiu-se a ideia apresentada em Pinheiro (2012), que considera o parâmetro de barreira no passo previsor. A justificativa para a utilização do parâmetro de barreira, segundo Pinheiro (2012) é o fato de que o mesmo possui influência nas condições de complementaridade, de modo que é interessante não deprezá-lo no passo previsor. Além disso, como o método proposto inclui uma estratégia de escolha entre direções do previsor ou corretor, a utilização do parâmetro de barreira se faz necessária em ambos os procedimentos, pois não se sabe previamente qual direção será escolhida. Assim, o sistema a ser resolvido no passo previsor é:

$$\begin{pmatrix} K + \beta I_n & 0 & J\mathbf{g}(\mathbf{x}^k)^t & J\mathbf{h}(\mathbf{x}^k)^t \\ 0 & N^k & 0 & \bar{S}_k \\ J\mathbf{g}(\mathbf{x}^k) & 0 & -\gamma I_m & 0 \\ J\mathbf{h}(\mathbf{x}^k) & I_p & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{\Delta}\mathbf{x}^k \\ \tilde{\Delta}\mathbf{s}^k \\ \tilde{\Delta}\boldsymbol{\lambda}^k \\ \tilde{\Delta}\boldsymbol{\nu}^k \end{pmatrix} = \begin{pmatrix} \mathbf{m}^k \\ \tilde{\pi}^k \\ \mathbf{t}^k \\ \mathbf{u}^k \end{pmatrix} \quad (5.40)$$

A fim de resolver (5.40), explora-se sua estrutura esparsa, o que permite a obtenção de expressões para cada uma das direções $\tilde{\Delta}\mathbf{x}^k$, $\tilde{\Delta}\mathbf{s}^k$, $\tilde{\Delta}\boldsymbol{\lambda}^k$ e $\tilde{\Delta}\boldsymbol{\nu}^k$. Da última linha do sistema (5.40), tem-se que:

$$J\mathbf{h}(\mathbf{x}^k)\tilde{\Delta}\mathbf{x}^k + \tilde{\Delta}\mathbf{s}^k = \mathbf{u}^k, \quad (5.41)$$

ou seja,

$$\tilde{\Delta}\mathbf{s}^k = \mathbf{u}^k - J\mathbf{h}(\mathbf{x}^k)\tilde{\Delta}\mathbf{x}^k. \quad (5.42)$$

A segunda linha do sistema (5.40) permite a obtenção da direção $\tilde{\Delta}\boldsymbol{\nu}^k$, como segue:

$$N^k \tilde{\Delta}\mathbf{s}^k + \bar{S}_k \tilde{\Delta}\boldsymbol{\nu}^k = \tilde{\pi}^k, \quad (5.43)$$

e, portanto,

$$\tilde{\Delta}\boldsymbol{\nu}^k = \bar{S}_k^{-1}(\tilde{\pi}^k - N^k \tilde{\Delta}\mathbf{s}^k). \quad (5.44)$$

Para calcular as duas direções remanescentes, $\tilde{\Delta}\mathbf{x}^k$ e $\tilde{\Delta}\boldsymbol{\lambda}^k$, considera-se a primeira linha do sistema (5.40):

$$(K + \beta I_n)\tilde{\Delta}\mathbf{x}^k + J\mathbf{g}(\mathbf{x}^k)^t \tilde{\Delta}\boldsymbol{\lambda}^k + J\mathbf{h}(\mathbf{x}^k)^t \tilde{\Delta}\boldsymbol{\nu}^k = \mathbf{m}^k. \quad (5.45)$$

Substitui-se a direção $\tilde{\Delta}\boldsymbol{\nu}^k$, obtida em (5.44) na equação (5.45):

$$(K + \beta I_n)\tilde{\Delta}\mathbf{x}^k + J\mathbf{g}(\mathbf{x}^k)^t\tilde{\Delta}\boldsymbol{\lambda}^k + J\mathbf{h}(\mathbf{x}^k)^t\bar{S}_k^{-1}(\tilde{\boldsymbol{\pi}}^k - \mathbf{N}^k\tilde{\Delta}\mathbf{s}^k) = \mathbf{m}^k. \quad (5.46)$$

Agora, substituímos o valor de $\tilde{\Delta}\mathbf{s}^k$, determinado em (5.42) na equação (5.46):

$$(K + \beta I_n)\tilde{\Delta}\mathbf{x}^k + J\mathbf{g}(\mathbf{x}^k)^t\tilde{\Delta}\boldsymbol{\lambda}^k + J\mathbf{h}(\mathbf{x}^k)^t\bar{S}_k^{-1}[\tilde{\boldsymbol{\pi}}^k - \mathbf{N}^k(\mathbf{u}^k - J\mathbf{h}(\mathbf{x}^k)\tilde{\Delta}\mathbf{x}^k)] = \mathbf{m}^k, \quad (5.47)$$

donde segue que,

$$\begin{aligned} [K + \beta I_n + J\mathbf{h}(\mathbf{x}^k)^t\bar{S}_k^{-1}\mathbf{N}^k J\mathbf{h}(\mathbf{x}^k)]\tilde{\Delta}\mathbf{x}^k + J\mathbf{g}(\mathbf{x}^k)^t\tilde{\Delta}\boldsymbol{\lambda}^k \\ + J\mathbf{h}(\mathbf{x}^k)^t\bar{S}_k^{-1}(\tilde{\boldsymbol{\pi}}^k - \mathbf{N}^k\mathbf{u}^k) = \mathbf{m}^k. \end{aligned} \quad (5.48)$$

Define-se:

$$\theta_k = K + \beta I_n + J\mathbf{h}(\mathbf{x}^k)^t\bar{S}_k^{-1}\mathbf{N}^k J\mathbf{h}(\mathbf{x}^k) \quad (5.49)$$

e

$$\tilde{\mathbf{p}}^k = J\mathbf{h}(\mathbf{x}^k)^t\bar{S}_k^{-1}(\tilde{\boldsymbol{\pi}}^k - \mathbf{N}^k\mathbf{u}^k). \quad (5.50)$$

Então, a equação (5.48) pode ser reescrita como:

$$\theta_k\tilde{\Delta}\mathbf{x}^k + J\mathbf{g}(\mathbf{x}^k)^t\tilde{\Delta}\boldsymbol{\lambda}^k = \mathbf{m}^k - \tilde{\mathbf{p}}^k. \quad (5.51)$$

Assim, juntando a terceira equação do sistema (5.40) com (5.51), temos o seguinte sistema, conhecido como *sistema reduzido*:

$$\begin{pmatrix} \theta_k & J\mathbf{g}(\mathbf{x}^k)^t \\ J\mathbf{g}(\mathbf{x}^k) & -\gamma I_m \end{pmatrix} \begin{pmatrix} \tilde{\Delta}\mathbf{x}^k \\ \tilde{\Delta}\boldsymbol{\lambda}^k \end{pmatrix} = \begin{pmatrix} \mathbf{m}^k - \tilde{\mathbf{p}}^k \\ \mathbf{t}^k \end{pmatrix} \quad (5.52)$$

Observe que da segunda equação de (5.52), pode-se obter:

$$J\mathbf{g}(\mathbf{x}^k)\tilde{\Delta}\mathbf{x}^k = \gamma\tilde{\Delta}\boldsymbol{\lambda}^k + \mathbf{t}^k. \quad (5.53)$$

Da igualdade (5.51), obtém-se:

$$J\mathbf{g}(\mathbf{x}^k)^t\tilde{\Delta}\boldsymbol{\lambda}^k = \mathbf{m}^k - \tilde{\mathbf{p}}^k - \theta_k\tilde{\Delta}\mathbf{x}^k. \quad (5.54)$$

Multiplicando à esquerda pela matriz θ_k^{-1} , tem-se que (5.54) é equivalente a:

$$\theta_k^{-1}J\mathbf{g}(\mathbf{x}^k)^t\tilde{\Delta}\boldsymbol{\lambda}^k = \theta_k^{-1}(\mathbf{m}^k - \tilde{\mathbf{p}}^k) - \tilde{\Delta}\mathbf{x}^k. \quad (5.55)$$

A multiplicação de (5.55) à esquerda por $J\mathbf{g}(\mathbf{x}^k)$, implica em:

$$J\mathbf{g}(\mathbf{x}^k)\theta_k^{-1}J\mathbf{g}(\mathbf{x}^k)^t\tilde{\Delta}\boldsymbol{\lambda}^k = J\mathbf{g}(\mathbf{x}^k)\theta_k^{-1}(\mathbf{m}^k - \tilde{\mathbf{p}}^k) - J\mathbf{g}(\mathbf{x}^k)\tilde{\Delta}\mathbf{x}^k. \quad (5.56)$$

Substituindo (5.53) em (5.56), tem-se que:

$$J\mathbf{g}(\mathbf{x}^k)\theta_k^{-1}J\mathbf{g}(\mathbf{x}^k)^t\tilde{\Delta}\boldsymbol{\lambda}^k = J\mathbf{g}(\mathbf{x}^k)\theta_k^{-1}(\mathbf{m}^k - \tilde{\mathbf{p}}^k) - (\gamma\tilde{\Delta}\boldsymbol{\lambda}^k + \mathbf{t}^k). \quad (5.57)$$

Agrupando os termos com a direção $\tilde{\Delta}\boldsymbol{\lambda}^k$, conclui-se que:

$$[J\mathbf{g}(\mathbf{x}^k)\theta_k^{-1}J\mathbf{g}(\mathbf{x}^k)^t + \gamma I_m]\tilde{\Delta}\boldsymbol{\lambda}^k = J\mathbf{g}(\mathbf{x}^k)\theta_k^{-1}(\mathbf{m}^k - \tilde{\mathbf{p}}^k) - \mathbf{t}^k. \quad (5.58)$$

Portanto, isolando $\tilde{\Delta}\boldsymbol{\lambda}^k$, obtém-se:

$$\tilde{\Delta}\boldsymbol{\lambda}^k = [J\mathbf{g}(\mathbf{x}^k)\theta_k^{-1}J\mathbf{g}(\mathbf{x}^k)^t + \gamma I_m]^{-1}[J\mathbf{g}(\mathbf{x}^k)\theta_k^{-1}(\mathbf{m}^k - \tilde{\mathbf{p}}^k) - \mathbf{t}^k]. \quad (5.59)$$

Com isto, pode-se calcular $\tilde{\Delta}\mathbf{x}^k$ utilizando (5.51):

$$\tilde{\Delta}\mathbf{x}^k = \theta_k^{-1}[\mathbf{m}^k - \tilde{\mathbf{p}}^k - J\mathbf{g}(\mathbf{x}^k)^t\tilde{\Delta}\boldsymbol{\lambda}^k]. \quad (5.60)$$

5.1.3.2 Procedimento Corretor

De posse das direções do passo previsor, procede-se com o cálculo das direções corrigidas. Para tanto, as direções determinadas no procedimento previsor são utilizadas, de modo a considerar os termos de segunda ordem no resíduo de complementaridade:

$$\boldsymbol{\pi}^k = -\bar{S}_k\boldsymbol{\nu}^k + \mu\boldsymbol{\delta}^k - \tilde{\Delta}\mathbf{s}^k \circ \tilde{\Delta}\boldsymbol{\nu}^k. \quad (5.61)$$

Neste caso, o sistema a ser resolvido é:

$$\begin{pmatrix} K + \beta I_n & 0 & J\mathbf{g}(\mathbf{x}^k)^t & J\mathbf{h}(\mathbf{x}^k)^t \\ 0 & N^k & 0 & \bar{S}_k \\ J\mathbf{g}(\mathbf{x}^k) & 0 & -\gamma I_m & 0 \\ J\mathbf{h}(\mathbf{x}^k) & I_p & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta\mathbf{x}^k \\ \Delta\mathbf{s}^k \\ \Delta\boldsymbol{\lambda}^k \\ \Delta\boldsymbol{\nu}^k \end{pmatrix} = \begin{pmatrix} \mathbf{m}^k \\ \boldsymbol{\pi}^k \\ \mathbf{t}^k \\ \mathbf{u}^k \end{pmatrix}, \quad (5.62)$$

Procedendo de modo análogo ao do passo previsor, pode-se determinar as seguintes direções corrigidas:

$$\Delta\boldsymbol{\lambda}^k = [J\mathbf{g}(\mathbf{x}^k)\theta_k^{-1}J\mathbf{g}(\mathbf{x}^k)^t + \gamma I_m]^{-1}[J\mathbf{g}(\mathbf{x}^k)\theta_k^{-1}(\mathbf{m}^k - \mathbf{p}^k) - \mathbf{t}^k] \quad (5.63)$$

$$\Delta\mathbf{x}^k = \theta_k^{-1}[\mathbf{m}^k - \mathbf{p}^k - J\mathbf{g}(\mathbf{x}^k)^t\Delta\boldsymbol{\lambda}^k] \quad (5.64)$$

$$\Delta \mathbf{s}^k = \mathbf{u}^k - J\mathbf{h}(\mathbf{x}^k)\Delta \mathbf{x}^k \quad (5.65)$$

$$\Delta \boldsymbol{\nu}^k = \bar{S}_k^{-1}(\boldsymbol{\pi}^k - N^k \Delta \mathbf{s}^k) \quad (5.66)$$

em que $\theta_k = K + \beta I_n + J\mathbf{h}(\mathbf{x}^k)^t \bar{S}_k^{-1} N^k J\mathbf{h}(\mathbf{x}^k)$ conforme definido em (5.49) e

$$\mathbf{p}^k = J\mathbf{h}(\mathbf{x}^k)^t \bar{S}_k^{-1}(\boldsymbol{\pi}^k - N^k \mathbf{u}^k). \quad (5.67)$$

5.1.4 Atualização do parâmetro de barreira, cálculo do novo ponto e do comprimento do passo

O parâmetro de barreira é atualizado pela seguinte heurística de redução:

$$\mu^{k+1} = \tau_1 \mu^k \quad (5.68)$$

em que $\tau_1 \in (0, 1)$. O valor inicial do parâmetro de barreira, μ^0 , deve ser tal que a condição $\mathbf{s}^0 > -\mu$ seja satisfeita, isto é, de modo que as folgas, no início do método, satisfaçam estritamente a restrição de não-negatividade relaxada (isto é, estão no interior da região relaxada). De acordo com Stanzani (2012), o valor τ_1 não deve ser escolhido muito pequeno, a fim de evitar o comportamento oscilatório do método, nem muito grande, evitando a parada prematura do método em pontos que não sejam ótimos.

A direção determinada pelo procedimento predictor nos métodos de pontos interiores clássicos é sempre uma direção de descida para uma função de mérito, conforme destacado por Shanno & Vanderbei (2000). A ideia de utilizar a matriz do sistema de direções várias vezes para determinar direções aperfeiçoadas ganhou popularidade a partir do método predictor-corretor desenvolvido por Mehrotra (1992). Este método, inicialmente desenvolvido para problemas de programação linear, logo foi estendido para problemas de programação quadrática convexa. Entretanto, conforme destacado por Shanno & Vanderbei (2000), em problemas de programação não-linear não-convexos, a direção do procedimento corretor pode não ser adequada.

Investigações realizadas no trabalho de Nocedal & Waltz (2005) apontam que a utilização do passo corretor pode ser prejudicial ao método, inclusive no caso linear e quadrático. Segundo o autor, este efeito prejudicial usualmente se manifesta no aumento do valor da complementaridade, a qual deveria se reduzir ao longo das iterações. Em seu trabalho, Shanno & Vanderbei (2000) apresentam um algoritmo em que a direção do procedimento corretor somente é utilizada quando as correções de posto sobre as matrizes do sistema não aumentam por um fator (igual a 100) na passagem de uma iteração para outra. Caso isto ocorra, o procedimento corretor é completamente abandonado, e somente as direções

do tipo previsor são utilizadas.

Tendo em vista as observações destes autores, neste trabalho propõe-se atualizar a solução com o passo previsor ou corretor, dependendo dos valores de complementaridade obtidos. Assim, após calcular as direções dos passos previsor e corretor, calculam-se duas soluções provisórias, uma para o previsor, expressa em (5.69)

$$\begin{aligned} \mathbf{x}_{\text{prev}}^+ &= \mathbf{x}^k + \alpha_{P,\text{prev}}^k \tilde{\Delta} \mathbf{x}^k \\ \mathbf{s}_{\text{prev}}^+ &= \mathbf{s}^k + \alpha_{P,\text{prev}}^k \tilde{\Delta} \mathbf{s}^k \\ \boldsymbol{\lambda}_{\text{prev}}^+ &= \boldsymbol{\lambda}^k + \alpha_{D,\text{prev}}^k \tilde{\Delta} \boldsymbol{\lambda}^k \\ \boldsymbol{\nu}_{\text{prev}}^+ &= \boldsymbol{\nu}^k + \alpha_{D,\text{prev}}^k \tilde{\Delta} \boldsymbol{\nu}^k \end{aligned} \quad (5.69)$$

e outra para o corretor, expressa em (5.70)

$$\begin{aligned} \mathbf{x}_{\text{cor}}^+ &= \mathbf{x}^k + \alpha_{P,\text{cor}}^k \Delta \mathbf{x}^k \\ \mathbf{s}_{\text{cor}}^+ &= \mathbf{s}^k + \alpha_{P,\text{cor}}^k \Delta \mathbf{s}^k \\ \boldsymbol{\lambda}_{\text{cor}}^+ &= \boldsymbol{\lambda}^k + \alpha_{D,\text{cor}}^k \Delta \boldsymbol{\lambda}^k \\ \boldsymbol{\nu}_{\text{cor}}^+ &= \boldsymbol{\nu}^k + \alpha_{D,\text{cor}}^k \Delta \boldsymbol{\nu}^k \end{aligned} \quad (5.70)$$

Os escalares $\alpha_{P,\text{prev}}$ e $\alpha_{D,\text{prev}}$ são denominados passos primal e dual, respectivamente, para o procedimento previsor. Já os escalares $\alpha_{P,\text{cor}}$ e $\alpha_{D,\text{cor}}$ são denominados passos primal e dual, respectivamente, para o procedimento corretor. Diversas estratégias para a escolha do passo são apresentadas na literatura. Neste trabalho, uma versão adaptada a partir de Granville (1994) para levar em conta a utilização da função barreira logarítmica modificada é utilizada.

Estes passos devem ser determinados a partir das condições de fronteira da região factível. Isto é fácil de ser feito para as variáveis duais, mas deve ser feito com um cuidado maior para as variáveis primais quando se utiliza a função barreira logarítmica modificada, tendo em vista que as condições de fronteira relacionadas às variáveis de folga, $\mathbf{s}$, variam a cada iteração, juntamente com o parâmetro de barreira. Por simplicidade de notação, omitiremos os termos indicativos do previsor e corretor na representação do passo e direções de busca na discussão a seguir.

Considerando que desejamos que o próximo ponto determinado seja factível, para cada coordenada $j = 1, \dots, p$ deve-se ter:

$$s_j^{k+1} > -\mu^{k+1}, \quad (5.71)$$

isto é,

$$s_j^k + \alpha_P^k \Delta s_j^k > -\tau_1 \mu^k, \quad (5.72)$$

da qual se pode concluir que

$$\alpha_P^k \Delta s_j^k > -(s_j^k + \tau_1 \mu^k). \quad (5.73)$$

Assumindo que $s_j^k + \tau_1 \mu^k > 0$, a desigualdade é trivialmente verdadeira quando $\Delta s_j^k \geq 0$. Assim, o passo somente é calculado quando $\Delta s_j^k < 0$. Neste caso, deve-se ter:

$$\alpha_P^k < -\frac{s_j^k + \tau_1 \mu^k}{\Delta s_j^k}. \quad (5.74)$$

A fim de garantir que a fronteira da região relaxada não seja atingida, um parâmetro $\sigma \in (0, 1)$ usualmente é utilizado. Dessa forma, o passo primal para o previsor e o corretor é calculado, respectivamente, por:

$$\alpha_{P,\text{prev}}^k = \sigma \min \left\{ 1, -\frac{(s_j^k + \tau_1 \mu^k)}{\tilde{\Delta} s_j^k} : \tilde{\Delta} s_j^k < 0 \right\}, \quad (5.75)$$

$$\alpha_{P,\text{cor}}^k = \sigma \min \left\{ 1, -\frac{(s_j^k + \tau_1 \mu^k)}{\Delta s_j^k} : \Delta s_j^k < 0 \right\}. \quad (5.76)$$

Já o passo dual para o previsor e o corretor é calculado, respectivamente, por:

$$\alpha_{D,\text{prev}}^k = \sigma \min \left\{ 1, -\frac{\nu_j^k}{\tilde{\Delta} \nu_j^k} : \tilde{\Delta} \nu_j^k < 0 \right\}, \quad (5.77)$$

$$\alpha_{D,\text{cor}}^k = \sigma \min \left\{ 1, -\frac{\nu_j^k}{\Delta \nu_j^k} : \Delta \nu_j^k < 0 \right\}. \quad (5.78)$$

No caso de $s_j^k + \tau_1 \mu^k \leq 0$ para algum j , problemas podem ocorrer. Este é o caso em que a solução atual deixaria de estar no interior da região factível relaxada ao atualizar o parâmetro de barreira pela heurística (5.68). Dois problemas principais podem vir a ocorrer diante desse cenário: o comprimento do passo estar limitado por alguma outra coordenada, de modo que impossibilitaria que o novo ponto se tornasse factível para o interior da nova região relaxada; ou porque a direção gerada pelo método possui a j -ésima componente negativa, o que significaria que nenhum passo dado permitiria que a desigualdade (5.73) ocorresse. Por este motivo, Nash et al. (1994) propõe uma extrapolação nos termos logarítmicos, de modo que a função barreira modificada seja definida para todo o espaço. Neste trabalho, entretanto, preferiu-se adotar a estratégia mais simples apresentada por Pinheiro (2012): realizar uma nova relaxação da região factível, se ocorrer $s_j^{k+1} \leq -\tau_1 \mu^k$ para algum $j = 1, \dots, p$. Neste caso, o parâmetro de barreira é atualizado por:

$$\mu^{k+1} = -(1 + \tau_2) \min_{j=1, \dots, p} \{s_j^{k+1}\}. \quad (5.79)$$

em que $\tau_2 > 0$ é um parâmetro.

Assim, uma das contribuições deste trabalho é de que o novo ponto seja determinado com base nos valores provisórios obtidos para as direções do previsor e do corretor, de acordo com o seguinte critério, em que $\chi > 0$ é o parâmetro que controla a escolha entre os passos previsor ou corretor, com base no valor de complementaridade:

- se $(\mathbf{s}_{\text{prev}}^+)^t \boldsymbol{\nu}_{\text{prev}}^+ < \chi (\mathbf{s}_{\text{cor}}^+)^t \boldsymbol{\nu}_{\text{cor}}^+$, então:

$$\begin{aligned} \mathbf{x}^{k+1} &= \mathbf{x}_{\text{prev}}^+ \\ \mathbf{s}^{k+1} &= \mathbf{s}_{\text{prev}}^+ \\ \boldsymbol{\lambda}^{k+1} &= \boldsymbol{\lambda}_{\text{prev}}^+ \\ \boldsymbol{\nu}^{k+1} &= \boldsymbol{\nu}_{\text{prev}}^+ \end{aligned} \tag{5.80}$$

- caso contrário:

$$\begin{aligned} \mathbf{x}^{k+1} &= \mathbf{x}_{\text{cor}}^+ \\ \mathbf{s}^{k+1} &= \mathbf{s}_{\text{cor}}^+ \\ \boldsymbol{\lambda}^{k+1} &= \boldsymbol{\lambda}_{\text{cor}}^+ \\ \boldsymbol{\nu}^{k+1} &= \boldsymbol{\nu}_{\text{cor}}^+ \end{aligned} \tag{5.81}$$

5.1.5 Critério de parada

Como a maioria dos métodos iterativos, o método PDPCBLM aqui apresentado não determina a solução exata para o problema de otimização não-linear. Assim, é necessário um critério de parada para decidir quando uma solução de boa qualidade, isto é, suficientemente próxima de um ótimo, foi obtida. Neste trabalho, o critério de parada se baseia em Pinheiro (2012), que considera que uma boa solução foi obtida quando os resíduos forem suficientemente pequenos. Assim, dada uma precisão $\varepsilon > 0$, o critério de parada é dado por:

- Factibilidade Primal:

$$\|\mathbf{t}^k\|_\infty \leq \varepsilon \text{ e } \|\mathbf{u}^k\|_\infty \leq \varepsilon. \tag{5.82}$$

- Factibilidade Dual:

$$\|\mathbf{m}^k\|_\infty \leq \varepsilon. \tag{5.83}$$

- Folgas Complementares:

$$\|\boldsymbol{\pi}^k\|_\infty \leq \varepsilon. \tag{5.84}$$

Estas três condições são equivalentes a dizer que:

$$\|\nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\|_{\infty} \leq \varepsilon. \quad (5.85)$$

5.1.6 Algoritmo do Método PDPCBLM

Tendo em vista todos os procedimentos que foram apresentados, o método PDPCBLM pode ser sintetizado no Algoritmo 7. O procedimento de inicialização das variáveis se baseia no trabalho de Pinheiro (2012), em que se assume que os resíduos são nulos para calcular uma estimativa inicial para as variáveis. Outras formas de inicialização de variáveis primais e duais podem ser utilizadas.

5.2 Suavização Hiperbólica

Nesta seção é apresentada a estratégia de suavização utilizada para a resolução do PDE-PV através do método PDPCBLM. Uma importante classe de funções comumente presentes em problemas de otimização é a das funções modulares. Por exemplo, funções de penalidade exata utilizam a norma da violação das restrições para computar o valor da penalização. A função máximo, como vimos, também é utilizada em funções de penalidade, e ela pode ser escrita como uma função envolvendo termos modulares, através da expressão $\max\{x, y\} = \frac{x+y+|x-y|}{2}$.

Problemas que possuem termos modulares em sua formulação usualmente possuem uma dificuldade adicional: a não-diferenciabilidade nos pontos em que o argumento do módulo se anula. Em particular, o PDE-PV se enquadra nesta categoria de problemas, pois a função objetivo de custo total envolvendo o efeito dos pontos de carregamento de válvula apresenta termos modulares, além de ser não-linear e não-convexa. Todas estas características fazem do PDE-PV um problema de otimização com uma grande quantidade de mínimos locais, o que dificulta a obtenção do mínimo global.

Os métodos clássicos de otimização, entre eles o PDPCBLM, requerem a hipótese de diferenciabilidade das funções envolvidas. Apesar de existirem métodos determinísticos para a resolução de problemas não-diferenciáveis, como métodos baseados em subgradientes, neste trabalho optou-se por outro caminho: a suavização dos termos modulares presentes na função objetivo do problema. Isto porque métodos baseados nos gradientes generalizados de Clarke não podem ser utilizados diretamente na análise de problemas cuja função não seja Lipschitz contínua. Ademais, os métodos para otimização não-diferenciável, apesar de estarem em pleno desenvolvimento, por vezes não são muito eficientes.

Algoritmo 7: Método PDPCBLM

- 1 Inicialização: Forneça o valor do parâmetro de barreira inicial $\mu^0 > 0$, seu fator de redução $\tau_1 \in (0, 1)$, fator da nova relaxação $\tau_2 > 0$, fator de escolha entre passo previsor e corretor $\chi > 0$, e estimadores dos multiplicadores de Lagrange $\delta^0 > \mathbf{0}$;
 - 2 Inicialize a variável primal, $\mathbf{x}^0$, bem como as variáveis para a correção de inércia conforme o passo 1 do Algoritmo 6. Faça $k \leftarrow 0$;
 - 3 Calcule: $f(\mathbf{x}^k)$, $\nabla f(\mathbf{x}^k)$, $\mathbf{g}(\mathbf{x}^k)$, $J\mathbf{g}(\mathbf{x}^k)$, $\mathbf{h}(\mathbf{x}^k)$ e $J\mathbf{h}(\mathbf{x}^k)$;
 - 4 Calcule $\mathbf{s}^0 \leftarrow -\mathbf{h}(\mathbf{x}^0)$ e, em seguida, monte as matrizes $\bar{S}_k$ e $\bar{S}_k^{-1}$;
 - 5 Calcule $\boldsymbol{\nu}^0 \leftarrow \mu^0 \bar{S}_k^{-1} \delta^0$ e $\boldsymbol{\lambda}^0 \leftarrow -[J\mathbf{g}(\mathbf{x}^0)J\mathbf{g}(\mathbf{x}^0)^t]^{-1}J\mathbf{g}(\mathbf{x}^0)[\nabla f(\mathbf{x}^0) + J\mathbf{h}(\mathbf{x}^0)^t\boldsymbol{\nu}^0]$;
 - 6 Determine a matriz N^k ;
 - 7 Construa matriz A_k e utilize a Correção de Inércia do Algoritmo 6 para determinar os parâmetros β e γ de modo que $\tilde{A}_k$ tenha a inércia desejada;
 - 8 Calcule os resíduos do previsor: $\mathbf{m}^k$ (5.11), $\tilde{\boldsymbol{\pi}}^k$ (5.39), $\mathbf{t}^k$ (5.19) e $\mathbf{u}^k$ (5.22);
 - 9 Calcule o vetor $\tilde{\mathbf{p}}^k$ e θ_k ;
 - 10 Calcule as direções do procedimento previsor: $\tilde{\Delta}\boldsymbol{\lambda}^k$ (5.59), $\tilde{\Delta}\mathbf{x}^k$ (5.60), $\tilde{\Delta}\mathbf{s}^k$ (5.42) e $\tilde{\Delta}\boldsymbol{\nu}^k$ (5.44);
 - 11 Calcule os passos $\alpha_{P,\text{prev}}^k$ (5.75) e $\alpha_{D,\text{prev}}^k$ (5.77), bem como a solução provisória do previsor, $\mathbf{x}_{\text{prev}}^+$, $\mathbf{s}_{\text{prev}}^+$, $\boldsymbol{\lambda}_{\text{prev}}^+$ e $\boldsymbol{\nu}_{\text{prev}}^+$, utilizando (5.69);
 - 12 Calcule o resíduo $\boldsymbol{\pi}^k$ (5.61) e o vetor $\mathbf{p}^k$ (5.67);
 - 13 Calcule as direções do procedimento corretor: $\Delta\boldsymbol{\lambda}^k$ (5.63), $\Delta\mathbf{x}^k$ (5.64), $\Delta\mathbf{s}^k$ (5.65) e $\Delta\boldsymbol{\nu}^k$ (5.66);
 - 14 Calcule os passos $\alpha_{P,\text{cor}}^k$ (5.76) e $\alpha_{D,\text{cor}}^k$ (5.78), bem como a solução provisória do corretor, $\mathbf{x}_{\text{cor}}^+$, $\mathbf{s}_{\text{cor}}^+$, $\boldsymbol{\lambda}_{\text{cor}}^+$ e $\boldsymbol{\nu}_{\text{cor}}^+$, utilizando (5.70);
 - 15 **enquanto** $\|\mathbf{m}^k\|_\infty > \varepsilon$ ou $\|\boldsymbol{\pi}^k\|_\infty > \varepsilon$ ou $\|\mathbf{t}^k\|_\infty > \varepsilon$ ou $\|\mathbf{u}^k\|_\infty > \varepsilon$ **faça**
 - 16 Atualize a solução, obtendo $\mathbf{x}^{k+1}$, $\mathbf{s}^{k+1}$, $\boldsymbol{\lambda}^{k+1}$ e $\boldsymbol{\nu}^{k+1}$, utilizando (5.80) ou (5.81), de acordo com os valor da complementaridade;
 - 17 Atualize o parâmetro de barreira utilizando (5.68) ou, caso exista algum $s_j^{k+1} \leq -\tau\mu^k$, utilize (5.79);
 - 18 Atualize os estimadores dos multiplicadores de Lagrange: $\delta^{k+1} = \boldsymbol{\nu}^{k+1}$;
 - 19 Faça $k \leftarrow k + 1$;
 - 20 Calcule: $f(\mathbf{x}^k)$, $\nabla f(\mathbf{x}^k)$, $\mathbf{g}(\mathbf{x}^k)$, $J\mathbf{g}(\mathbf{x}^k)$, $\mathbf{h}(\mathbf{x}^k)$ e $J\mathbf{h}(\mathbf{x}^k)$;
 - 21 Monte as matrizes $\bar{S}_k$, $\bar{S}_k^{-1}$ e N^k ;
 - 22 Construa matriz A_k e utilize a Correção de Inércia do Algoritmo 6 para determinar os parâmetros β e γ de modo que $\tilde{A}_k$ tenha a inércia desejada;
 - 23 Calcule os resíduos do previsor: $\mathbf{m}^k$ (5.11), $\tilde{\boldsymbol{\pi}}^k$ (5.39), $\mathbf{t}^k$ (5.19) e $\mathbf{u}^k$ (5.22);
 - 24 Calcule o vetor $\tilde{\mathbf{p}}^k$ e θ_k ;
 - 25 Calcule as direções do procedimento previsor: $\tilde{\Delta}\boldsymbol{\lambda}^k$ (5.59), $\tilde{\Delta}\mathbf{x}^k$ (5.60), $\tilde{\Delta}\mathbf{s}^k$ (5.42) e $\tilde{\Delta}\boldsymbol{\nu}^k$ (5.44);
 - 26 Calcule os passos $\alpha_{P,\text{prev}}^k$ (5.75) e $\alpha_{D,\text{prev}}^k$ (5.77), bem como a solução provisória do previsor, $\mathbf{x}_{\text{prev}}^+$, $\mathbf{s}_{\text{prev}}^+$, $\boldsymbol{\lambda}_{\text{prev}}^+$ e $\boldsymbol{\nu}_{\text{prev}}^+$, utilizando (5.69);
 - 27 Calcule o resíduo $\boldsymbol{\pi}^k$ (5.61) e o vetor $\mathbf{p}^k$ (5.67);
 - 28 Calcule as direções do procedimento corretor: $\Delta\boldsymbol{\lambda}^k$ (5.63), $\Delta\mathbf{x}^k$ (5.64), $\Delta\mathbf{s}^k$ (5.65) e $\Delta\boldsymbol{\nu}^k$ (5.66);
 - 29 Calcule os passos $\alpha_{P,\text{cor}}^k$ (5.76) e $\alpha_{D,\text{cor}}^k$ (5.78), bem como a solução provisória do corretor, $\mathbf{x}_{\text{cor}}^+$, $\mathbf{s}_{\text{cor}}^+$, $\boldsymbol{\lambda}_{\text{cor}}^+$ e $\boldsymbol{\nu}_{\text{cor}}^+$, utilizando (5.70);
 - 30 Retorne a solução ótima: $\mathbf{x}^* = \mathbf{x}^k$, $\mathbf{s}^* = \mathbf{s}^k$, $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}^k$ e $\boldsymbol{\nu}^* = \boldsymbol{\nu}^k$.
-

De acordo com Chen (2012), aproximações suaves para problemas de otimização têm sido estudadas há décadas, incluindo problemas de complementaridade, desigualdades variacionais, problemas de complementaridade de cones de segunda ordem, programação semidefinida, programação semi-infinita, controle ótimo, otimização de autovalores, métodos de penalidade e problemas de programação matemática com restrições de equilíbrio. Tais métodos de suavização são eficientes não apenas em problemas com função objetivo não-diferenciável, mas também em problemas com restrições não-diferenciáveis.

A título de exemplo, no trabalho de Chen & Mangasarian (1996), é apresentada uma forma de suavização para a função $(x)_+ = \max\{0, x\}$, através de um procedimento de suavização paramétrico relacionado a uma função densidade de probabilidade, com um parâmetro η . À medida em que este parâmetro η se aproxima de zero, a função suave para $(x)_+$ se aproxima de sua versão não-diferenciável. A forma de suavização proposta pelos autores parte do fato que $(x)_+ = \int_{-\infty}^x \sigma(y) dy$, em que σ é a função degrau unitário. A função degrau unitário, por sua vez, pode ser escrita como $\sigma(x) = \int_{-\infty}^x \delta(y) dy$, em que δ é a função delta de Dirac. Isto indica que a função $(x)_+$ pode ser obtida pela integração dupla da função delta de Dirac. A fim de obter uma representação suave, os autores aproximam a função delta de Dirac através das funções densidade de probabilidade, parametrizadas em η .

No trabalho de He et al. (2008), o PDE-PV é resolvido através de uma estratégia híbrida envolvendo evolução diferencial e programação quadrática sequencial. A fim de aplicar a programação quadrática sequencial, os autores empregam uma função de suavização entrópica. Nesta estratégia, dadas funções $\psi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciáveis, com $i = 1, \dots, q$, defina a função $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ por $\psi(\mathbf{x}) = \max_{i=1, \dots, q} \psi_i(\mathbf{x})$. Esta função pode ser suavizada por

$$\psi(\mathbf{x}, \eta) = \eta \ln \left(\sum_{i=1}^q e^{\frac{\psi_i(\mathbf{x})}{\eta}} \right). \quad (5.86)$$

De acordo com os autores, pode-se mostrar que

$$\psi(\mathbf{x}) \leq \psi(\mathbf{x}, \eta) \leq \psi(\mathbf{x}) + \eta \ln q, \quad \forall \mathbf{x} \in \mathbb{R}^n \text{ e } \eta > 0. \quad (5.87)$$

Desta forma, $\psi(\mathbf{x}, \eta) \rightarrow \psi(\mathbf{x})$ quando $\eta \rightarrow 0$. Esta função também é apresentada no trabalho de Bertsekas (1999).

Assim, seguindo esta mesma linha de resolução para o problema, apresenta-se a definição de função suavizante de acordo com Chen (2012).

Definição 5.2. Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função contínua. Dizemos que $\tilde{f} : \mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}$ é uma *função suavizante* para f , se $\tilde{f}(\cdot, \eta)$ é uma função continuamente diferenciável em

$\mathbb{R}^n$ para qualquer $\eta > 0$ fixado, e se para todo $\mathbf{x} \in \mathbb{R}^n$,

$$\lim_{\mathbf{z} \rightarrow \mathbf{x}, \eta \downarrow 0} \tilde{f}(\mathbf{z}, \eta) = f(\mathbf{x}). \quad (5.88)$$

Dessa maneira, pode-se definir um método suave para minimização da função f utilizando $\tilde{f}$ conforme o Algoritmo 8. A vantagem dos métodos de suavização é que através dos mesmos podemos resolver problemas de otimização continuamente diferenciáveis, para os quais existe uma rica teoria e métodos de solução eficientes, com garantias de obtenção de um mínimo local ou de um ponto estacionário do problema não-diferenciável original, através da atualização do parâmetro de suavização η . A eficiência, contudo, depende da função de suavização adotada, do método de solução para o problema suave, e do esquema de atualização do parâmetro η_k .

Algoritmo 8: Otimização Suave

- 1 Inicialização: Defina uma função de suavização paramétrica $\tilde{f} : \mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}$ para aproximar f ;
- 2 Iteração interna: Utilize um algoritmo para encontrar uma solução aproximada (não há necessidade de ser uma solução exata) para o problema de otimização suave:

$$\min_{\mathbf{x} \in \Omega} \tilde{f}(\mathbf{x}, \eta_k) \quad (5.89)$$

para algum $\eta_k > 0$ fixo;

- 3 Iteração externa: Atualize η_k para garantir a convergência do método de suavização para um ponto de mínimo local ou ponto estacionário do problema não-diferenciável original.
-

Neste trabalho, ao associar o método PDPCBLM a uma função de suavização, o parâmetro de suavização foi atualizado pela seguinte regra:

$$\eta_{k+1} = \rho_k \eta_k \quad (5.90)$$

em que

$$\rho_k = \frac{1}{\rho_1 + \frac{\rho_2}{k}}, \quad (5.91)$$

com $\rho_1 > 1$ e $\rho_2 \geq 0$. Esta atualização é realizada juntamente com a atualização do parâmetro de barreira, e visa garantir um decréscimo mais acentuado do parâmetro η nas primeiras iterações, e mais fraca ao final do processo. Isto porque se a atualização for realizada de modo inapropriado, o método pode apresentar instabilidades, devido à proximidade da natureza não diferenciável do problema.

5.2.1 A técnica de suavização hiperbólica

Tendo em vista a definição de função de suavização apresentada, nesta seção é introduzida a técnica de suavização hiperbólica, conforme os trabalhos de Xavier & Oliveira (2005), Junior (2010), Souza (2010), Xavier (2010), Souza et al. (2011) e Xavier & Xavier (2011). Na abordagem de suavização hiperbólica, a função módulo $|\cdot| : \mathbb{R} \rightarrow \mathbb{R}$ é suavizada pela função

$$\vartheta(x, \eta) = \sqrt{x^2 + \eta}, \quad (5.92)$$

em que $\eta > 0$ é o parâmetro de suavização. A Figura 5.1 apresenta a função ϑ para diferentes valores do parâmetro de suavização η .

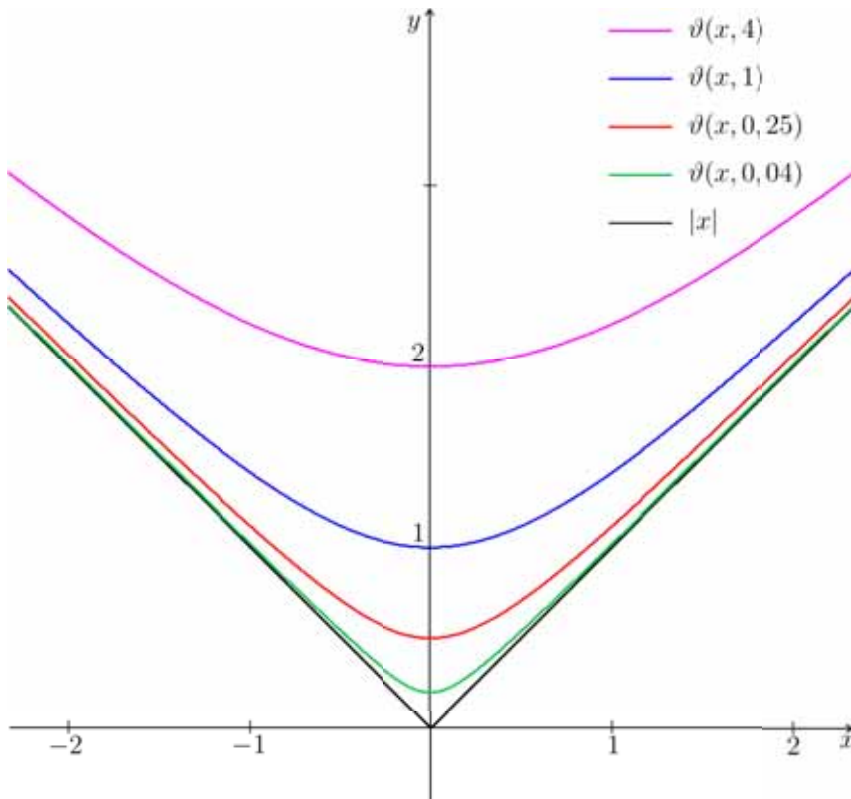


Figura 5.1: Função ϑ para diferentes valores de η .

Seja, agora, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável, e considere a função $|f(\mathbf{x})|$. Suponha, ainda, que esta função possua alguma não-diferenciabilidade, isto é, um ponto em que $f(\mathbf{x})$ muda de sinal. Claramente, esta função pode ser vista como a seguinte composição: $|f(\mathbf{x})| = (|\cdot| \circ f)(\mathbf{x})$. Assim, uma forma suavizada desta função pode ser obtida utilizando a função $\vartheta_\eta \circ f$, em que $\vartheta_\eta(x) = \vartheta(x, \eta)$.

A seguir, são apresentadas as principais propriedades da técnica de suavização hiperbólica.

5.2.2 Propriedades da suavização hiperbólica

A função ϑ definida pela igualdade (5.92) goza das seguintes propriedades:

- (i) $\vartheta(\cdot, \eta)$ é uma função de classe C^∞ , para qualquer $\eta > 0$ fixado;
- (ii) $\lim_{z \rightarrow x, \eta \downarrow 0} \vartheta(z, \eta) = |x|$;
- (iii) $\vartheta(x, \eta) > |x|$, para todo $\eta > 0$;
- (iv) se $\eta_1 > \eta_2$, então $\vartheta(x, \eta_1) > \vartheta(x, \eta_2)$, para todo $x \in \mathbb{R}$;
- (v) $\vartheta(x, \eta) - |x| \leq \sqrt{\eta}$ para todo $\eta > 0$;
- (vi) Se $\{\eta_k\}$ é uma sequência tal que $\lim_{k \rightarrow \infty} \eta_k = 0$, então a sequência de funções $\{\vartheta(\cdot, \eta_k)\}$ converge uniformemente para a função $|\cdot|$.

As propriedades (i) e (ii) indicam que a função ϑ satisfaz a definição de função suavizante proposta por Chen (2012). A propriedade (iii) indica que a função ϑ estima por excesso a função módulo, para qualquer valor adotado para o parâmetro de suavização. A propriedade (iv) significa que a função ϑ é crescente no parâmetro de suavização, isto é, para x fixo, um aumento no parâmetro de suavização provocará um aumento na função aproximante. A propriedade (v) estabelece um limitante para o erro ao utilizar a aproximação da função módulo pela função ϑ , o qual apenas depende do parâmetro η . A propriedade (vi) é uma decorrência imediata da propriedade (v), indicando que uma sequência de funções geradas através da redução do parâmetro de suavização converge uniformemente para a função módulo. A demonstração dessas propriedades é trivial. Para demonstrar a propriedade (v), basta notar que, para quaisquer $x, y \in \mathbb{R}_+$, vale a desigualdade $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$.

Ao resolver um problema suavizado, o parâmetro η desempenha um papel crucial, uma vez que é ele que controla o grau da suavização. Um valor pequeno para η significa a resolução de um problema próximo da não-diferenciabilidade e, portanto, bastante difícil e mal-condicionado próximo das regiões em que o termo modular se anula. Por outro lado, a escolha de um valor grande para η implica na suavização excessiva, descaracterizando a função objetivo aproximada, que fica distante da função objetivo original. Por este motivo, a escolha deste parâmetro e sua atualização deve ser feita de maneira cuidadosa.

Segundo Souza (2010), a função ϑ também possui uma característica convexificadora, que pode eliminar alguns mínimos locais que estejam mais distantes do mínimo global, de acordo com o valor de η . Por este motivo, a escolha do parâmetro de suavização deve expressar um compromisso harmônico entre a simplificação do problema e a coerência da solução. A título de exemplificação, a Figura 5.2 apresenta o gráfico da função

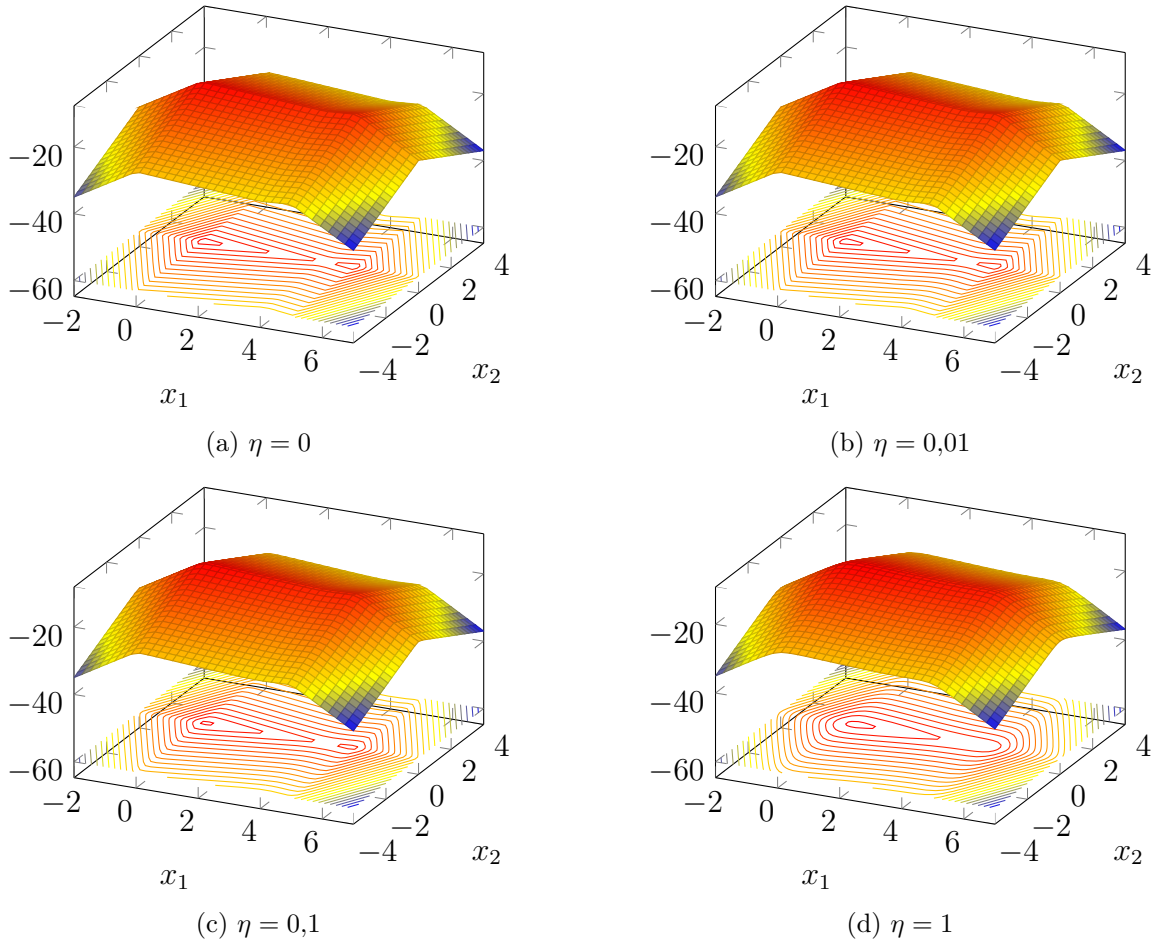


Figura 5.2: Gráfico da função $f(x_1, x_2) = -|x_1| - |3x_1 - 15| + \frac{2}{3}|x_1 - 4| - |x_2|$ com suas curvas de nível, em suas versões suavizadas, para diferentes valores do parâmetro η .

$f(x_1, x_2) = -|x_1| - |3x_1 - 15| + \frac{2}{3}|x_1 - 4| - |x_2|$ e suas versões obtidas através de suavização hiperbólica, para diferentes valores do parâmetro η .

Por fim, apresentamos o enunciado do resultado apresentado por Santos (1997) apud Junior (2010), que justifica do ponto de vista teórico a utilização da suavização e sua convergência para um ponto de mínimo local do problema original: *Para qualquer ponto de mínimo local do problema original e para qualquer vizinhança $\delta > 0$ especificada, existe um valor limitante $\bar{\tau}$ tal que, para todo $\tau_l < \bar{\tau}$, existe um ponto de mínimo local do problema suavizado $\bar{\mathbf{x}}_l$ situado dentro dessa vizinhança δ desse ponto de mínimo local $\bar{\mathbf{x}}$, ou seja, $\|\bar{\mathbf{x}}_l - \bar{\mathbf{x}}\| < \delta$, para todo $\tau_l < \bar{\tau}$.*

Capítulo 6

Resultados Numéricos

Neste capítulo são apresentados os resultados numéricos da aplicação do método PDPC-BLM ao PDE-PV nas três variantes apresentadas no Capítulo 3. O método foi programado em linguagem Matlab, utilizando o Matlab 2011a. A execução do programa foi realizada em um computador com processador Intel Core i5, com 4GB de memória RAM e sistema operacional Windows 7 Home Premium. Como a linguagem de programação é interpretada, não foram realizadas considerações acerca de tempo computacional.

O algoritmo proposto é uma extensão daquele proposto por Pinheiro (2012), com algumas variações: o procedimento de correção de inércia, que se baseia na proposta de Nocedal & Wright (2006), na forma de fatoração da matriz, através da fatoração LDL^t , o critério de escolha entre passo previsor ou corretor, e no cálculo do comprimento do passo, que utiliza fatores menores que 1, ao invés de 1,01, e calcula o passo pela fronteira da região relaxada na iteração seguinte. Inclui-se, também, devido às características do PDE-PV, o procedimento de suavização da função objetivo, através da técnica de suavização hiperbólica.

A fim de evitar problemas de mau condicionamento, quando o parâmetro de suavização é muito baixo, adotou-se a seguinte estratégia: fixou-se um valor $\bar{\eta} > 0$, de modo que a partir do momento em que $\eta^k \leq \bar{\eta}$, o parâmetro não é mais reduzido. Notou-se que a escolha do valor $\bar{\eta}$ tem um papel fundamental no refinamento da solução obtida. Os resultados obtidos foram comparados com outros trabalhos disponíveis na literatura, em particular Samed (2004) e Ravi et al. (2006).

Além disso, nas tabelas que exprimem os resultados das iterações, a coluna Δf_{abs}^k apresenta a diferença absoluta entre a função objetivo original e a função aproximada:

$$\Delta f_{\text{abs}}^k = |f(\mathbf{x}^k) - f_{\eta^k}(\mathbf{x}^k)|. \quad (6.1)$$

Já a coluna Δf_{rel}^k apresenta a diferença relativa percentual entre a função objetivo original e a função aproximada, sendo calculada da seguinte forma:

$$\Delta f_{\text{rel}}^k = \left| \frac{f(\mathbf{x}^k) - f_{\eta^k}(\mathbf{x}^k)}{f(\mathbf{x}^k)} \right| \cdot 100. \quad (6.2)$$

A coluna “corretor” indica se a direção do corretor foi utilizada na iteração: um sinal “−” indica que o passo e direção do previsor foram utilizados, enquanto um sinal “+” indica a utilização do passo e direção do corretor.

6.1 PDE-PV

Nesta seção apresentam-se os resultados obtidos na resolução do PDE-PV. Foram explorados três sistemas distintos, com 3, 13 e 19 geradores. Verificou-se que a convergência do método para cada caso está associada à escolha dos parâmetros de inicialização.

6.1.1 Caso 3 Geradores – Parte I

Os dados deste caso foram extraídos de Sinha et al. (2003), sendo que os mesmos também foram utilizados no trabalho de Samed (2004). Os coeficientes da curva de custo, incluindo os efeitos dos pontos de carregamento de carregamento de válvula, bem como os limites operacionais das unidades geradoras são apresentados na Tabela 6.1. A formulação do modelo do PDE-PV para este caso possui:

- 3 variáveis;
- 1 restrição de igualdade;
- 6 restrições de desigualdade.

O modelo é expresso por:

$$\begin{aligned} \text{minimizar} \quad & 0,001562P_{G,1}^2 + 7,92P_{G,1} + 561 + |300 \text{sen}(0,0315(100 - P_{G,1}))| \\ & + 0,00482P_{G,2}^2 + 7,97P_{G,2} + 78 + |150 \text{sen}(0,063(50 - P_{G,2}))| \\ & + 0,00194P_{G,3}^2 + 7,85P_{G,3} + 310 + |200 \text{sen}(0,042(100 - P_{G,3}))| \\ \text{sujeito a} \quad & P_{G,1} + P_{G,2} + P_{G,3} = 850 \\ & 100 \leq P_{G,1} \leq 600 \\ & 50 \leq P_{G,2} \leq 200 \\ & 100 \leq P_{G,3} \leq 400 \end{aligned} \quad (6.3)$$

Tabela 6.1: Características do sistema – PDE-PV: 3 geradores

Gerador (k)	$\overline{P_{G,k}}$ (MW)	$\overline{P_{G,k}}$ (MW)	a_k (\$/MW ²)	b_k (\$/MW)	c_k (\$)	e_k (\$)	f_k (1/MW)
1	100	600	0,001562	7,92	561	300	0,0315
2	50	200	0,00482	7,97	78	150	0,063
3	100	400	0,00194	7,85	310	200	0,042
Demanda: $D = 850$ MW							
Fonte: Sinha et al. (2003)							

Tabela 6.2: Inicialização – PDE-PV: 3 geradores (Parte I)

Fatores de aumento/redução da correção de inércia	$\kappa_1 = \kappa_2 = 10$	
Fatores de redução do parâmetro de suavização	$\rho_1 = 5$ e $\rho_2 = 0$	
Limitante da atualização do parâmetro de suavização	$\bar{\eta} = 10^{-4}$	
Parâmetro de barreira/relaxação inicial	$\mu^0 = 20$	
Fatores de atualização do parâmetro de barreira	$\tau_1 = \tau_2 = 0,2$	
Fator de controle da escolha entre passo previsor ou corretor	$\chi = 1$	
Fator de redução do passo	$\sigma = 0,9995$	
Precisão de parada	$\varepsilon = 10^{-6}$	
Potência inicial das unidades geradoras (MW)	1	220
	2	110
	3	280
Estimadores dos multiplicadores de Lagrange	$\delta^0 = \mathbb{1}_p$	

Os parâmetros apresentados na Tabela 6.2 foram utilizados na inicialização do método.

O vetor $\mathbb{1}_p$ denota um vetor coluna com p linhas, cujas componentes são todas iguais 1, em que p é o número de restrições de desigualdade. Ao aplicar o método PDPCBLM para a resolução do problema com os parâmetros especificados, verificou-se que o mesmo convergiu em 13 iterações, conforme a Tabela 6.3. Na referida tabela, o valor $f(\mathbf{x}^k)$ denota o valor da função objetivo original, enquanto $f_{\eta^k}(\mathbf{x}^k)$ denota o valor da função suavizada, pelo parâmetro η^k indicado na respectiva coluna. Da mesma tabela, infere-se que, na solução ótima, a diferença percentual relativa entre a função objetivo original (não-diferenciável) e a função objetivo aproximada (suave) é de apenas 0,012199927%.

Na Tabela (6.3), a norma dos resíduos, na solução ótima, é igual a $3,8071715 \times 10^{-7}$. Nota-se, também, que na 8ª iteração ocorreu penalização do parâmetro de barreira. A maioria das iterações deram passo máximo, excetuando as iterações 5, 6 e 7. Outro ponto que merece destaque é o fato de que houve alternância entre a escolha das direções do previsor e do corretor. O método utiliza, inicialmente, as direções do procedimento previsor nas iterações 1 e 2. Nas duas iterações seguintes, utiliza-se a direção do corretor. Na 5ª iteração, o método volta a utilizar as direções geradas no procedimento previsor

Tabela 6.3: Resultados (Iterações) – PDE-PV: 3 geradores (Parte I)

Iteração k	$f(\mathbf{x}^k)$	$f_{\eta^k}(\mathbf{x}^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	6511,7011	6852,8035	341,1024	5,2382993	240	1	20	0	0	–
1	8592,5064	8996,2693	403,76291	4,6990121	82,661543	1	20	0,9995	0,9995	–
2	8580,249	8708,5463	128,29736	1,4952638	3,6635996	0,2	4	0,9995	0,9995	–
3	8536,2999	8579,9362	43,636269	0,51118481	5,3458391	0,04	0,8	0,9995	0,9995	+
4	8419,7459	8447,4791	27,733233	0,32938326	8,6503266	0,008	0,16	0,9995	0,9995	+
5	8256,0524	8270,3375	14,285052	0,17302521	9,0898045	0,0016	0,032	0,9995	0,7881581	–
6	8234,4123	8239,7617	5,349344	0,064963276	3,247482	0,00032	0,0064	0,10717635	0,71692281	–
7	8234,0095	8235,3314	1,3219199	0,016054389	1,9926299	$6,4 \times 10^{-5}$	0,00128	0,9995	0,99008296	–
8	8234,197	8235,2053	1,0083157	0,012245464	0,5589004	$6,4 \times 10^{-5}$	0,32672583	0,9995	0,9995	–
9	8234,2281	8235,2075	0,97945053	0,011894868	0,0089980771	$6,4 \times 10^{-5}$	0,065345166	0,9995	0,9995	–
10	8234,2114	8235,2055	0,99414378	0,012073333	0,0021973929	$6,4 \times 10^{-5}$	0,013069033	0,9995	0,9995	+
11	8234,2032	8235,2052	1,0019436	0,01216807	0,00050966613	$6,4 \times 10^{-5}$	0,0026138066	0,9995	0,9995	+
12	8234,2008	8235,2051	1,0043567	0,01219738	$4,8015957 \times 10^{-5}$	$6,4 \times 10^{-5}$	0,00052276132	0,9995	0,9995	+
13	8234,2006	8235,2051	1,0045664	0,012199927	$3,8071715 \times 10^{-7}$	$6,4 \times 10^{-5}$	0,00010455226	0,9995	0,9995	+

Tabela 6.4: Resultados – PDE-PV: 3 geradores (Parte I)

Método	$P_{G,1}$ (MW)	$P_{G,2}$ (MW)	$P_{G,3}$ (MW)	Função Objetivo (\$)
AGH (Samed, 2004)	299,7595	150,4285	399,8120	8234,556
AGHCOE (Samed, 2004)	249,2786	101,4205	499,3009*	8233,979*
PDPCBLM	300,06712	149,94901	399,98387	8234,200

* A solução da autora apresenta valor melhor da função objetivo, mas é infactível.

para, finalmente, na 10^a iteração, voltar a utilizar a direção do corretor, até atingir a convergência.

Na Figura 6.1, pode-se ver a trajetória de soluções geradas pelo método proposto, na resolução desta instância do PDE-PV. Já na Figura 6.2 pode-se verificar a trajetória dos multiplicadores de Lagrange, da restrição de igualdade, relativa à satisfação da demanda, e das restrições de desigualdade, relacionadas à canalização das variáveis. Por fim, a Figura 6.3 apresenta o comportamento do parâmetro de correção de inércia ao longo das iterações do método. Nota-se que no início do método correções mais significativas são necessárias e que, à medida em que o método se aproxima da solução ótima, o parâmetro de correção é reduzido. Isto ocorre porque na região de mínimo local a função é convexa.

A escolha do ponto inicial $\mathbf{P}^0 = (220, 110, 280)^t$ foi realizada de modo que o mesmo pertencesse ao interior da região definida pelas restrições de desigualdade originais. Isto significa que, neste caso, a relaxação representada pela função barreira logarítmica modificada não exerce nenhuma influência quanto à factibilização do ponto inicial. A Tabela 6.4 apresenta o despacho ótimo obtido pelo método PDPCBLM, junto aos resultados apresentados em Samed (2004). Nesta tabela, pode-se notar que o método AGHCOE da autora obteve um valor da função objetivo inferior ao do método PDPCBLM. Porém, o limite da unidade geradora 3 foi violado, conforme destacado na Tabela 6.4. O método PDPCBLM determinou uma solução com o valor da função objetivo igual a \$8234,20. Trabalhos mais recentes, como o de Meng et al. (2010) indicam que o melhor valor mínimo obtido para a função objetivo é de \$8234,07, que é ligeiramente inferior ao obtido pelo método PDPCBLM. De fato, isto ocorreu devido ao valor do limitante $\bar{\eta}$, conforme veremos na próxima seção.

6.1.2 Caso 3 Geradores – Parte II

O problema do caso anterior foi resolvido novamente, utilizando um limitante $\bar{\eta}$ menor. Isto permite que a função aproximante apresente valores mais próximos ao da função original. Em contrapartida, sua resolução se torna mais difícil, em virtude da proximidade da natureza não-diferenciável do problema original, quando o valor de η é baixo.

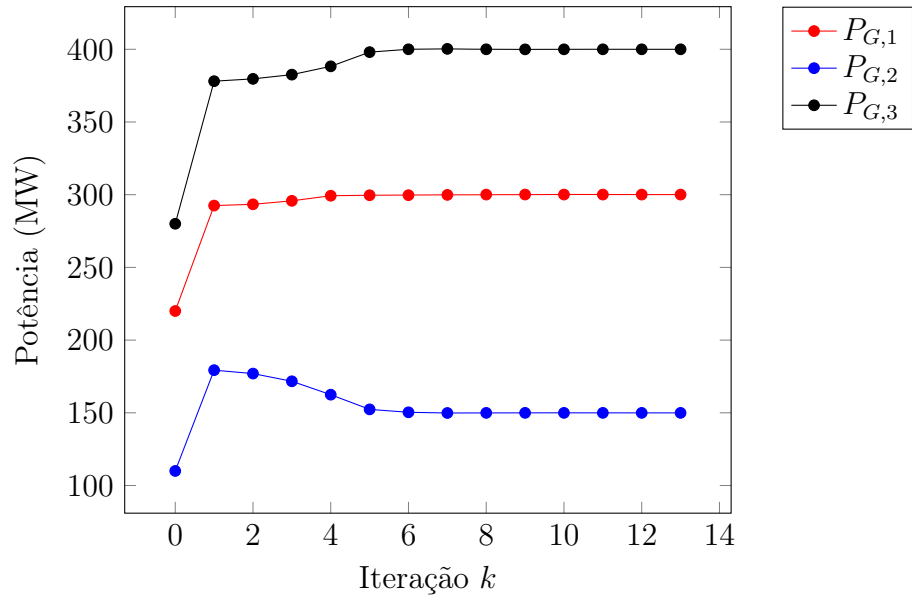


Figura 6.1: Sequência de potências geradas pelo método PDPCBLM para o PDE-PV de 3 geradores (Parte I).

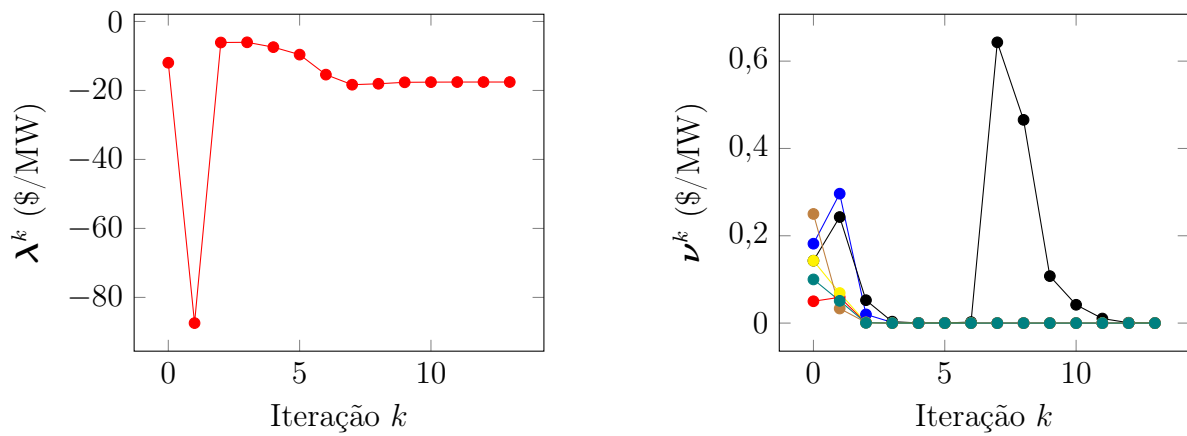


Figura 6.2: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PV de 3 geradores (Parte I).

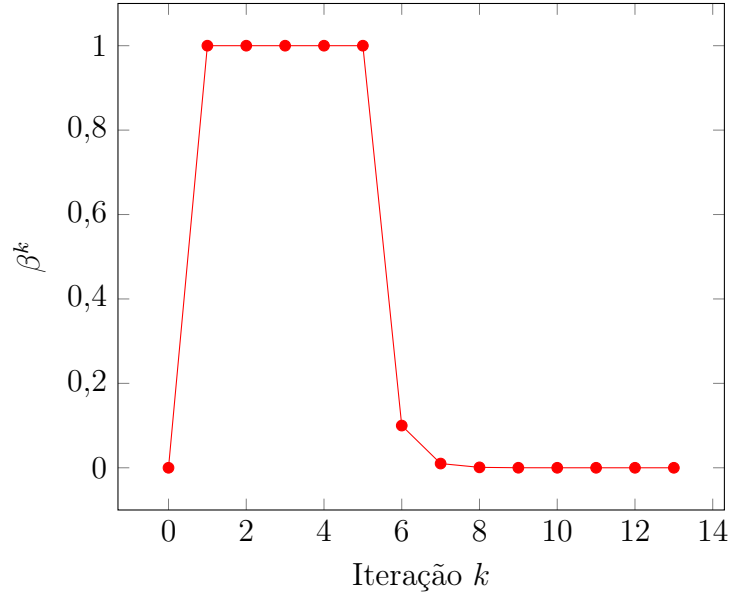


Figura 6.3: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PV de 3 geradores (Parte I).

Tabela 6.5: Inicialização – PDE-PV: 3 geradores (Parte II)

Fatores de aumento/redução da correção de inércia	$\kappa_1 = \kappa_2 = 10$	
Fatores de redução do parâmetro de suavização	$\rho_1 = 2,5$ e $\rho_2 = 5$	
Limitante da atualização do parâmetro de suavização	$\bar{\eta} = 10^{-8}$	
Parâmetro de barreira/relaxação inicial	$\mu^0 = 10$	
Fatores de atualização do parâmetro de barreira	$\tau_1 = \tau_2 = 0,5$	
Fator de controle da escolha entre passo previsor ou corretor	$\chi = 1$	
Fator de redução do passo	$\sigma = 0,9995$	
Precisão de parada	$\varepsilon = 10^{-6}$	
Potência inicial das unidades geradoras (MW)	1	220
	2	110
	3	280
Estimadores dos multiplicadores de Lagrange	$\delta^0 = \mathbb{1}_p$	

Os parâmetros apresentados na Tabela 6.5 foram utilizados na inicialização do método.

Assim, as únicas mudanças realizadas além do limitante $\bar{\eta}$ foram nos fatores de redução do parâmetro η , no parâmetro de barreira inicial, μ^0 , e seus fatores de atualização, τ_1 e τ_2 . Como já mencionado, a redução do limitante $\bar{\eta}$ permite a obtenção de soluções mais refinadas, porém ao custo de um número maior de iterações. Neste caso, verificou-se que o método convergiu em 21 iterações, cujos dados estão sumarizados na Tabela 6.6. Pode-se notar a ocorrência de oscilações na norma dos resíduos. Isto ocorre porque o parâmetro η já é bastante baixo e este efeito oscilatório somente acaba a partir do momento em que o limitante é atingido. Uma busca unidimensional mais robusta, como a de Armijo, ou um critério para aceitação de passo baseado em alguma função mérito ou filtro poderia ajudar na amenização deste efeito. Assim como no primeiro caso, ocorreram penalizações

no parâmetro de barreira em algumas iterações, a saber as iterações 10, 11, 12, 13 e 14. A diferença relativa entre a função objetivo original e aproximada, na solução ótima determinada, é de $2,9103339 \times 10^{-5}\%$.

A restrição de demanda deste problema é expressa por $P_{G,1} + P_{G,2} + P_{G,3} = 850$, a qual permite isolar, por exemplo, a variável $P_{G,3}$. Com isto, pode-se reduzir o problema de otimização à forma:

$$\begin{aligned}
\text{Minimizar} \quad & 0,001562P_{G,1}^2 + 7,92P_{G,1} + 561 + |300 \sin(0,0315(100 - P_{G,1}))| \\
& + 0,00482P_{G,2}^2 + 7,97P_{G,2} + 78 + |150 \sin(0,063(50 - P_{G,2}))| \\
& + 0,00194(850 - P_{G,1} - P_{G,2})^2 + 7,85(850 - P_{G,1} - P_{G,2}) + 310 \\
& + |200 \sin(0,042(100 - (850 - P_{G,1} - P_{G,2})))| \tag{6.4} \\
\text{sujeito a} \quad & 100 \leq P_{G,1} \leq 600 \\
& 50 \leq P_{G,2} \leq 200 \\
& 100 \leq 850 - P_{G,1} - P_{G,2} \leq 400
\end{aligned}$$

o qual é um problema de otimização apenas nas variáveis $P_{G,1}$ e $P_{G,2}$. A trajetória de soluções $P_{G,1}$ e $P_{G,2}$ obtidas na resolução do problema equivalente (6.4) e suas curvas de nível são apresentadas na Figura 6.4. Nota-se que, para estes parâmetros, o método foi capaz de refinar a solução até atingir o melhor valor conhecido na literatura para a solução ótima, ou seja, \$8234,07. Da mesma forma que no caso da Parte I, as Figuras 6.5, 6.6 e 6.7 apresentam a trajetória de soluções geradas pelo método, a trajetória dos multiplicadores de Lagrange e o comportamento do parâmetro de correção de inércia, respectivamente, ao longo das iterações do método. A Tabela 6.7 apresenta as soluções ótimas refinadas obtidas pelo método PDPCBLM. O valor máximo necessário do parâmetro de correção de inércia foi igual 1, assim como na Parte I.

6.1.3 Caso 3 Geradores – Parte III

Mais uma vez, o PDE-PV de 3 geradores foi resolvido, agora visando investigar a capacidade da função barreira modificada tratar a infactibilidade, quando o ponto inicial viola alguma restrição de desigualdade. Tendo em vista que o objetivo é analisar a capacidade de convergência em uma situação na qual a relaxação promovida pela função barreira modificada é necessária, e não a capacidade de refinar a solução, adotou-se o mesmo conjunto de parâmetros da Tabela 6.2, excetuando o ponto inicial, que foi alterado para:

$$\mathbf{P}^0 = \begin{pmatrix} P_{G,1}^0 & P_{G,2}^0 & P_{G,3}^0 \end{pmatrix}^t = \begin{pmatrix} 230 & 40 & 280 \end{pmatrix}^t.$$

Tabela 6.6: Resultados (Iterações) – PDE-PV: 3 geradores (Parte II)

Iteração k	$f(\mathbf{x}^k)$	$f_{\eta}(\mathbf{x}^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	6511,7011	6852,8035	341,1024	5,2382993	240	1	10	0	0	–
1	8592,8828	8996,277	403,39413	4,6945145	83,311444	1	10	0,9995	0,9995	–
2	8577,6469	8672,3574	94,71054	1,1041553	3,9601397	0,13333333	5	0,9995	0,9995	–
3	8526,8127	8561,0041	34,191337	0,40098613	5,8299234	0,026666667	2,5	0,9995	0,9995	+
4	8406,7587	8431,9206	25,161881	0,29930538	8,7030731	0,0064	1,25	0,9995	0,9995	+
5	8240,5679	8258,5075	17,939585	0,2176984	8,7663285	0,0017066667	0,625	0,9995	0,80583677	–
6	8235,2777	8241,7831	6,5053616	0,078993834	11,694752	0,00048761905	0,3125	0,9995	0,97055432	–
7	8234,2471	8236,3729	2,1258223	0,025816839	0,23152415	0,00014628571	0,15625	0,9995	0,32059339	–
8	8233,7888	8234,918	0,75616954	0,0091833212	0,73675708	4,5511111 $\times 10^{-5}$	0,078125	0,7621882	0,99903682	–
9	8233,7888	8234,4998	0,71095287	0,0086345775	1,4754229	1,4563556 $\times 10^{-5}$	0,0390625	0,9995	0,99710159	–
10	8233,8204	8234,0369	0,21645619	0,002628867	0,48323153	4,7662545 $\times 10^{-6}$	0,89018993	0,9995	0,9995	+
11	8233,7189	8233,9156	0,19670097	0,0023889687	0,50772896	1,5887515 $\times 10^{-6}$	0,82572648	0,9995	0,9995	–
12	8233,7166	8233,7846	0,067925789	0,00082497118	0,48363361	5,3773128 $\times 10^{-7}$	1,0834109	0,9995	0,9995	+
13	8233,7181	8233,743	0,024848744	0,00030179251	0,37984229	1,8436501 $\times 10^{-7}$	1,0734311	0,9995	0,9995	+
14	8233,7887	8233,7963	0,0076306136	9,2674392 $\times 10^{-5}$	0,44762133	6,3913204 $\times 10^{-8}$	1,0602661	0,9995	0,9995	+
15	8234,1142	8234,1309	0,016686035	0,00020264518	13,30527	2,2369621 $\times 10^{-8}$	0,8495685	0,9995	0,9995	–
16	8234,0766	8234,0807	0,0040996352	4,9788645 $\times 10^{-5}$	1,0241975	7,8951605 $\times 10^{-9}$	0,42478425	0,9995	0,9995	+
17	8234,0734	8234,0764	0,0030311936	3,6812807 $\times 10^{-5}$	0,33022371	7,8951605 $\times 10^{-9}$	0,21239213	0,9995	0,9995	–
18	8234,0738	8234,0763	0,0025452795	3,0911546 $\times 10^{-5}$	0,071344325	7,8951605 $\times 10^{-9}$	0,10619606	0,9995	0,9995	+
19	8234,0739	8234,0763	0,002408155	2,9246216 $\times 10^{-5}$	0,0054960911	7,8951605 $\times 10^{-9}$	0,053098031	0,9995	0,9995	+
20	8234,0739	8234,0763	0,0023964792	2,9104417 $\times 10^{-5}$	4,1407242 $\times 10^{-5}$	7,8951605 $\times 10^{-9}$	0,026549016	0,9995	0,9995	+
21	8234,0739	8234,0763	0,0023963904	2,9103339 $\times 10^{-5}$	2,2924681 $\times 10^{-8}$	7,8951605 $\times 10^{-9}$	0,013274508	0,9995	0,9995	+

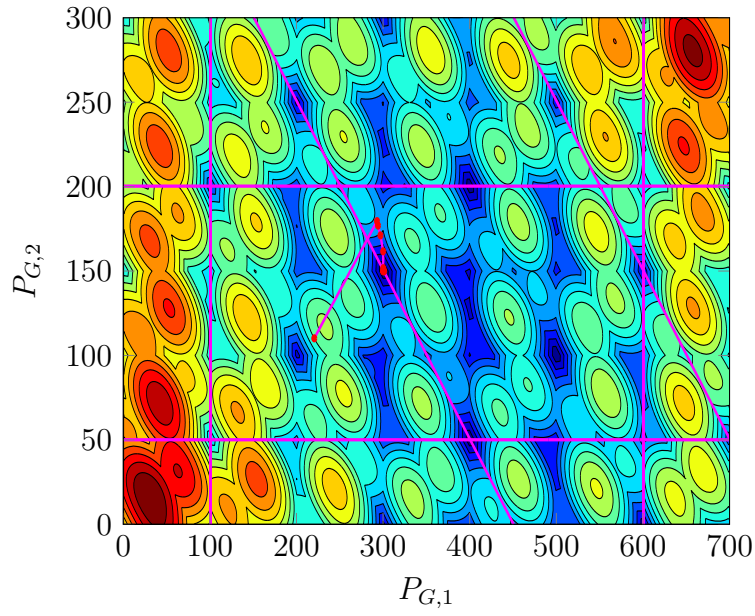


Figura 6.4: Trajetória dos pontos $(P_{G,1}, P_{G,2})$ – PDE-PV: 3 geradores (Parte II).

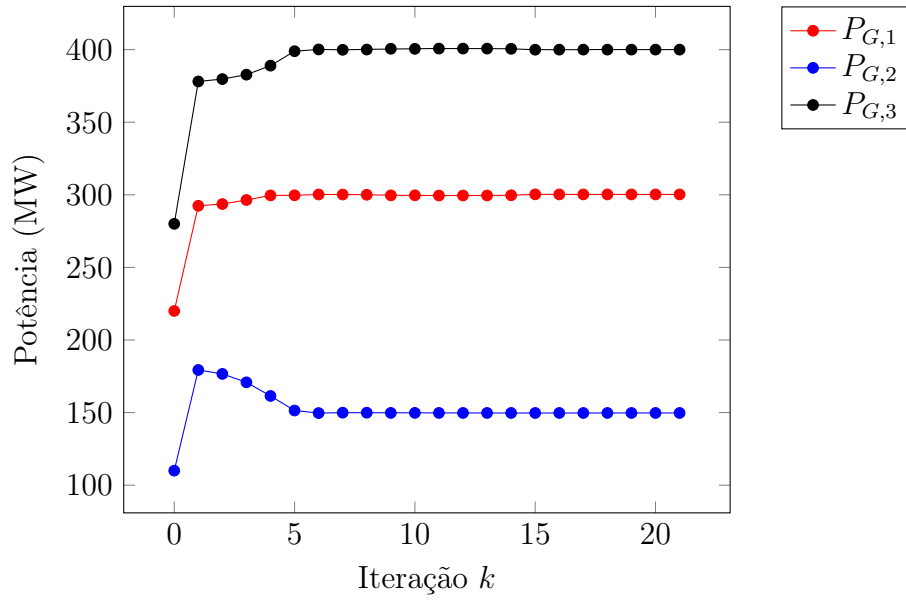


Figura 6.5: Sequência de potências geradas pelo método PDPCBLM para o PDE-PV de 3 geradores (Parte II).

Tabela 6.7: Resultados – PDE-PV: 3 geradores (Parte II)

Método	$P_{G,1}$ (MW)	$P_{G,2}$ (MW)	$P_{G,3}$ (MW)	Função Objetivo (\$)
AGH (Samed, 2004)	299,7595	150,4285	399,8120	8234,556
AGHCOE (Samed, 2004)	249,2786	101,4205	499,3009*	8233,979*
QPSO (Meng et al., 2010)	300,27	149,73	400	8234,07
PDPCBLM	300,26298	149,73702	400	8234,07391

* A solução da autora apresenta valor melhor da função objetivo, mas é infactível.

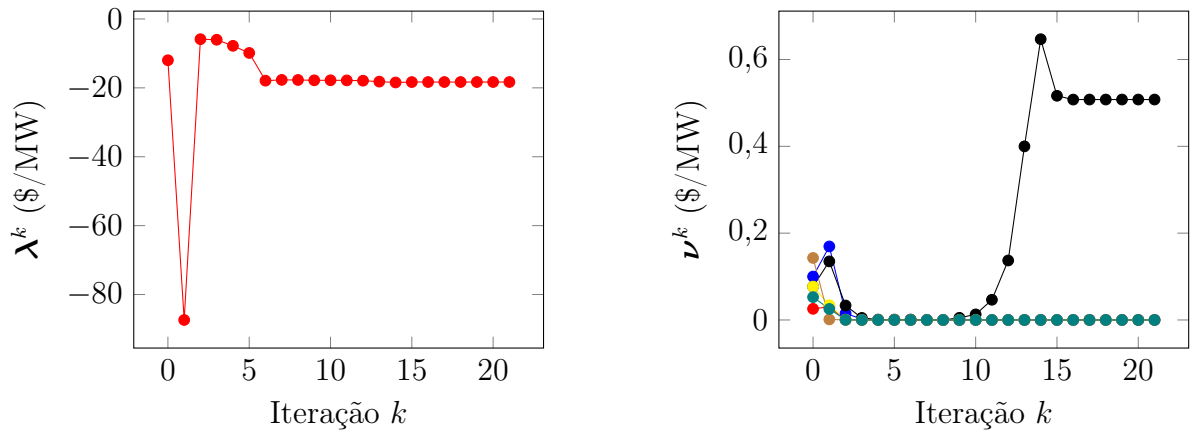


Figura 6.6: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PV de 3 geradores (Parte II).

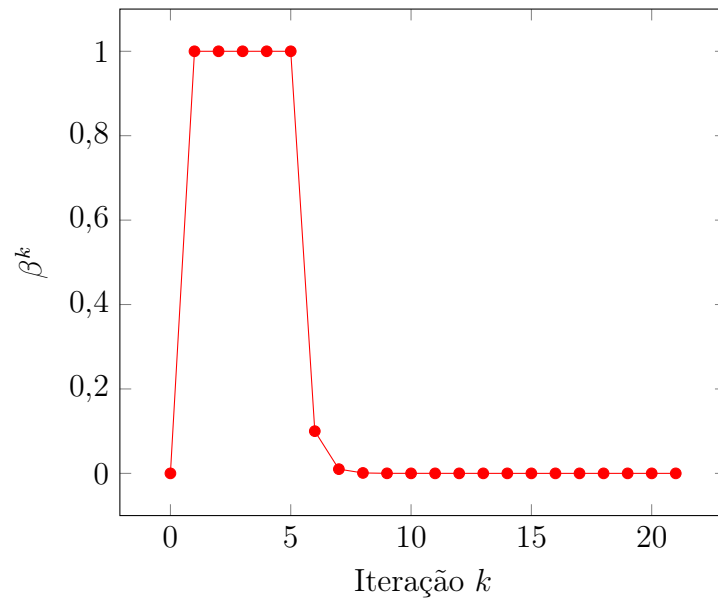


Figura 6.7: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PV de 3 geradores (Parte II).

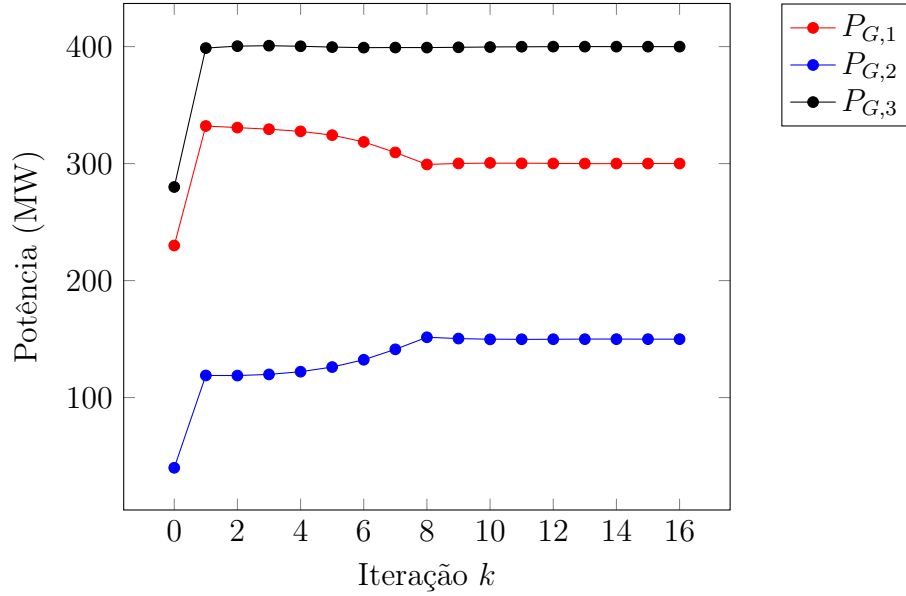


Figura 6.8: Sequência de potências geradas pelo método PDPCBLM para o PDE-PV de 3 geradores (Parte III).

Observa-se que, neste ponto inicial, a potência inicial $P_{G,2}^0 = 40$ viola o limite mínimo de geração da unidade geradora 2. O método convergiu em 16 iterações, e os resultados das iterações estão resumidos na Tabela 6.8. A diferença relativa entre a função objetivo original e aproximada, na solução ótima determinada, é de 0,012199947%. A análise da trajetória das potências geradas, apresentada na Figura 6.8 indica que logo no início do processo iterativo o método consegue trazer para o interior da região o ponto infactível para as restrições de desigualdade. Já as Figuras 6.9 e 6.10 apresentam a trajetória dos multiplicadores de Lagrange e do parâmetro de correção de inércia, respectivamente, ao longo das iterações.

Um ponto interessante a se notar, em todos os casos de 3 geradores, e que fica mais evidente na Tabela 6.8 de resultados da Parte III é que, usualmente, o passo dual não é máximo quando a direção do procedimento predictor é utilizada, enquanto o oposto é mais comum nas iterações em que a direção do procedimento corretor é utilizada.

6.1.4 Caso 13 Geradores – Parte I

Os dados deste caso foram extraídos de Samed (2004). Os coeficientes da curva de custo, incluindo os efeitos dos pontos de carregamento de válvula, bem como os limites operacionais das unidades geradoras são apresentados na Tabela 6.10. A formulação do modelo do PDE-PV para este caso possui:

- 13 variáveis;

Tabela 6.8: Resultados (Iterações) – PDE-PV: 3 geradores (Parte III)

Iteração k	$f(\mathbf{x}^k)$	$f_{\eta^k}(\mathbf{x}^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	6054,2469	6367,8651	313,61821	5,1801359	1246,7794	1	20	0	0	–
1	8607,4104	9007,0697	399,65928	4,6432	28,450851	1	20	0,9995	0,12659104	–
2	8608,6855	8737,5512	128,86565	1,496926	1,8591072	0,2	4	0,9995	0,9995	–
3	8608,0211	8647,2608	39,239749	0,45585098	1,9587152	0,04	0,8	0,9995	0,99685064	–
4	8598,7856	8611,8607	13,075052	0,15205696	2,8093315	0,008	0,94853387	0,9995	0,9995	+
5	8574,4857	8580,2977	5,8119933	0,067782413	4,5507407	0,0016	0,35200027	0,9995	0,14380445	–
6	8515,1452	8518,3103	3,165024	0,037169348	6,9794011	0,00032	0,070400053	0,9995	0,5497217	–
7	8385,2545	8386,8822	1,6276959	0,019411407	8,6524974	$6,4 \times 10^{-5}$	0,014080011	0,9995	0,9995	+
8	8239,634	8242,2742	2,6401604	0,032042204	10,307366	$6,4 \times 10^{-5}$	0,0028160021	0,9995	0,9995	+
9	8234,7666	8235,9177	1,1510217	0,013977587	2,8071709	$6,4 \times 10^{-5}$	0,00056320043	0,9995	0,97710027	–
10	8234,2825	8235,7645	1,4819935	0,017997846	6,1613342	$6,4 \times 10^{-5}$	0,00011264009	0,9995	0,9995	+
11	8234,1854	8235,2806	1,0951875	0,013300496	1,8020633	$6,4 \times 10^{-5}$	$2,2528017 \times 10^{-5}$	0,9995	0,9995	+
12	8234,1658	8235,2193	1,0535569	0,012794944	0,64540722	$6,4 \times 10^{-5}$	$4,5056034 \times 10^{-6}$	0,70580439	0,9989196	–
13	8234,19	8235,2057	1,0157345	0,012335573	0,14833618	$6,4 \times 10^{-5}$	$9,0112068 \times 10^{-7}$	0,9995	0,0063205433	–
14	8234,2001	8235,2051	1,0050454	0,012205744	0,03038242	$6,4 \times 10^{-5}$	$1,8022414 \times 10^{-7}$	0,9995	0,41521868	–
15	8234,2006	8235,2051	1,0045695	0,012199964	$2,8222596 \times 10^{-5}$	$6,4 \times 10^{-5}$	$3,6044827 \times 10^{-8}$	0,9995	0,9995	+
16	8234,2006	8235,2051	1,0045681	0,012199947	$1,4227901 \times 10^{-8}$	$6,4 \times 10^{-5}$	$7,2089654 \times 10^{-9}$	0,9995	0,9995	+

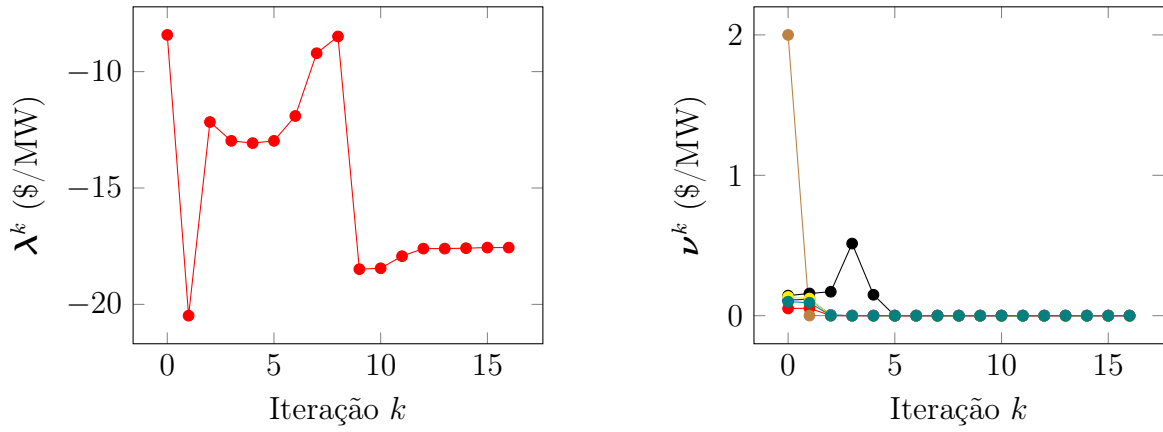


Figura 6.9: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PV de 3 geradores (Parte III).

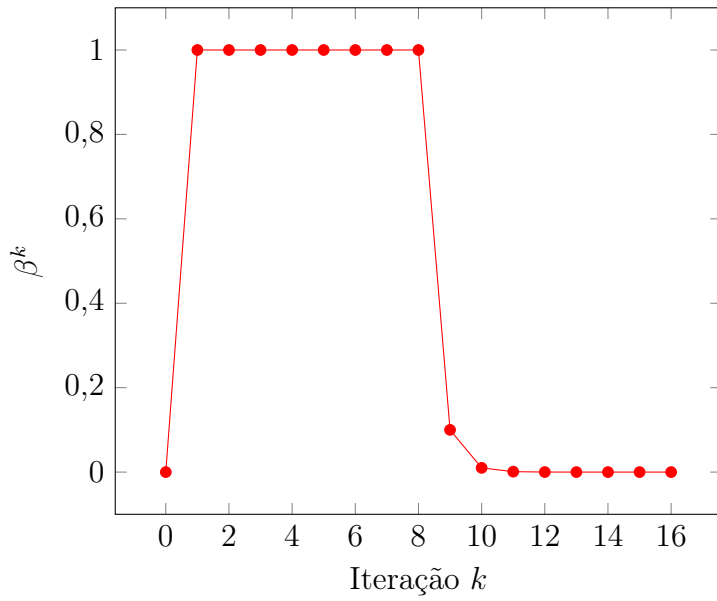


Figura 6.10: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PV de 3 geradores (Parte III).

Tabela 6.9: Resultados – PDE-PV: 3 geradores (Parte III)

Método	$P_{G,1}$ (MW)	$P_{G,2}$ (MW)	$P_{G,3}$ (MW)	Função Objetivo (\$)
AGH (Samed, 2004)	299,7595	150,4285	399,8120	8234,556
AGHCOE (Samed, 2004)	249,2786	101,4205	499,3009*	8233,979*
QPSO (Meng et al., 2010)	300,27	149,73	400	8234,07
PDPCBLM	300,06712	149,94901	399,98387	8234,20056

* A solução da autora apresenta valor melhor da função objetivo, mas é infactível.

Tabela 6.10: Características do sistema – PDE-PV: 13 geradores

Gerador (k)	$\overline{P_{G,k}}$ (MW)	$\overline{P_{G,k}}$ (MW)	a_k (\$/MW ²)	b_k (\$/MW)	c_k (\$)	e_k (\$)	f_k (1/MW)
1	0	680	0,00028	8,1	550	300	0,035
2	0	360	0,00056	8,1	309	200	0,042
3	0	360	0,00056	8,1	307	200	0,042
4	60	180	0,00324	7,74	240	150	0,063
5	60	180	0,00324	7,74	240	150	0,063
6	60	180	0,00324	7,74	240	150	0,063
7	60	180	0,00324	7,74	240	150	0,063
8	60	180	0,00324	7,74	240	150	0,063
9	60	180	0,00324	7,74	240	150	0,063
10	40	120	0,00284	8,6	126	100	0,084
11	40	120	0,00284	8,6	126	100	0,084
12	55	120	0,00284	8,6	126	100	0,084
13	55	120	0,00284	8,6	126	100	0,084

Demanda: $D = 2520$ MW

Fonte: Samed (2004)

- 1 restrição de igualdade;
- 26 restrições de desigualdade.

Os parâmetros apresentados na Tabela 6.11 foram utilizados na inicialização do método.

Neste caso, a escolha do ponto inicial foi feita de modo a ficar próxima de satisfazer a demanda total, de 2520 MW (a soma das potências iniciais é de 2500 MW) e de modo que geradores com limitantes superiores de geração iguais fossem inicializados com o mesmo valor de geração. A Tabela 6.12 apresenta os resultados obtidos, nas 18 iterações necessárias para a convergência. Nota-se que não houve penalização do parâmetro de barreira e , assim como no caso de 3 geradores, a redução significativa dos resíduos somente ocorreu a partir do momento em que o parâmetro η atingiu o limitante $\bar{\eta}$. A diferença relativa entre a função objetivo original e aproximada, na solução ótima determinada, é de 0,016131898%.

Neste caso, notou-se que o método foi bastante sensível à escolha dos parâmetros e , nos testes realizados, uma escolha inadequada dos fatores de atualização do parâmetro de suavização levavam à divergência. Por isso, escolheu-se um valor baixo para ρ_1 . Outra dificuldade é a quantidade de mínimos locais deste tipo de problema. Apenas para se ter uma ideia, Borckmans et al. (2014) verificou que após executar o seu algoritmo 10^6 vezes, o método determinou cerca de 2×10^5 candidatos a solução, em uma instância com apenas 15 geradores. Por isso, escolheu-se uma solução relativamente próxima à reportada no trabalho de Pothiya et al. (2010). Ainda assim, o método determinou um mínimo local próximo, diferente daquele que se esperava obter. As Figuras 6.11, 6.12 e 6.13 apresentam

Tabela 6.11: Inicialização – PDE-PV: 13 geradores (Parte I)

Fatores de aumento/redução da correção de inércia	$\kappa_1 = \kappa_2 = 10$	
Fatores de redução do parâmetro de suavização	$\rho_1 = 2$ e $\rho_2 = 0$	
Limitante da atualização do parâmetro de suavização	$\bar{\eta} = 10^{-4}$	
Parâmetro de barreira/relaxação inicial	$\mu^0 = 20$	
Fatores de atualização do parâmetro de barreira	$\tau_1 = \tau_2 = 0,5$	
Fator de controle da escolha entre passo previsor ou corretor	$\chi = 1$	
Fator de redução do passo	$\sigma = 0,9995$	
Precisão de parada	$\varepsilon = 10^{-6}$	
Potência inicial das unidades geradoras (MW)	1	600
	2-3	300
	4-9	150
	10-13	100
Estimadores dos multiplicadores de Lagrange	$\delta^0 = \mathbb{1}_p$	

a trajetória das potências geradas, dos multiplicadores de Lagrange e do parâmetro de correção de inércia, respectivamente. No caso do gráfico das potências geradas, verificou-se que aqueles que possuem características iguais apresentaram a mesma trajetória de otimização, de modo que algumas linhas estão sobrepostas.

Pela análise da Tabela 6.13, nota-se que a solução ótima obtida pelo método PDPCBLM melhora perceptivelmente as apresentadas no trabalho de Samed (2004). Por outro lado, apresenta valor próximo, mas superior, ao apresentado no trabalho de Pothiya et al. (2010), que utilizou a heurística de colônia de formigas para resolver o problema.

6.1.5 Caso 13 Geradores – Parte II

Da mesma forma que foi feita no caso de 3 geradores, investigamos o efeito da mudança dos parâmetros e a possibilidade de refinamento da solução obtida. A inicialização foi realizada considerando alguns aspectos do resultado obtido na Parte I. Os parâmetros apresentados na Tabela 6.14 foram utilizados na inicialização do método.

Neste caso, alterou-se os parâmetros relacionados à suavização e o parâmetro de barreira inicial, visando obter uma solução refinada. A Tabela 6.15 apresenta os resultados obtidos, nas 33 iterações necessárias para a convergência. Novamente, a redução significativa dos resíduos somente ocorreu a partir do momento em que o parâmetro η atingiu o limitante $\bar{\eta}$.

Como se pode notar nas Figuras 6.15 e 6.16, o método precisou de correções de inércia mais significativas no início do processo iterativo: entre as iterações 1 e 6. Este fato, provavelmente, também está relacionado ao comportamento dos multiplicadores de Lagrange no início do processo. Os multiplicadores aparentemente encontraram uma trajetória de

Tabela 6.12: Resultados (Iterações) – PDE-PV: 13 geradores (Parte I)

Iteração k	$f(\mathbf{x}^k)$	$f_{\eta^k}(\mathbf{x}^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	25079,52	26326,608	1247,088	4,9725353	20	1	20	0	0	–
1	24990,256	26378,724	1388,4678	5,5560365	19,252998	1	20	0,9995	0,9995	–
2	24839,408	25797,716	958,30766	3,8580133	4,121365	0,5	10	0,9995	0,9995	–
3	24823,161	25411,527	588,36603	2,3702301	4,0964471	0,25	5	0,9995	0,9995	+
4	24770,632	25137,938	367,30582	1,4828278	4,4637123	0,125	2,5	0,9995	0,9995	+
5	24581,419	24850,317	268,89872	1,0939105	7,9971766	0,0625	1,25	0,9995	0,9995	+
6	24177,073	24490,359	313,28587	1,2957973	2,3586428	0,03125	0,625	0,9995	0,72908907	–
7	24175,542	24386,697	211,15477	0,87342309	0,67929897	0,015625	0,3125	0,9995	0,99875182	–
8	24175,657	24314,973	139,31632	0,57626695	0,89705081	0,0078125	0,15625	0,9995	0,9995	+
9	24175,699	24265,235	89,536529	0,37035756	1,0485578	0,00390625	0,078125	0,9995	0,9995	+
10	24175,7	24231,328	55,628001	0,23009882	1,2313708	0,001953125	0,0390625	0,9995	0,9995	+
11	24175,649	24208,865	33,216485	0,13739646	1,3756707	0,0009765625	0,01953125	0,9995	0,9995	+
12	24175,512	24194,59	19,077417	0,078912153	1,3629412	0,00048828125	0,009765625	0,9995	0,9995	+
13	24175,223	24185,934	10,710993	0,044305664	1,2221221	0,00024414063	0,0048828125	0,9995	0,9995	+
14	24174,656	24180,83	6,174093	0,025539528	1,0338903	0,00012207031	0,0024414063	0,9995	0,9995	+
15	24173,691	24177,724	4,0329926	0,016683396	0,2840569	$6,1035156 \times 10^{-5}$	0,0012207031	0,9995	0,9995	+
16	24173,808	24177,715	3,9069244	0,016161808	0,011258933	$6,1035156 \times 10^{-5}$	0,00061035156	0,9995	0,9995	+
17	24173,816	24177,715	3,8996503	0,016131712	0,0003804442	$6,1035156 \times 10^{-5}$	0,00030517578	0,9995	0,9995	+
18	24173,816	24177,715	3,8996954	0,016131898	$5,7081625 \times 10^{-7}$	$6,1035156 \times 10^{-5}$	0,00015258789	0,9995	0,9995	+

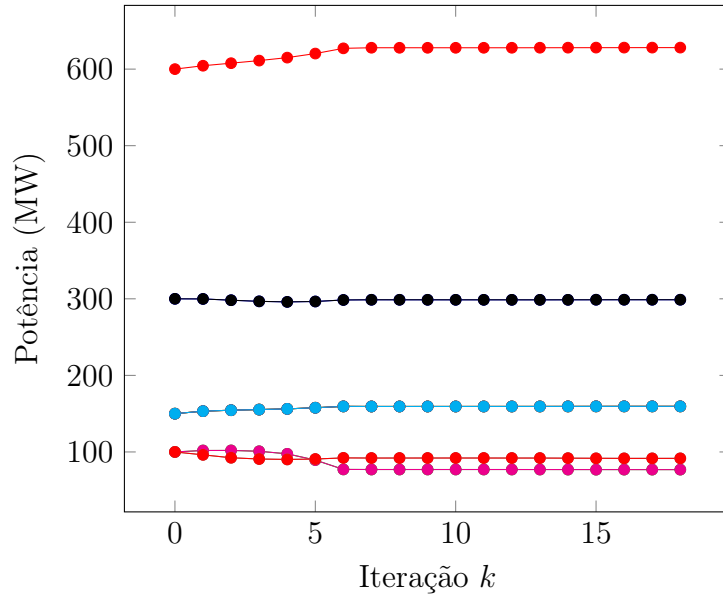


Figura 6.11: Sequência de potências geradas pelo método PDPCBLM para o PDE-PV de 13 geradores (Parte I).

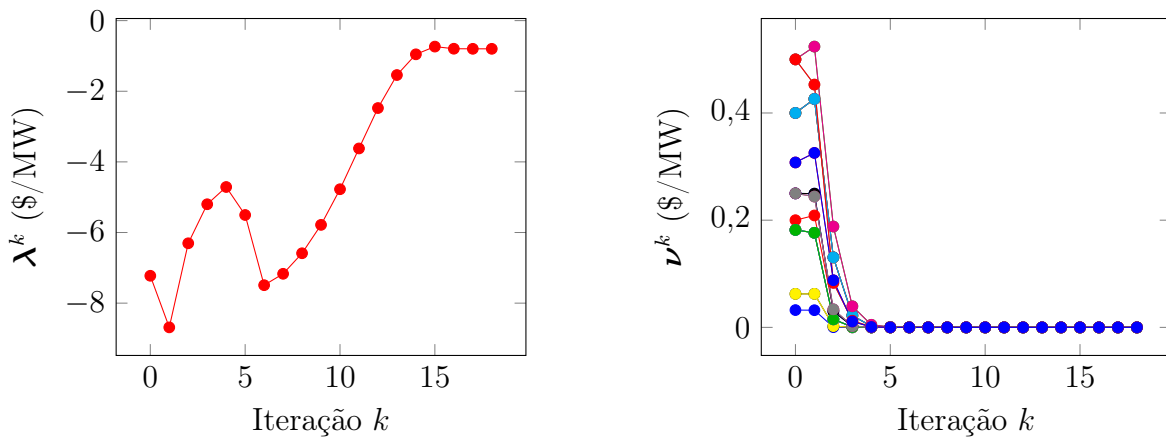


Figura 6.12: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PV de 13 geradores (Parte I).

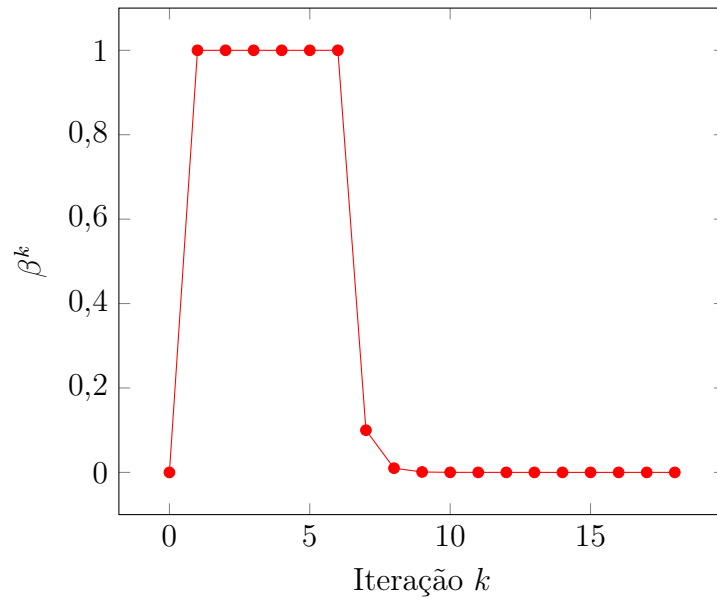


Figura 6.13: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PV de 13 geradores (Parte I).

Tabela 6.13: Resultados – PDE-PV: 13 geradores (Parte I)

Potência Gerada (MW)	Métodos			
	AGH (Samed, 2004)	AGHCOE (Samed, 2004)	ACO (Pothiya et al., 2010)	PDPCBLM
$P_{G,1}$	643,5755	617,1943	628,32	628,08088
$P_{G,2}$	328,6986	300,4246	299,06	298,79312
$P_{G,3}$	310,0123	297,2810	299,17	298,79312
$P_{G,4}$	124,9541	165,6300	159,73	159,53795
$P_{G,5}$	145,0035	159,0019	159,73	159,53795
$P_{G,6}$	150,8724	164,5568	159,73	159,53795
$P_{G,7}$	160,4046	140,7621	159,73	159,53795
$P_{G,8}$	172,1755	157,1320	159,72	159,53795
$P_{G,9}$	176,5106	165,9530	159,72	159,53795
$P_{G,10}$	62,44434	76,51676	75,47	76,92247
$P_{G,11}$	92,61867	74,81078	77,33	76,92247
$P_{G,12}$	62,40476	80,65879	92,10	91,63009
$P_{G,13}$	90,32517	120,0779	90,59	91,63009
Função Objetivo (\$)	25374,14	24799,03	24169,63	24173,81

Tabela 6.14: Inicialização – PDE-PV: 13 geradores (Parte II)

Fatores de aumento/redução da correção de inércia	$\kappa_1 = \kappa_2 = 10$	
Fatores de redução do parâmetro de suavização	$\rho_1 = 1,5$ e $\rho_2 = 2$	
Limitante da atualização do parâmetro de suavização	$\bar{\eta} = 10^{-7}$	
Parâmetro de barreira/relaxação inicial	$\mu^0 = 10$	
Fatores de atualização do parâmetro de barreira	$\tau_1 = \tau_2 = 0,5$	
Fator de controle da escolha entre passo previsor ou corretor	$\chi = 1$	
Fator de redução do passo	$\sigma = 0,9995$	
Precisão de parada	$\varepsilon = 10^{-6}$	
Potência inicial das unidades geradoras (MW)	1	600
	2-3	300
	4-9	150
	10-13	100
Estimadores dos multiplicadores de Lagrange	$\delta^0 = \mathbb{1}_p$	

convergência a partir da 10^a iteração, isto é, assim que as correções de inércia mais significativas já haviam ocorrido. Além disso, pode-se perceber que, no final do processo de convergência, o parâmetro de correção de inércia se elevou, um pouco, novamente. A justificativa mais coerente para este fenômeno é o fato de que, conforme o processo de convergência avança, o problema torna-se mais mal-condicionado, em especial nas regiões de não-diferenciabilidade em que, provavelmente, encontram-se os pontos de mínimo. Por fim, a Tabela 6.16 apresenta os valores das gerações de potência ativa determinadas pelo método PDPCBLM.

Apesar do aumento no número de iterações, em decorrência da redução do limitante $\bar{\eta}$, observa-se que, neste caso, mais uma vez foi possível melhorar as soluções apresentadas, inclusive em relação ao trabalho de Pothiya et al. (2010), o que não tinha acontecido na Parte I, em que a solução obtida era um pouco pior em termos de custo.

6.1.6 Caso 19 Geradores

Os dados deste caso foram extraídos de Balamurugan et al. (2014), que resolveu o problema de despacho econômico e ambiental combinados, através de um método de evolução diferencial. Os coeficientes da curva de custo, incluindo os efeitos dos pontos de carregamento de válvula, bem como os limites operacionais das unidades geradoras são apresentados na Tabela 6.17. A formulação do modelo do PDE-PV para este caso possui:

- 19 variáveis;
- 1 restrição de igualdade;
- 38 restrições de desigualdade.

Tabela 6.15: Resultados (Iterações) – PDE-PV: 13 geradores (Parte II)

Iteração k	$f(x^k)$	$f_{\eta^k}(x^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(x^k, s^k, \lambda^k, \nu^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	25112,236	26346,289	1234,0534	4,9141518	20	1	10	0	0	–
1	25030,071	26396,805	1366,03673	5,4603673	9,7688699	1	10	0,9995	0,9995	–
2	24786,569	25469,044	682,47467	2,7534051	4,9662059	0,28571429	5	0,9995	0,9995	+
3	24768,286	25119,066	350,78022	1,4162475	5,3286847	0,11428571	2,5	0,93838208	0,99777029	–
4	24714,133	24926,065	211,93232	0,85753491	4,8431636	0,052747253	1,25	0,9995	0,034036271	+
5	24544,487	24702,504	158,0168	0,64379751	8,0443604	0,026373626	12,294926	0,9995	0,9995	–
6	24167,405	24379,394	211,98872	0,87716791	8,7970656	0,013880856	6,1474632	0,9995	0,9995	+
7	24172,797	24313,789	140,99251	0,58326935	13,116587	0,007571376	3,0737316	0,46667348	0,86665489	–
8	24145,674	24264,243	118,5683	0,49105398	3,8659475	0,0042399706	1,5368658	0,9995	0,0024735065	–
9	24152,564	24231,807	79,242782	0,32809263	7,3283819	0,0024228403	0,76843289	0,71869204	0,9995	+
10	24145,887	24208,161	62,273242	0,25790414	4,6446985	0,0014068105	0,38421645	0,9995	0,0014010906	–
11	24147,661	24190,645	42,983893	0,17800437	5,308421	0,00082753559	0,19210822	0,7875063	0,9995	+
12	24145,982	24177,65	31,668181	0,131153	1,5344786	0,00049204819	0,096054112	0,9995	0,0023364031	–
13	24145,983	24168,334	22,351096	0,092566518	0,91376809	0,00029522891	0,048027056	0,9995	0,9995	+
14	24145,971	24161,511	15,539567	0,064356769	0,97528402	0,00017851051	0,024013528	0,9995	0,9995	–
15	24145,952	24156,576	10,624583	0,044001509	0,98653359	0,00010865857	0,012006764	0,9995	0,9995	–
16	24145,919	24153,072	7,1533119	0,029625346	0,94251505	6,6525654 $\times 10^{-5}$	0,006003382	0,9995	0,99945851	–
17	24145,865	24150,633	4,7679548	0,019746465	0,85750416	4,0938864 $\times 10^{-5}$	0,003001691	0,9995	0,99933996	–
18	24145,777	24148,961	3,1835178	0,013184574	0,75596812	2,5307662 $\times 10^{-5}$	0,0015008455	0,9995	0,9995	+
19	24145,638	24147,815	2,1772036	0,0090169645	0,65978429	1,5708204 $\times 10^{-5}$	0,00075042275	0,9995	0,9995	+
20	24145,441	24147,012	1,5716795	0,0065092184	0,58197104	9,7854384 $\times 10^{-6}$	0,00037521137	0,9995	0,9995	+
21	24145,229	24146,427	1,1976228	0,0049600804	0,53888976	6,115899 $\times 10^{-6}$	0,00018760569	0,9995	0,9995	+
22	24145,059	24145,985	0,92551739	0,0038331544	0,5293591	3,8338471 $\times 10^{-6}$	9,3802844 $\times 10^{-5}$	0,9995	0,9995	+
23	24144,92	24145,644	0,72419215	0,0029993562	0,52169114	2,4098468 $\times 10^{-6}$	4,6901422 $\times 10^{-5}$	0,9995	0,9995	+
24	24144,809	24145,378	0,56945414	0,0023584951	0,51787206	1,5185336 $\times 10^{-6}$	2,3450711 $\times 10^{-5}$	0,9995	0,9995	+
25	24144,72	24145,17	0,44996555	0,0018636188	0,51477716	9,5907384 $\times 10^{-7}$	1,1725355 $\times 10^{-5}$	0,9995	0,9995	+
26	24144,65	24145,007	0,35636146	0,0014759438	0,51273909	6,0700876 $\times 10^{-7}$	5,8626777 $\times 10^{-6}$	0,9995	0,9995	+
27	24144,594	24144,877	0,28289447	0,0011716679	0,51100145	3,8493238 $\times 10^{-7}$	2,9313389 $\times 10^{-6}$	0,9995	0,9995	+
28	24144,55	24144,775	0,22496484	0,00093174169	0,50933815	2,4454528 $\times 10^{-7}$	1,4656694 $\times 10^{-6}$	0,9995	0,9995	+
29	24144,515	24144,694	0,17913943	0,0007419467	0,50808497	1,5561972 $\times 10^{-7}$	7,3283472 $\times 10^{-7}$	0,9995	0,9995	+
30	24144,487	24144,629	0,14281865	0,00059151662	0,5184629	9,9186197 $\times 10^{-8}$	3,6641736 $\times 10^{-7}$	0,9995	0,9995	+
31	24144,492	24144,629	0,13762337	0,0005699904	0,0049952938	9,9186197 $\times 10^{-8}$	1,8320868 $\times 10^{-7}$	0,9995	0,95319696	–
32	24144,492	24144,629	0,1373605	0,00056891031	1,6021599 $\times 10^{-5}$	9,9186197 $\times 10^{-8}$	9,1604339 $\times 10^{-8}$	0,9995	0,99644866	–
33	24144,492	24144,629	0,13735958	0,00056890648	9,7776436 $\times 10^{-9}$	9,9186197 $\times 10^{-8}$	4,580217 $\times 10^{-8}$	0,9995	0,99949032	–

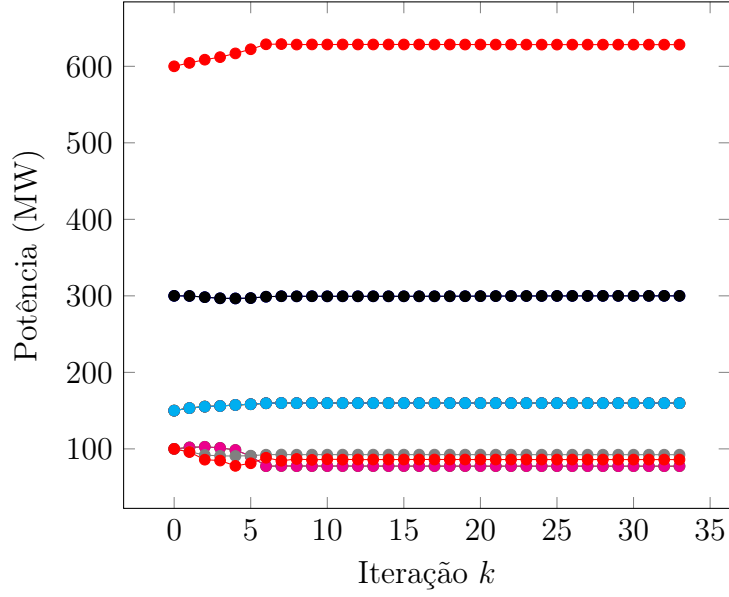


Figura 6.14: Sequência de potências geradas pelo método PDPCBLM para o PDE-PV de 13 geradores (Parte II).

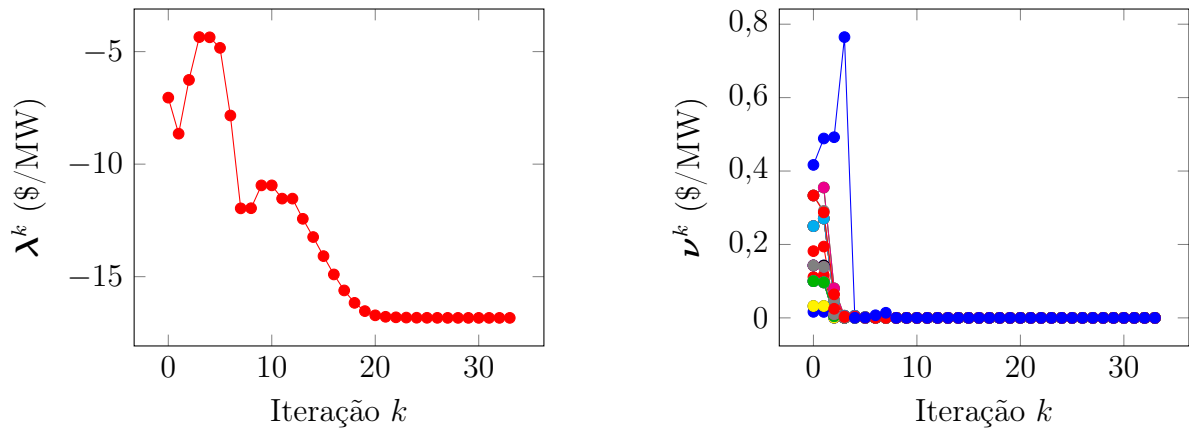


Figura 6.15: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PV de 13 geradores (Parte II).

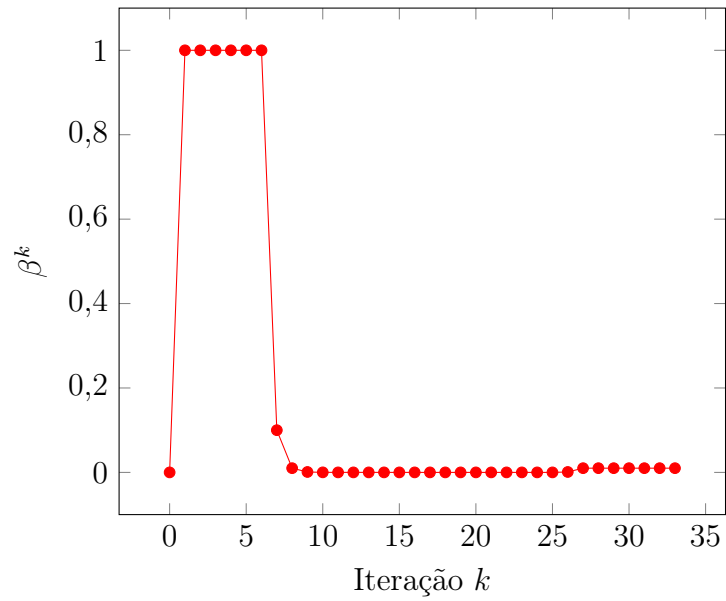


Figura 6.16: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PV de 13 geradores (Parte II).

Tabela 6.16: Resultados – PDE-PV: 13 geradores (Parte II)

Potência Gerada (MW)	Métodos			PDPCBLM
	AGH (Samed, 2004)	AGHCOE (Samed, 2004)	ACO (Pothiya et al., 2010)	
$P_{G,1}$	643,5755	617,1943	628,32	628,33045
$P_{G,2}$	328,6986	300,4246	299,06	299,99307
$P_{G,3}$	310,0123	297,2810	299,17	299,99307
$P_{G,4}$	124,9541	165,6300	159,73	159,74125
$P_{G,5}$	145,0035	159,0019	159,73	159,74125
$P_{G,6}$	150,8724	164,5568	159,73	159,74125
$P_{G,7}$	160,4046	140,7621	159,73	159,74125
$P_{G,8}$	172,1755	157,1320	159,72	159,74125
$P_{G,9}$	176,5106	165,9530	159,72	159,74125
$P_{G,10}$	62,44434	76,51676	75,47	77,40922
$P_{G,11}$	92,61867	74,81078	77,33	77,40922
$P_{G,12}$	62,40476	80,65879	92,10	92,40855
$P_{G,13}$	90,32517	120,0779	90,59	86,00891
Função Objetivo (\$)	25374,14	24799,03	24169,63	24144,49

Tabela 6.17: Características do sistema – PDE-PV: 19 geradores

Gerador (k)	$\frac{P_{G,k}}{(\text{MW})}$	$\overline{P_{G,k}}$ (MW)	a_k (\$/MW ²)	b_k (\$/MW)	c_k (\$)	e_k (\$)	f_k (1/MW)
1	100	300	0,0097	6,8	119	90	0,72
2	120	438	0,0055	4	90	79	0,05
3	100	250	0,0055	4	45	0	0
4	8	25	0,0025	0,85	0	0	0
5	50	63,75	0	5,28	0,891	0	0
6	150	300	0,008	3,5	110	0	0
7	50	63,75	0	5,439	21	0	0
8	100	500	0,0075	6	88	50	0,52
9	200	600	0,0085	6	55	0	0
10	15	40	0,009	5,2	90	0	0
11	50	150	0,0045	1,6	65	0	0
12	25	75	0,0025	0,85	78	58	0,02
13	50	63,75	0	2,55	49	0	0
14	0	95	0,0045	1,6	85	0	0
15	20	220	0,0065	4,7	80	92	0,75
16	15	80	0,0045	1,4	90	0	0
17	15	80	0,0025	0,85	10	0	0
18	50	230	0,0045	1,6	25	0	0
19	400	500	0,008	5,5	90	0	0
Demanda: $D = 2908$ MW							

Fonte: Balamurugan et al. (2014)

Os parâmetros apresentados na Tabela 6.18 foram utilizados na inicialização do método.

A escolha do ponto inicial, neste caso, foi o valor médio entre os limitantes mínimo e máximo de saída de potência ativa de cada unidade geradora. A Tabela 6.19 apresenta os resultados obtidos, nas 23 iterações necessárias para a convergência. Na referida tabela, percebe-se que o parâmetro de barreira sofreu penalização nas iterações 4, 5, 6, 7, 8 e 9. O mesmo fato verificado nos casos anteriores se repete: a redução significativa dos resíduos somente ocorreu a partir do momento em que o parâmetro η atingiu o limitante $\bar{\eta}$, na iteração 21. A diferença relativa entre a função objetivo original e aproximada, na solução ótima determinada, é de 0,0015661004%.

Este caso difere dos demais, devido ao fato de haver geradores com características variadas: lineares, quadráticas e não-diferenciáveis. Além disso, o artigo de Balamurugan et al. (2014) não resolveu apenas o PDE, mas sim uma versão do mesmo combinada com o despacho ambiental. Por outro lado, na revisão bibliográfica realizada verificou-se que nenhum outro autor resolveu este exemplo. Por isso, a fim de verificar a consistência dos resultados, o problema suavizado foi resolvido utilizando quatro diferentes pacotes computacionais, considerados estado da arte em otimização não-linear. Como não podemos alterar o valor do parâmetro η ao longo das iterações no pacote computacional, seu valor

Tabela 6.18: Inicialização – PDE-PV: 19 geradores

Fatores de aumento/redução da correção de inércia	$\kappa_1 = \kappa_2 = 10$
Fatores de redução do parâmetro de suavização	$\rho_1 = 2$ e $\rho_2 = 0$
Limitante da atualização do parâmetro de suavização	$\bar{\eta} = 10^{-6}$
Parâmetro de barreira/relaxação inicial	$\mu^0 = 10$
Fatores de atualização do parâmetro de barreira	$\tau_1 = \tau_2 = 0,1$
Fator de controle da escolha entre passo previsor ou corretor	$\chi = 1$
Fator de redução do passo	$\sigma = 0,9995$
Precisão de parada	$\varepsilon = 10^{-6}$
	1 200
	2 279
	3 175
	4 16,5
	5 56,875
	6 225
	7 56,875
	8 300
	9 400
Potência inicial das unidades geradoras (MW)	10 27,5
	11 100
	12 50
	13 56,875
	14 47,5
	15 120
	16 47,5
	17 47,5
	18 140
	19 450
Estimadores dos multiplicadores de Lagrange	$\delta^0 = 10 \cdot \mathbb{1}_p$

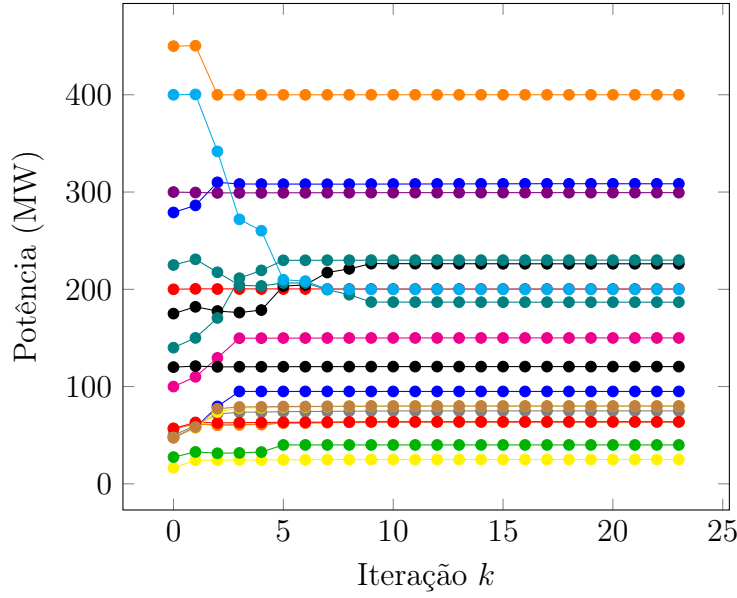


Figura 6.17: Sequência de potências geradas pelo método PDPCBLM para o PDE-PV de 19 geradores.

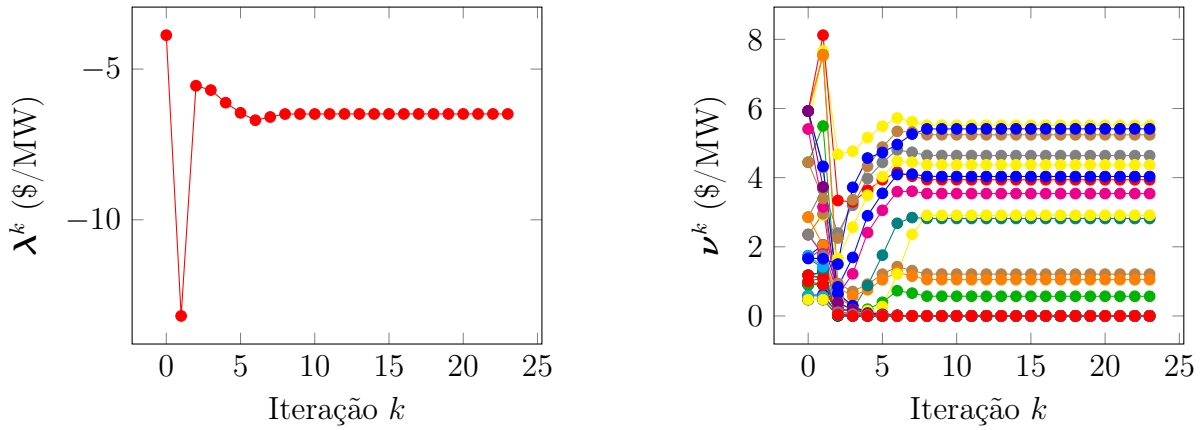


Figura 6.18: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PV de 19 geradores.

foi fixado no valor final $9,5367432 \times 10^{-7}$. Todos os pacotes foram executados com suas opções no *default* e iniciados no mesmo ponto do método PDPCBLM. Os pacotes utilizados foram: IPOPT (*Interior Point OPTimizer*), KNITRO (*Nonlinear Interior point Trust Region Optimization*), MINOS (*Modular In-core Nonlinear Optimization System*) e SNOPT (*Sparse Nonlinear OPTimizer*), em suas versões de demonstração, através da interface do GAMS 24.3.3. As Figuras 6.17, 6.18 e 6.19 apresentam a trajetória das potências geradas, dos multiplicadores de Lagrange e do parâmetro de correção de inércia, respectivamente, pelo método aqui proposto.

Pela análise da Tabela 6.20, nota-se que a solução ótima obtida pelo método PDPCBLM é consistente quando comparada com as determinadas pelos pacotes computacionais testados. Além disso, a solução obtida é melhor do que as determinadas pelos pacotes,

Tabela 6.19: Resultados (Iterações) – PDE-PV: 19 geradores

Iteração k	$f(\mathbf{x}^k)$	$f_{\eta^k}(\mathbf{x}^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	19406,556	19645,351	238,7945	1,2304837	111,875	1	10	0	0	–
1	19751,814	20012,572	260,75881	1,3201766	114,94867	1	10	0,9995	0,9995	–
2	18589,072	18793,134	204,0618	1,0977514	70,437495	0,5	1	0,93185112	0,98413362	+
3	17932,899	18076,777	143,87812	0,80231376	194,84732	0,25	0,1	0,32327857	0,48796326	–
4	17856,22	17951,939	95,719225	0,53605535	69,481793	0,125	0,082884383	0,054316582	0,50893798	–
5	17618,959	17685,067	66,107677	0,37520762	85,214007	0,0625	0,078583539	0,2834969	0,40829824	–
6	17613,504	17656,738	43,233637	0,24545733	34,15217	0,03125	0,056497633	0,023061547	0,66101797	–
7	17570,217	17602,105	31,887448	0,1814858	20,953807	0,015625	0,055204361	0,45274914	0,78839698	–
8	17562,609	17585,13	22,520415	0,12822932	16,357784	0,0078125	0,0304028	0,41166065	0,9995	+
9	17552,571	17570,42	17,848812	0,10168774	1,4727845	0,00390625	0,018274785	0,9995	0,9995	+
10	17552,188	17564,705	12,516336	0,071309263	1,7244641	0,001953125	0,0018274785	0,9995	0,9995	+
11	17551,844	17560,682	8,8381653	0,050354625	1,7193544	0,0009765625	0,00018274785	0,9995	0,9995	+
12	17551,603	17557,845	6,2421808	0,035564733	1,7195347	0,00048828125	$1,8274785 \times 10^{-5}$	0,9995	0,9995	+
13	17551,433	17555,843	4,4102748	0,02512772	1,7196246	0,00024414063	$1,8274785 \times 10^{-6}$	0,9995	0,9995	+
14	17551,312	17554,428	3,1167244	0,017757786	1,7196888	0,00012207031	$1,8274785 \times 10^{-7}$	0,9995	0,9995	+
15	17551,226	17553,429	2,2029517	0,012551554	1,7197325	$6,1035156 \times 10^{-5}$	$1,8274785 \times 10^{-8}$	0,9995	0,9995	+
16	17551,166	17552,723	1,5572694	0,0088727403	1,7197639	$3,0517578 \times 10^{-5}$	$1,8274785 \times 10^{-9}$	0,9995	0,9995	+
17	17551,123	17552,224	1,1009294	0,0062727004	1,719786	$1,5258789 \times 10^{-5}$	$1,8274785 \times 10^{-10}$	0,9995	0,9995	+
18	17551,093	17551,871	0,77836145	0,0044348318	1,7198017	$7,6293945 \times 10^{-6}$	$1,8274785 \times 10^{-11}$	0,9995	0,9995	+
19	17551,072	17551,622	0,55032807	0,003135581	1,7198127	$3,8146973 \times 10^{-6}$	$1,8274785 \times 10^{-12}$	0,9995	0,9995	+
20	17551,057	17551,446	0,38911241	0,0022170313	1,7198205	$1,9073486 \times 10^{-6}$	$1,8274785 \times 10^{-13}$	0,9995	0,9995	+
21	17551,046	17551,321	0,27512988	0,0015675982	0,027897035	$9,5367432 \times 10^{-7}$	$1,8274785 \times 10^{-14}$	0,9995	0,9995	+
22	17551,046	17551,321	0,2748674	0,0015661027	$2,1528469 \times 10^{-5}$	$9,5367432 \times 10^{-7}$	$1,8274785 \times 10^{-15}$	0,9995	0,99949921	–
23	17551,046	17551,321	0,27486701	0,0015661004	$1,0809213 \times 10^{-8}$	$9,5367432 \times 10^{-7}$	$1,8274785 \times 10^{-16}$	0,9995	0,9995	+

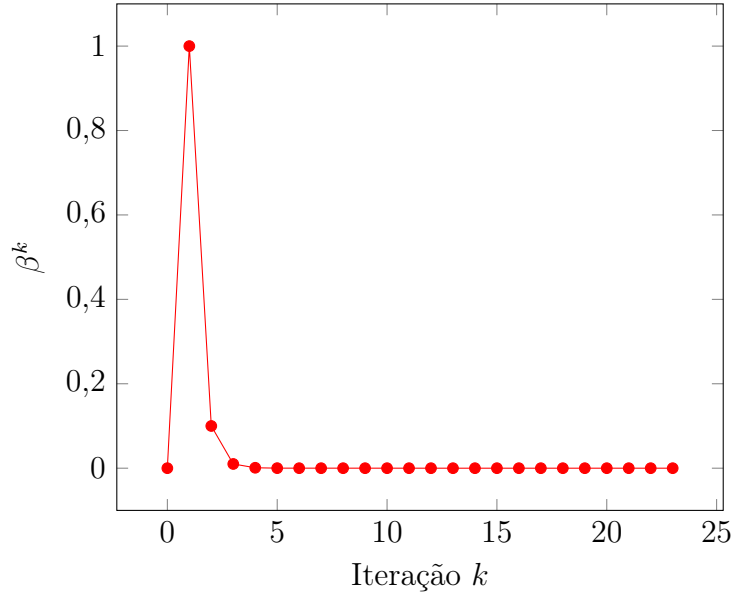


Figura 6.19: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PV de 19 geradores.

Tabela 6.20: Resultados – PDE-PV: 19 geradores

Potência Gerada (MW)	IPOPT	KNITRO	Métodos		
			MINOS	SNOPT	PDPCBLM
$P_{G,1}$	204,72	100	300	204,72	200,35634
$P_{G,2}$	245,665	434,164	371,317	182,841	308,49096
$P_{G,3}$	250	250	201,548	250	226,22324
$P_{G,4}$	25	25	25	25	25
$P_{G,5}$	63,75	63,75	63,75	63,75	63,75
$P_{G,6}$	211,234	300	169,814	260,122	186,77848
$P_{G,7}$	63,75	63,75	63,75	63,75	63,75
$P_{G,8}$	305,412	214,789	178,54	323,536	299,37001
$P_{G,9}$	200	222,796	200	200	200
$P_{G,10}$	40	40	40	40	40
$P_{G,11}$	150	150	150	150	150
$P_{G,12}$	75	75	75	75	75
$P_{G,13}$	63,75	63,75	63,75	63,75	63,75
$P_{G,14}$	95	95	95	95	95
$P_{G,15}$	124,72	20	120,531	120,531	120,53097
$P_{G,16}$	80	80	80	80	80
$P_{G,17}$	80	80	80	80	80
$P_{G,18}$	230	230	230	230	230
$P_{G,19}$	400	400	400	400	400
Função Objetivo (\$)	17565,957	17407,540	17820,054	17689,949	17551,046
Número de Iterações	107	163	119	241	23

excetuando o KNITRO, que obteve a solução de melhor qualidade. O elevado número de iterações dos pacotes mostra a dificuldade de resolução do problema, quando o valor do parâmetro η é baixo. Para se ter uma ideia, no trabalho de Ferreira (2013), o pacote KNITRO foi utilizado para resolver problemas de FPO. Os resultados apresentados pela autora indicam que o pacote resolveu o caso do sistema IEEE 118 barras, com seus parâmetros no modo *default*, em 10 iterações. O FPO para o sistema IEEE 118 barras possui um número maior de restrições e variáveis que este caso do PDE-PV. Ainda assim o pacote realizou um número significativamente menor de iterações para resolver o FPO do que o que precisou para resolver esta instância de 19 geradores do PDE-PV.

6.2 PDE-PPV

Agora, apresentam-se os resultados obtidos na resolução de três instâncias distintas do Problema de Despacho Econômico com representação das Perdas e Ponto de Carregamento de Válvula (PDE-PPV), a saber o sistema de 6 barras de Wood & Wollenberg (1996), o sistema IEEE-14 barras e o sistema IEEE-30 barras. Os resultados aqui obtidos serão comparados aos apresentados por Ravi et al. (2006). No artigo dos referidos autores, o PDE-PPV é resolvido através de um método heurístico, porém um ciclo externo atualiza os coeficientes-B de acordo com o método apresentado por Saadat (1999) e que foi explicado no Apêndice A.

Neste trabalho, a fim de realizar uma comparação justa, a mesma estratégia foi adotada: um ciclo externo utiliza o método de Newton para calcular a geração de potência ativa da barra *slack*, a geração de potência reativa, as magnitudes e ângulos de tensão do sistema para, em seguida, calcularem-se os coeficientes-B. Após o cálculo dos coeficientes-B, o método PDPCBLM é utilizado para resolver o PDE-PPV. A solução obtida é, então, utilizada em um novo fluxo de carga (exceto a geração da barra *slack*), para calcular novos valores dos coeficientes-B. Considerou-se que a otimalidade é atingida quando, ao atualizar os coeficientes-B para o do novo ponto, a restrição de igualdade continua satisfeita com uma precisão $\bar{\varepsilon} = 10^{-6}$. Outro critério de parada viável seria a diferença entre as gerações de potência ativa da barra *slack*, determinada pelo método de pontos interiores e o fluxo de carga, como feito no trabalho de Ravi et al. (2006). A Figura 6.20 apresenta um esquema da metodologia considerada para os testes.

Assim, ocorre a resolução de uma sequência de instâncias do PDE-PPV, com coeficientes-B atualizados. Para inicializar o algoritmo, os coeficientes-B do primeiro problema a ser resolvido são calculados com base no banco de dados. Para a resolução do fluxo de carga, empregou-se o pacote Matpower 4.1, o qual encontra-se disponível no *site*:

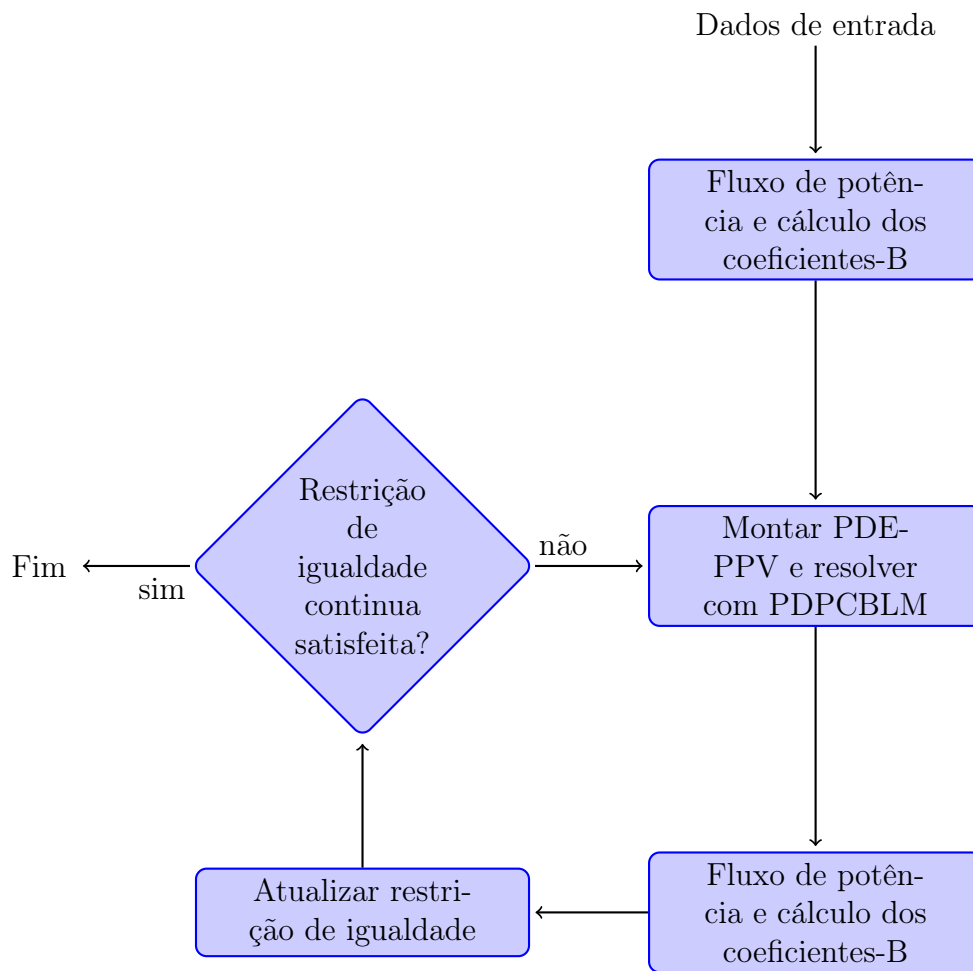


Figura 6.20: Esquema de resolução do PDE-PPV.

<http://www.pserc.cornell.edu/matpower/>. O Matpower é um pacote de rotinas em Matlab, para a resolução de problemas de fluxo de carga e FPO. O ciclo interno do método PDPCBLM é sempre inicializado com a solução final do ciclo interno anterior, incluindo os estimadores dos multiplicadores de Lagrange. As únicas variáveis reinicializadas são aquelas relacionadas à correção de inércia.

6.2.1 Sistema 6 Barras de Wood e Wollenberg

Para a resolução do PDE-PPV associado ao sistema de 6 barras, considerou-se que o valor da demanda de potência ativa e reativa eram de, respectivamente, 65 MW e 65 MVar, conforme o primeiro padrão de carga estudado no trabalho de Ravi et al. (2006) (no trabalho de Wood & Wollenberg (1996), as potências são iguais a 70). Os dados do sistema podem ser encontrados no Apêndice C. Os coeficientes de custo das unidades geradoras são apresentados na Tabela 6.21. A formulação do modelo do PDE-PPV para este caso possui:

Tabela 6.21: Características do sistema – PDE-PPV: 6 Barras

Gerador (k)	$\frac{P_{G,k}}{(\text{MW})}$	$\overline{P_{G,k}}$ (MW)	a_k (\$/MW ²)	b_k (\$/MW)	c_k (\$)	e_k (\$)	f_k (1/MW)
1	50	200	0,00533	11,669	213,1	150	0,063
2	37,5	150	0,00889	10,333	200	75	0,0315
3	45	180	0,00741	10,833	240	110	0,042
Demanda (soma das demandas nas barras): $D = 195$ MW							
Fonte: Ravi et al. (2006)							

Tabela 6.22: Inicialização – PDE-PPV: 6 Barras

Fatores de aumento/redução da correção de inércia	$\kappa_1 = \kappa_2 = 10$	
Fatores de redução do parâmetro de suavização	$\rho_1 = 2,5$ e $\rho_2 = 0$	
Limitante da atualização do parâmetro de suavização	$\bar{\eta} = 10^{-5}$	
Parâmetro de barreira/relaxação inicial	$\mu^0 = 1$	
Fatores de atualização do parâmetro de barreira	$\tau_1 = \tau_2 = 0,5$	
Fator de controle da escolha entre passo previsor ou corretor	$\chi = 1$	
Fator de redução do passo	$\sigma = 0,9995$	
Precisão de parada	$\varepsilon = 10^{-6}$	
Potência inicial das unidades geradoras (MW)	1	125
	2	93,75
	3	112,5
Estimadores dos multiplicadores de Lagrange	$\delta^0 = \mathbb{1}_p$	

- 3 variáveis;
- 1 restrição de igualdade;
- 6 restrições de desigualdade.

Os parâmetros apresentados na Tabela 6.22 foram utilizados na inicialização do método. As gerações iniciais foram tomadas como sendo o valor médio entre o mínimo e o máximo das unidades geradoras. A Tabela 6.23 apresenta os resultados obtidos, tendo sido realizado um total de 22 iterações. As linhas separando algumas iterações indicam a convergência do método PDPCBLM no ciclo interno e a subsequente atualização dos coeficientes-B.

Desta tabela, é perceptível que o primeiro ciclo interno é o mais custoso. Isto porque é nele que o método busca encontrar uma região de mínimo local. Nos ciclos internos posteriores, nota-se que o número de iterações passa a diminuir, pois o método já se encontra próximo do ponto de mínimo local. As Figuras 6.21, 6.22 e 6.23 apresentam a sequência de potências geradas, multiplicadores de Lagrange e parâmetro de correção de inércia, respectivamente, ao longo das iterações do método. Nota-se que a mudança de coeficientes-B na passagem da 15^a para 16^a iteração permitiu a correção da potência $P_{G,3}$, enquanto as demais permaneceram praticamente no mesmo patamar. A partir daí,

Tabela 6.23: Resultados (Iterações) – PDE-PPV: 6 Barras

Iteração k	$f(\mathbf{x}^k)$	$f_{jk}(\mathbf{x}^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	4811,0313	4986,3824	175,35107	3,6447709	118,76398	1	1	0	0	–
1	3188,4951	3361,0303	172,53524	5,411181	44,029723	1	1	0,9995	0,9995	–
2	3204,3929	3290,2514	85,858491	2,6793996	5,8025513	0,4	0,5	0,9995	0,9995	+
3	3087,3965	3173,2589	85,862433	2,7810627	8,4013598	0,16	0,25	0,056159184	0,10480428	–
4	3034,0382	3107,3245	73,286324	2,4154714	0,76574186	0,064	0,125	0,068620393	0,24656448	–
5	3035,5358	3081,1907	45,654972	1,5040169	0,79593105	0,0256	0,41782268	0,9995	0,9995	–
6	3035,3951	3062,7337	27,338536	0,90065822	1,013797	0,01024	0,20891134	0,20029791	0,99410184	–
7	3035,654	3051,6602	16,006204	0,52727368	0,4194437	0,004096	0,10445567	0,69552944	0,9995	–
8	3035,5778	3045,1899	9,6121385	0,31664939	0,33732884	0,0016384	0,052227835	0,9995	0,9995	+
9	3035,4206	3041,3935	5,9728571	0,19677197	0,19397608	0,00065536	0,059841921	0,9995	0,9995	+
10	3035,351	3039,1037	3,7526773	0,1236324	0,074264618	0,000262144	0,02992096	0,9995	0,9995	+
11	3035,3331	3037,6889	2,3558072	0,077612808	0,021984723	0,0001048576	0,01496048	0,9995	0,9995	+
12	3035,3297	3036,8092	1,479525	0,048743471	0,0059865365	4,194304 $\times 10^{-5}$	0,0074802401	0,9995	0,9995	+
13	3035,3291	3036,2599	0,93081579	0,030666058	0,0024008891	1,6777216 $\times 10^{-5}$	0,00374012	0,9995	0,9995	+
14	3035,329	3035,9156	0,58660103	0,019325781	8,4348544 $\times 10^{-6}$	6,7108864 $\times 10^{-6}$	0,00187006	0,9995	0,9995	+
15	3035,329	3035,9156	0,58662934	0,019326713	5,3411894 $\times 10^{-7}$	6,7108864 $\times 10^{-6}$	0,00093503001	0,9995	0,9995	+
16	3003,4141	3003,9969	0,58285123	0,01940629	0,02915303	6,7108864 $\times 10^{-6}$	0,000467515	0,9995	0,93958934	–
17	3003,4312	3004,0154	0,58427428	0,01945356	1,5341077 $\times 10^{-5}$	6,7108864 $\times 10^{-6}$	0,0002337575	0,9995	0,99947525	–
18	3003,4312	3004,0154	0,58427687	0,019453646	7,6705593 $\times 10^{-9}$	6,7108864 $\times 10^{-6}$	0,00011687875	0,9995	0,9995	+
19	3003,6482	3004,2324	0,58427659	0,019452231	2,3836514 $\times 10^{-5}$	6,7108864 $\times 10^{-6}$	5,8439376 $\times 10^{-5}$	0,9995	0,99910767	–
20	3003,6483	3004,2325	0,58428274	0,019452435	1,1918791 $\times 10^{-8}$	6,7108864 $\times 10^{-6}$	2,9219688 $\times 10^{-5}$	0,9995	0,9995	+
21	3003,6458	3004,2301	0,58428268	0,019452449	1,4317624 $\times 10^{-7}$	6,7108864 $\times 10^{-6}$	1,4609844 $\times 10^{-5}$	0,9995	0,99949563	–
22	3003,6458	3004,2301	0,58428269	0,01945245	1,5447199 $\times 10^{-9}$	6,7108864 $\times 10^{-6}$	7,3049219 $\times 10^{-6}$	0,9995	0,9995	–

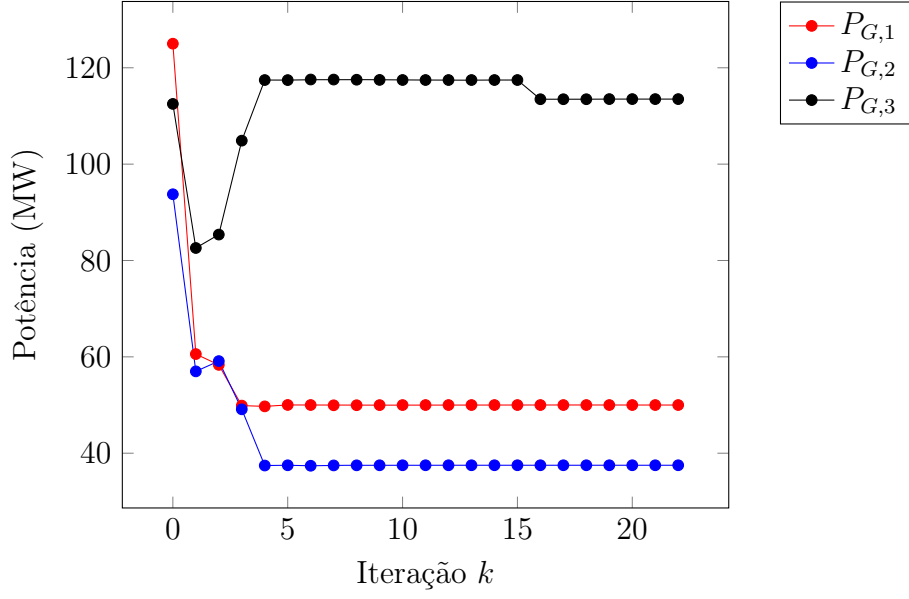


Figura 6.21: Sequência de potências geradas pelo método PDPCBLM para o PDE-PPV de 6 Barras.

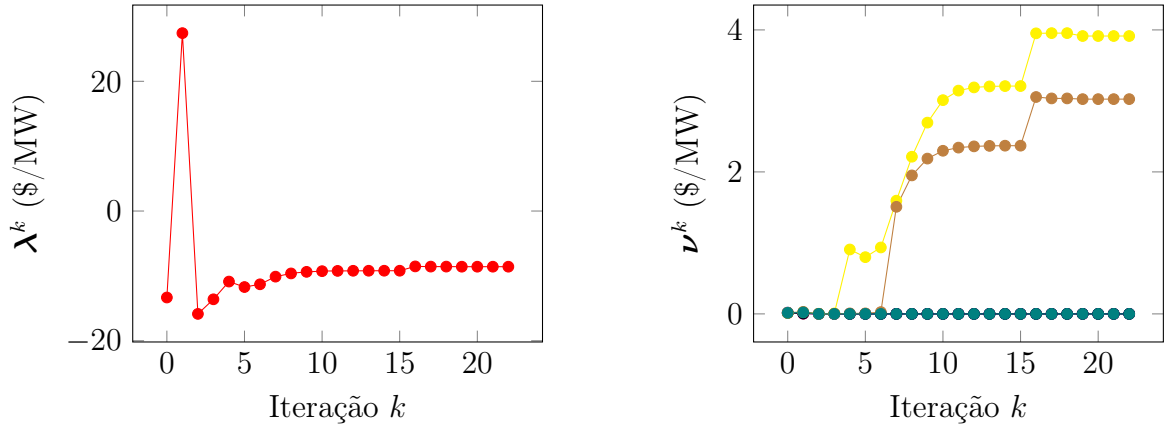


Figura 6.22: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PPV de 6 Barras.

as iterações seguintes apenas refinam a solução, até que a precisão desejada seja obtida. Por fim, a Tabela 6.24 apresenta o despacho a ser realizado, as perdas e a função objetivo de custo total, em comparação ao apresentado por Ravi et al. (2006), enquanto a Tabela 6.25 apresenta os valores das magnitudes e ângulos de tensão obtidos no último fluxo de carga executado para calcular os coeficientes-B. Nota-se que o método apresentou uma melhora na função objetivo, além de ser factível para o sistema elétrico, uma vez que o despacho realizado satisfaz o fluxo de carga do ciclo externo com a precisão definida. A diferença relativa entre a função objetivo original e aproximada, calculadas na solução ótima determinada, foi de 0,01945245%.

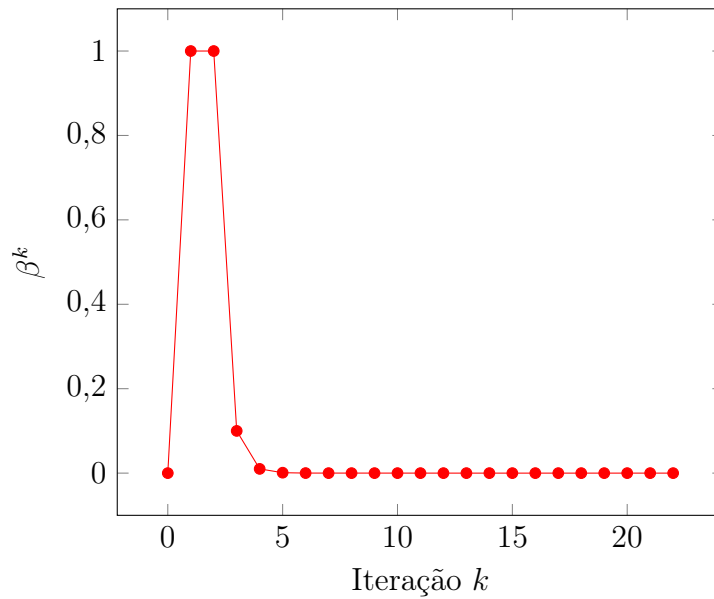


Figura 6.23: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PPV de 6 Barras.

Tabela 6.24: Resultados – PDE-PPV: 6 Barras

Potência Gerada (MW)	Métodos	
	IFEP (Ravi et al., 2006)	PDPCBLM
$P_{G,1}$	50	50
$P_{G,2}$	37,5	37,5
$P_{G,3}$	114,01	113,5015
Perdas L (MW)	6,51	6,0015
Função Objetivo (\$)	3007,77	3003,64

Tabela 6.25: Magnitudes e Ângulos de Tensão – PDE-PPV: 6 Barras

Barra	V	θ (graus)
1	1,05	0
2	1,05	-0,9670899
3	1,07	1,3237134
4	0,99263929	-2,1507633
5	0,99074451	-2,075729
6	1,0105419	-1,401307

Tabela 6.26: Características do sistema – PDE-PPV: 14 Barras

Gerador (k)	$\frac{P_{G,k}}{(\text{MW})}$	$\overline{P_{G,k}}$ (MW)	a_k (\$/MW ²)	b_k (\$/MW)	c_k (\$)	e_k (\$)	f_k (1/MW)
1	10	100	0,00375	2	0	80	0,073
2	20	50	0,0175	1,75	0	30	0,032
3	15	80	0,0625	1	0	50	0,051
6	10	45	0,00834	3,25	0	25	0,026
8	10	45	0,025	3	0	25	0,026
Demanda (soma das demandas nas barras): $D = 289$ MW							

Fonte: Ravi et al. (2006)

6.2.2 Sistema IEEE-14 Barras

Para a resolução do PDE-PPV associado ao sistema de 14 barras, considerou-se os dados do sistema que podem ser encontrados no Apêndice C. A demanda de potência ativa na barra 8 foi alterada de 0 MW para 30 MW, conforme apresentado no artigo de referência utilizado por Ravi et al. (2006), e também adotado no artigo de Arul et al. (2013). Outra alteração foi realizada nas potências geradas iniciais do banco de dados, igual ao ponto inicial do método PDPCBLM. Isto porque o banco de dados possuía gerações com valor nulo, o que inviabilizava o cálculo dos coeficientes-B, já que este supõe que todas as barras de geração estão gerando alguma quantidade. Isto não representa problema, já que nenhum dos geradores possui limitante mínimo igual a 0, isto é, todos estão sempre ligados. Os coeficientes de custo das unidades geradoras são apresentados na Tabela 6.26. A formulação do modelo do PDE-PPV para este caso possui:

- 5 variáveis;
- 1 restrição de igualdade;
- 10 restrições de desigualdade.

Novamente, as gerações iniciais foram tomadas como sendo o valor médio entre o mínimo e o máximo das unidades geradoras, e os demais parâmetros de inicialização são apresentados na 6.27. A Tabela 6.28 apresenta os resultados obtidos. Neste caso, o método convergiu após 24 iterações, ao longo de 4 ciclos externos. Observações análogas às verificadas no caso de 6 barras podem ser feitas: o número de iterações internas se reduz à medida em que novos ciclos externos são iniciados.

As Figuras 6.24, 6.25 e 6.26 apresentam a sequência de potências geradas, multiplicadores de Lagrange e parâmetro de correção de inércia, respectivamente, ao longo das iterações do método. É interessante perceber que, neste caso, não foi necessária a utilização da correção de inércia. Este fato provavelmente também se manifesta no comportamento

Tabela 6.27: Inicialização – PDE-PPV: 14 Barras

Fatores de aumento/redução da correção de inércia	$\kappa_1 = \kappa_2 = 10$
Fatores de redução do parâmetro de suavização	$\rho_1 = 3$ e $\rho_2 = 0$
Limitante da atualização do parâmetro de suavização	$\bar{\eta} = 10^{-7}$
Parâmetro de barreira/relaxação inicial	$\mu^0 = 5$
Fatores de atualização do parâmetro de barreira	$\tau_1 = \tau_2 = 0,4$
Fator de controle da escolha entre passo previsor ou corretor	$\chi = 1$
Fator de redução do passo	$\sigma = 0,9995$
Precisão de parada	$\varepsilon = 10^{-6}$

Potência inicial das unidades geradoras (MW)	1	55
	2	35
	3	47,5
	6	27,5
	8	27,5

Estimadores dos multiplicadores de Lagrange	$\delta^0 = 10 \cdot \mathbb{1}_p$
---	------------------------------------

menos errático dos multiplicadores de Lagrange, em especial na ordem de grandeza dos multiplicadores das restrições de desigualdade. Também é importante notar que alguns multiplicadores ν não são nulos, já que existem restrições ativas na solução ótima determinada. Isto está em concordância com as condições de KKT. A diferença relativa entre o valor da função objetivo original e aproximada, calculadas na solução ótima determinada, foi de 0,00094470632%. Assim como nas instâncias do PDE-PV, pode-se notar novamente que o comprimento do passo tende a ser máximo em iterações que a direção do corretor é utilizada. Por outro lado, quando a direção do previsor é utilizada, o comprimento do passo tende a não ser o valor máximo.

Na Tabela 6.29, nota-se que o valor obtido pelo método PDPCBLM foi um pouco superior ao do autor Ravi et al. (2006). Entretanto, ao substituir a solução do autor na função objetivo, observamos que o resultado do autor também possui valor 1070,56. Como as demais casas decimais não são apresentadas, não pudemos ter certeza se o valor no artigo realmente foi obtido. O comando `rand` do Matlab foi utilizado para gerar pequenas perturbações na solução do autor e calcular o valor da função objetivo, mas nenhum valor determinado foi inferior a 1070. De qualquer modo, a solução obtida pelo método PDPCBLM é de boa qualidade. A Tabela 6.30 apresenta o valor das magnitudes e ângulos de tensão no último fluxo de carga.

6.2.3 Sistema IEEE-30 Barras

Para a resolução do PDE-PPV associado ao sistema de 30 barras, consideram-se os dados do sistema que podem ser encontrados no Apêndice C. A mesma alteração feita para o sistema IEEE-14 foi aplicada neste caso: as potências geradas do banco de dados foram

Tabela 6.28: Resultados (Iterações) – PDE-PPV: 14 Barras

Iteração k	$f(\mathbf{x}^k)$	$f_{\eta^k}(\mathbf{x}^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	686,70723	828,68545	141,97822	20,675219	98,043044	1	5	0	0	–
1	1131,5943	1240,7854	109,19113	9,6493174	823,71323	1	5	0,82255532	0,94910229	–
2	1005,437	1055,4897	50,052694	4,978203	33,700745	0,33333333	2	0,28693141	0,16765642	–
3	1020,793	1041,8118	21,01881	2,0590669	8,8732072	0,11111111	1,9656798	0,31035632	0,9995	+
4	1062,3486	1074,0497	11,70117	1,1014385	1,7378231	0,037037037	1,4334834	0,9995	0,70609798	–
5	1062,1773	1067,5428	5,3655089	0,50514251	2,5801022	0,012345679	24,738735	0,9995	0,9995	+
6	1061,8596	1065,2261	3,3665482	0,31704269	3,852449	0,0041152263	22,524897	0,9995	0,9995	+
7	1066,3846	1067,9468	1,5621584	0,14649108	3,8251861	0,0013717421	17,254137	0,9995	0,9995	+
8	1069,2442	1070,3387	1,0944738	0,10235957	2,1115615	0,00045724737	9,9505618	0,9995	0,9995	+
9	1071,1742	1071,6617	0,48754205	0,04551473	1,0811557	0,00015241579	3,9802247	0,9995	0,9995	+
10	1071,747	1072,1104	0,36342628	0,033909708	0,058619886	5,0805263 $\times 10^{-5}$	1,5920899	0,9995	0,9995	+
11	1071,8522	1072,0109	0,15868059	0,014804335	1,0809897	1,6935088 $\times 10^{-5}$	0,63683596	0,9995	0,9995	+
12	1071,7908	1071,9131	0,12232788	0,01141341	0,035028312	5,6450293 $\times 10^{-6}$	0,25473438	0,9995	0,9995	+
13	1071,7918	1071,8443	0,052573964	0,0049052406	1,0811178	1,8816764 $\times 10^{-6}$	0,10189375	0,9995	0,9995	+
14	1071,7696	1071,8103	0,040710094	0,0037983999	0,039743886	6,2722547 $\times 10^{-7}$	0,040757501	0,9995	0,9995	–
15	1071,7698	1071,7873	0,017490572	0,0016319337	1,0809792	2,0907516 $\times 10^{-7}$	0,016303001	0,9995	0,9995	–
16	1071,7624	1071,7776	0,013559753	0,0012651827	1,2378687	6,9691719 $\times 10^{-8}$	0,0065212002	0,9995	0,9995	–
17	1071,7646	1071,7753	0,010649175	0,0009361138	0,1961475	6,9691719 $\times 10^{-8}$	0,0026084801	0,9995	0,9995	–
18	1071,7651	1071,7752	0,010118375	0,00094408507	0,0089793165	6,9691719 $\times 10^{-8}$	0,001043392	0,9995	0,9995	+
19	1071,7652	1071,7752	0,010092822	0,00094170086	2,5831748 $\times 10^{-5}$	6,9691719 $\times 10^{-8}$	0,00041735681	0,9995	0,9995	+
20	1071,7652	1071,7752	0,010092748	0,00094169398	1,3094772 $\times 10^{-8}$	6,9691719 $\times 10^{-8}$	0,00016694273	0,9995	0,9995	+
21	1070,5684	1070,5785	0,010113936	0,00094472582	0,00039129112	6,9691719 $\times 10^{-8}$	6,677709 $\times 10^{-5}$	0,9995	0,9919037	–
22	1070,5679	1070,578	0,010113756	0,00094470943	1,9564537 $\times 10^{-7}$	6,9691719 $\times 10^{-8}$	2,6710836 $\times 10^{-5}$	0,9995	0,9995	+
23	1070,5625	1070,5726	0,010113672	0,00094470633	4,0041789 $\times 10^{-7}$	6,9691719 $\times 10^{-8}$	1,0684334 $\times 10^{-5}$	0,9995	0,99946626	–
24	1070,5625	1070,5726	0,010113671	0,00094470632	1,997023 $\times 10^{-9}$	6,9691719 $\times 10^{-8}$	4,2737338 $\times 10^{-6}$	0,9995	0,9995	+

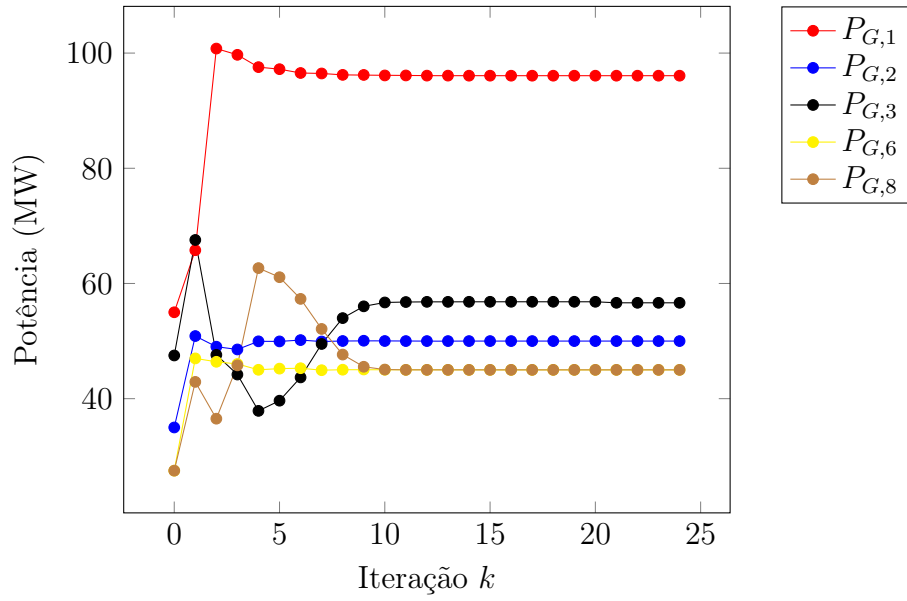


Figura 6.24: Sequência de potências geradas pelo método PDPCBLM para o PDE-PPV de 14 Barras.

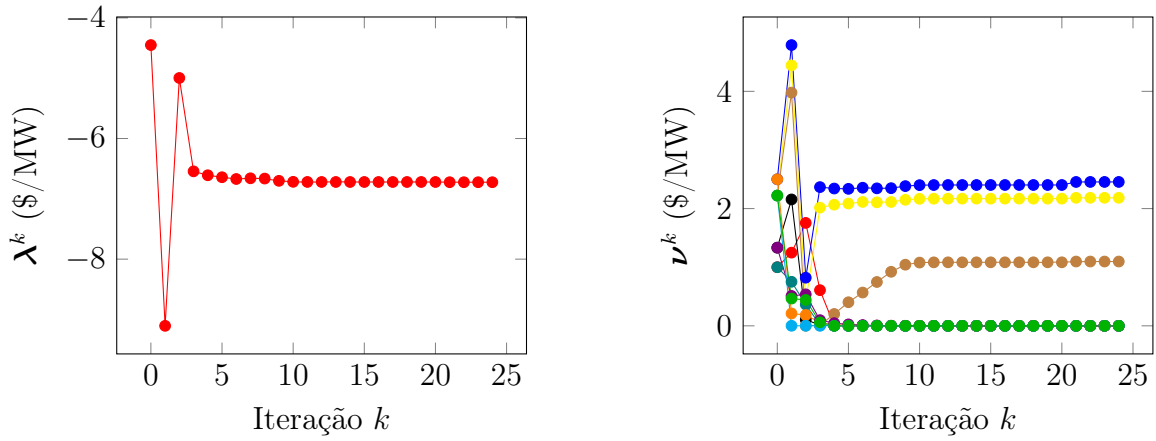


Figura 6.25: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PPV de 14 Barras.

Tabela 6.29: Resultados – PDE-PPV: 14 Barras

Potência Gerada (MW)	Métodos	
	IFEP (Ravi et al., 2006)	PDPCBLM
$P_{G,1}$	96,28	96,0739
$P_{G,2}$	50	50
$P_{G,3}$	56,39	56,6433
$P_{G,6}$	45	45
$P_{G,8}$	44,99	45
Perdas L (MW)	3,66	3,7172
Função Objetivo (\$)	1064,27	1070,56

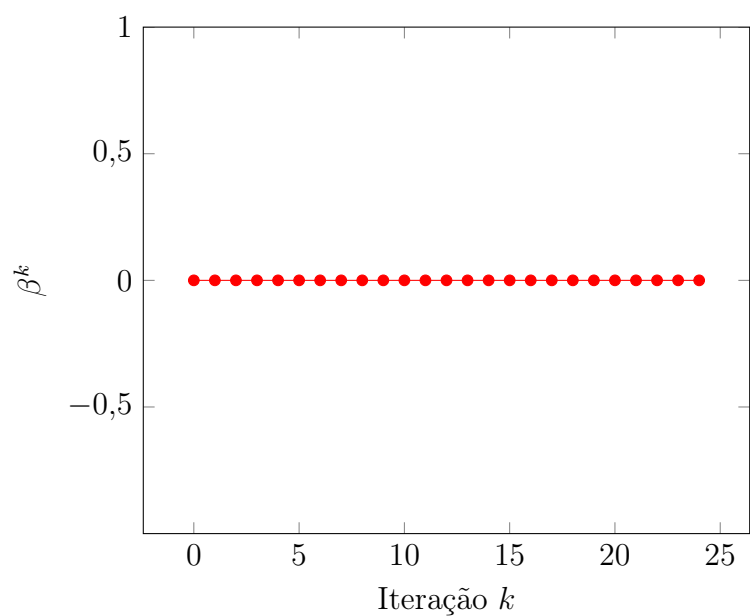


Figura 6.26: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PPV de 14 Barras.

Tabela 6.30: Magnitudes e Ângulos de Tensão – PDE-PPV: 14 Barras

Barra	V	θ (graus)
1	1,06	0
2	1,045	-1,7976297
3	1,01	-5,0501767
4	1,0254816	-4,8321096
5	1,0285438	-3,8673193
6	1,07	-4,9614839
7	1,063556	-5,9663423
8	1,09	-4,6603299
9	1,0555513	-7,4013548
10	1,0502495	-7,2589876
11	1,0560701	-6,2539031
12	1,0553895	-5,9358079
13	1,0499288	-6,1402754
14	1,0349325	-7,8559872

Tabela 6.31: Características do sistema – PDE-PPV: 30 Barras

Gerador (k)	$\frac{P_{G,k}}{(\text{MW})}$	$\overline{P_{G,k}}$ (MW)	a_k (\$/MW ²)	b_k (\$/MW)	c_k (\$)	e_k (\$)	f_k (1/MW)
1	50	200	0,00375	2	0	120	0,073
2	20	80	0,0175	1,75	0	50	0,032
5	15	50	0,0625	1	0	30	0,051
8	10	35	0,00834	3,25	0	25	0,026
11	10	30	0,025	3	0	25	0,026
13	12	40	0,025	3	0	30	0,048
Demanda (soma das demandas nas barras): $D = 283,4$ MW							

Fonte: Ravi et al. (2006)

Tabela 6.32: Inicialização – PDE-PPV: 30 Barras

Fatores de aumento/redução da correção de inércia	$\kappa_1 = \kappa_2 = 10$	
Fatores de redução do parâmetro de suavização	$\rho_1 = 2$ e $\rho_2 = 0$	
Limitante da atualização do parâmetro de suavização	$\bar{\eta} = 10^{-5}$	
Parâmetro de barreira/relaxação inicial	$\mu^0 = 5$	
Fatores de atualização do parâmetro de barreira	$\tau_1 = \tau_2 = 0,5$	
Fator de controle da escolha entre passo previsor ou corretor	$\chi = 1$	
Fator de redução do passo	$\sigma = 0,9995$	
Precisão de parada	$\varepsilon = 10^{-6}$	
Potência inicial das unidades geradoras (MW)	1	125
	2	50
	5	32,5
	8	22,5
	11	20
	13	26
Estimadores dos multiplicadores de Lagrange	$\delta^0 = 10 \cdot \mathbb{1}_p$	

alteradas para ficarem iguais ao ponto inicial do método PDPCBLM, a fim de evitar divisão por 0 no algoritmo de cálculo dos coeficientes-B. Os coeficientes de custo das unidades geradoras são apresentados na Tabela 6.31. A formulação do modelo do PDE-PPV para este caso possui:

- 6 variáveis;
- 1 restrição de igualdade;
- 12 restrições de desigualdade.

As gerações iniciais foram tomadas como sendo o valor médio entre o mínimo e o máximo da saída de potência das unidades geradoras. Os parâmetros de inicialização são apresentados na 6.32. A Tabela 6.33 apresenta os resultados obtidos. Neste caso, o método convergiu após 29 iterações, ao longo de 6 ciclos externos. A diferença relativa entre a função objetivo original e aproximada, na solução ótima determinada, é de 0,040000394%.

Tabela 6.33: Resultados (Iterações) – PDE-PPV: 30 Barras

Iteração k	$f(x^k)$	$f_{\eta^k}(x^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(x^k, s^k, \lambda^k, \nu^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	964,77584	1118,7395	153,96364	15,958488	39,065434	1	5	0	0	–
1	956,06846	1135,6225	179,55399	18,780453	34,656631	1	5	0,9995	0,9995	–
2	943,88671	1055,8455	111,95881	11,861467	6,0923931	0,5	2,5	0,9995	0,9995	+
3	912,4479	984,51766	72,069756	7,8985065	5,4340472	0,25	1,25	0,9995	0,9995	+
4	999,6616	1027,3071	27,645488	2,7654847	2,590805	0,125	0,625	0,9995	0,36222123	–
5	986,37773	1001,9154	15,537666	1,5752247	3,7801321	0,0625	0,3125	0,9995	0,9995	+
6	938,67548	948,69397	10,018492	1,0673009	6,6009588	0,03125	0,15625	0,9995	0,9995	+
7	860,34269	876,29914	15,956447	1,8546617	4,1036118	0,015625	0,078125	0,9995	0,9995	+
8	855,26163	868,35197	13,090343	1,5305659	0,25807088	0,0078125	0,0390625	0,9995	0,9995	+
9	852,94659	861,94737	9,0007825	1,0552575	0,32246927	0,00390625	0,01953125	0,9995	0,9995	+
10	852,91475	858,76744	5,8526891	0,6861986	0,7956353	0,001953125	0,009765625	0,029482042	0,9995	+
11	852,42289	856,67402	4,2511291	0,49871128	0,32365254	0,0009765625	0,0048828125	0,97115381	0,9995	–
12	852,3656	855,25609	2,8904868	0,33911349	0,34086378	0,00048828125	0,60348961	0,9995	0,9995	+
13	852,24023	854,26497	2,0247407	0,23757863	0,34225998	0,00024414063	0,30174481	0,9995	0,9995	+
14	852,16194	853,57261	1,4106682	0,16553992	0,34303979	0,00012207031	0,1508724	0,9995	0,9995	+
15	852,09887	853,09037	0,9915049	0,11636031	0,3435482	6,1035156 $\times 10^{-5}$	0,075436202	0,9995	0,9995	+
16	852,05682	852,75301	0,69618495	0,081706399	0,34382836	3,0517578 $\times 10^{-5}$	0,037718101	0,9995	0,9995	+
17	852,02583	852,51645	0,49062582	0,057583445	0,34401008	1,5258789 $\times 10^{-5}$	0,01885905	0,9995	0,9995	+
18	852,00437	852,35012	0,34575267	0,040581091	0,036016962	7,6293945 $\times 10^{-6}$	0,0094295252	0,9995	0,9995	+
19	852,00596	852,35008	0,34411585	0,040388902	0,0011687617	7,6293945 $\times 10^{-6}$	0,0047147626	0,9995	0,9995	–
20	852,00602	852,35008	0,34406359	0,040382765	1,9215811 $\times 10^{-6}$	7,6293945 $\times 10^{-6}$	0,0023573813	0,9995	0,9995	+
21	852,00602	852,35008	0,3440635	0,040382755	9,6444025 $\times 10^{-10}$	7,6293945 $\times 10^{-6}$	0,0011786907	0,9995	0,9995	+
22	852,93166	853,27306	0,34139045	0,040025533	0,0048078917	7,6293945 $\times 10^{-6}$	0,00058934533	0,9995	0,88728496	–
23	852,93377	853,27514	0,34136953	0,040022982	2,3915459 $\times 10^{-6}$	7,6293945 $\times 10^{-6}$	0,00029467266	0,9995	0,9995	+
24	852,93377	853,27514	0,34136949	0,040022978	1,3567688 $\times 10^{-9}$	7,6293945 $\times 10^{-6}$	0,00014733633	0,9995	0,9995	+
25	852,98657	853,32779	0,34121571	0,040002472	1,7568237 $\times 10^{-5}$	7,6293945 $\times 10^{-6}$	7,3668166 $\times 10^{-5}$	0,9995	0,99173017	–
26	852,98661	853,32782	0,34121559	0,040002455	8,7831653 $\times 10^{-9}$	7,6293945 $\times 10^{-6}$	3,6834083 $\times 10^{-5}$	0,9995	0,9995	+
27	852,99076	853,33196	0,34120135	0,040000591	4,6167804 $\times 10^{-7}$	7,6293945 $\times 10^{-6}$	1,8417041 $\times 10^{-5}$	0,9995	0,99889414	–
28	852,99109	853,33229	0,34119996	0,040000413	3,5645144 $\times 10^{-8}$	7,6293945 $\times 10^{-6}$	9,2085207 $\times 10^{-6}$	0,9995	0,9995	+
29	852,99112	853,33232	0,34119981	0,040000394	3,1606646 $\times 10^{-9}$	7,6293945 $\times 10^{-6}$	4,6042604 $\times 10^{-6}$	0,9995	0,9995	+

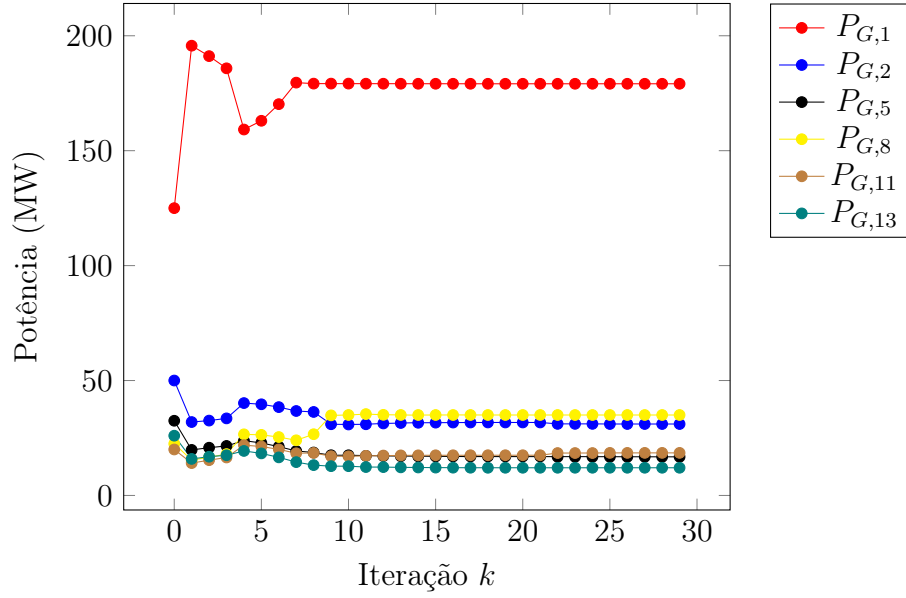


Figura 6.27: Sequência de potências geradas pelo método PDPCBLM para o PDE-PPV de 30 Barras.

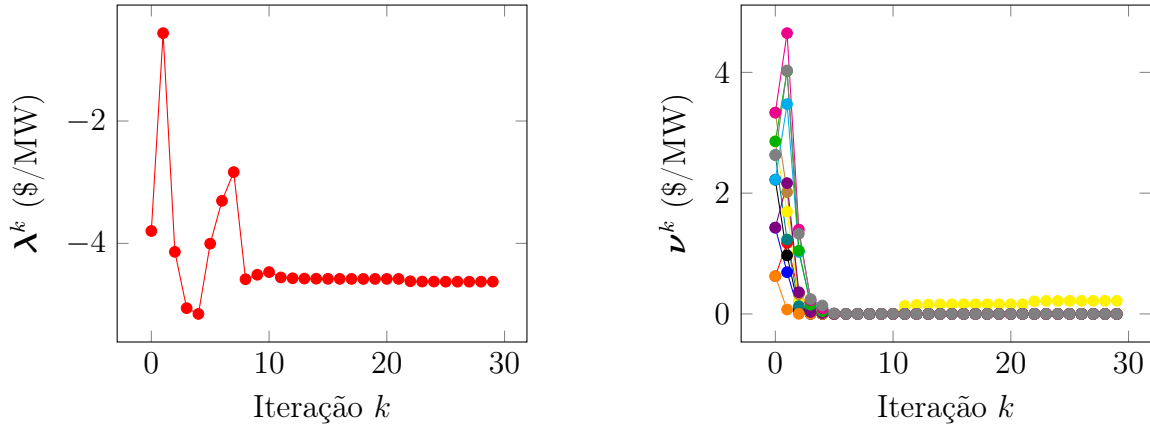


Figura 6.28: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PPV de 30 Barras.

Nas Figuras 6.27, 6.28 e 6.29 podemos observar a sequência de potências geradas, multiplicadores de Lagrange e parâmetro de correção de inércia, respectivamente. Neste caso, o gerador da barra 8 está gerando sua potência máxima na solução ótima encontrada, o que justifica a existência de um multiplicador ν_j que não está tendendo a 0 na solução ótima. Cabe salientar que, neste caso, o método PDPCBLM conseguiu determinar uma solução melhor que a apresentada por Ravi et al. (2006), conforme indica a Tabela 6.34. Por fim, a Tabela 6.35 apresenta as magnitudes e ângulos de tensão em cada barra, no último fluxo de potência executado.

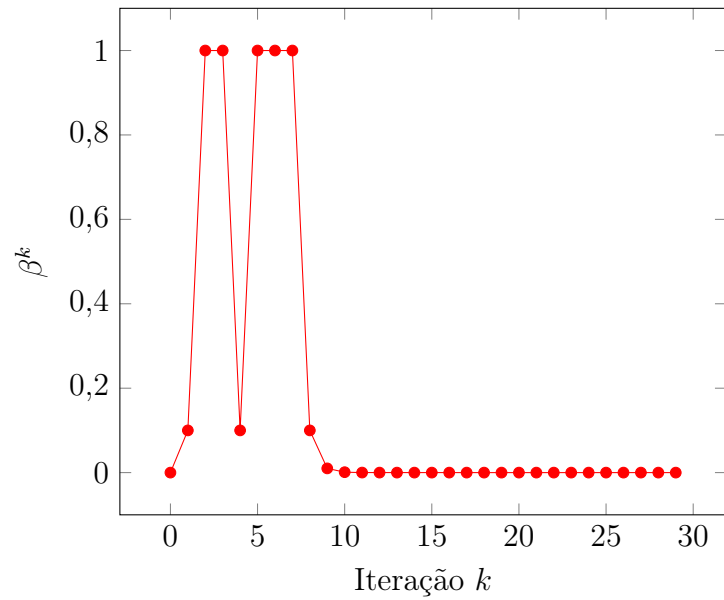


Figura 6.29: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PPV de 30 Barras.

Tabela 6.34: Resultados – PDE-PPV: 30 Barras

Potência Gerada (MW)	Métodos	
	IFEP (Ravi et al., 2006)	PDPCBLM
$P_{G,1}$	136,11	179,1104
$P_{G,2}$	64	31,082
$P_{G,5}$	27,53	16,8028
$P_{G,8}$	23,14	35
$P_{G,11}$	16,80	18,5226
$P_{G,13}$	23,32	12,0462
Perdas L (MW)	7,49	9,1641
Função Objetivo (\$)	868,03	852,9911

Tabela 6.35: Magnitudes e Ângulos de Tensão – PDE-PPV: 30 Barras

Barra	V	θ (graus)
1	1,06	0
2	1,045	-3,6748744
3	1,0260585	-4,9045768
4	1,0175714	-6,0106255
5	1,01	-10,245512
6	1,0141101	-7,0137254
7	1,004684	-8,8655974
8	1,01	-6,9217869
9	1,0537672	-8,6469167
10	1,0488773	-10,557223
11	1,082	-6,7104989
12	1,060151	-9,9686714
13	1,071	-9,1176085
14	1,0453852	-10,846284
15	1,0408937	-10,928985
16	1,0478766	-10,479711
17	1,0435515	-10,746999
18	1,0316081	-11,488131
19	1,0292347	-11,630031
20	1,0333703	-11,418717
21	1,0364614	-11,027654
22	1,0369691	-11,023444
23	1,0304072	-11,328567
24	1,0248042	-11,519377
25	1,0197181	-11,351736
26	1,0020837	-11,769402
27	1,025157	-10,990686
28	1,0097908	-7,4112508
29	1,005359	-12,215962
30	0,99390794	-13,09533

6.3 PDE-PV-RT

Nesta seção, apresentamos os resultados obtidos ao resolver o PDE-PV-RT, para os sistemas de 6 barras de Wood e Wollenberg, IEEE-14 barras e IEEE-30 barras. O objetivo é comparar os resultados da resolução através das perdas calculadas de forma aproximada pelos coeficientes-B, com a resolução do modelo em que as equações de balanço de potência ativa são consideradas explicitamente. Para o PDE-PV-RT, foram consideradas duas situações em cada caso: sem sobrecarga e com sobrecarga em uma linha de transmissão.

6.3.1 Sistema 6 Barras de Wood e Wollenberg – Parte I

Como o modelo do PDE-PV-RT supõe conhecidas as magnitudes de tensão, para que uma comparação justa fosse realizada as magnitudes de tensão foram fixadas no valor do último fluxo de carga executado no caso PDE-PPV de 6 Barras. Além disso, como neste problema existem restrições de balanço de potência, as unidades utilizadas foram pu, na base de 100 MVA. Também, a função objetivo foi rescalada por um fator de 1000, isto é, dividiu-se a função objetivo por 1000 para a resolução do problema. Isto não acarreta diferenças na solução obtida, e somente foi feito para que não houvesse grandes diferenças entre as magnitudes das variáveis e o valor da função objetivo. Além disso, o valor do ângulo de tensão na barra *slack* foi fixado em 0. Os coeficientes da função objetivo são os mesmos apresentados na Tabela 6.21, devidamente convertidos para pu.

O objetivo desta seção é resolver o PDE-PV-RT, quando as linhas de transmissão estão descongestionadas, a fim de verificar se a solução do PDE-PPV de 6 Barras é coerente. Para tanto assumiu-se que o limite para fluxo nas linhas é de 1 pu. Com isto, o sistema de transmissão irá operar com folga, sem que haja congestionamento nas linhas de transmissão. Este valor foi escolhido porque, ao executar o último fluxo de carga no PDE-PPV de 6 Barras, constatou-se que o maior fluxo de potência ativa era de aproximadamente 0,6 pu. A formulação do modelo do PDE-PV-RT para este caso possui:

- 8 variáveis;
- 6 restrições de igualdade;
- 28 restrições de desigualdade.

Para inicialização do método, utilizou-se os parâmetros na Tabela 6.36. O método convergiu em 8 iterações, e os resultados estão resumidos na Tabela 6.37.

As Figuras 6.30, 6.31, 6.32 e 6.33 a seguir apresentam a trajetória das potências geradas, dos ângulos de tensão, dos multiplicadores de Lagrange e do parâmetro de correção de

Tabela 6.36: Inicialização – PDE-PV-RT: 6 Barras (Parte I)

Fatores de aumento/redução da correção de inércia	$\kappa_1 = \kappa_2 = 10$	
Fatores de redução do parâmetro de suavização	$\rho_1 = 3$ e $\rho_2 = 0$	
Limitante da atualização do parâmetro de suavização	$\bar{\eta} = 10^{-5}$	
Parâmetro de barreira/relaxação inicial	$\mu^0 = 1$	
Fatores de atualização do parâmetro de barreira	$\tau_1 = \tau_2 = 0,1$	
Fator de controle da escolha entre passo previsor ou corretor	$\chi = 1$	
Fator de redução do passo	$\sigma = 0,9995$	
Precisão de parada	$\varepsilon = 10^{-6}$	
Potência inicial das unidades geradoras (pu)	1	1,25
	2	0,9375
	3	1,125
Vetor de ângulos nas barras	$\theta = \mathbf{0}$	
Estimadores dos multiplicadores de Lagrange	$\delta^0 = \mathbb{1}_p$	

inércia, ao longo do processo iterativo. Um ponto interessante a se notar é que a solução obtida, convertida em MW, convergiu para a mesma solução do caso PDE-PPV de 6 Barras, conforme indica a Tabela 6.38. Além disso, a diferença relativa entre a função objetivo original e aproximada, calculadas na solução determinada, foi de 0,016050064%.

6.3.2 Sistema 6 Barras de Wood e Wollenberg – Parte II

Da mesma forma que na seção anterior, o PDE-PV-RT foi resolvido, considerando as magnitudes de tensão obtidas no fluxo de potência final do PDE-PPV de 6 barras, apresentadas na Tabela 6.25. O objetivo desta seção é investigar o que ocorre quando há sobrecarga em alguma linha de transmissão. Dessa forma, reduziu-se o fluxo máximo de potência ativa permitido na linha 3–6, de 1 pu para 0,5 pu, uma vez que na solução encontrada previamente, na Parte I, o fluxo era de aproximadamente 0,6 pu na referida linha. Esta linha foi escolhida porque era a que apresentava maiores valores para o fluxo de potência. Para inicialização do método, utilizou-se os mesmos parâmetros da Parte I (Tabela 6.36). O método convergiu em 8 iterações, e os resultados estão resumidos na Tabela 6.39.

As Figuras 6.34, 6.35, 6.36 e 6.37 a seguir apresentam a trajetória das potências geradas, dos ângulos de tensão, dos multiplicadores de Lagrange e do parâmetro de correção de inércia, ao longo do processo iterativo. Neste caso, a diferença relativa entre a função original e aproximada, na solução ótima, foi de 0,018483127%. Deve-se notar que, neste caso, o método encontrou outra solução, de custo maior. Porém, para essa nova solução, nenhuma linha se encontra congestionada, isto é, o fluxo na linha 3–6 na solução determinada não se encontra no limite máximo. A Tabela 6.40 apresenta a solução ótima determinada, em que os ângulos de tensão estão expressos em graus.

Tabela 6.37: Resultados (Iterações) – PDE-PV-RT: 6 Barras (Parte I)

Iteração k	$f(\mathbf{x}^k)$	$f_{\eta^k}(\mathbf{x}^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	4,8110313	4,8139361	0,0029047961	0,060377824	1,1275092	0,01	1	0	0	–
1	3,1165463	3,1182646	0,0017183234	0,055135499	3,8374375	0,01	1	0,9995	0,9995	–
2	3,1064892	3,1100869	0,0035977527	0,11581411	1,1249452	0,0033333333	0,1	0,9995	0,9995	–
3	3,0079241	3,0100772	0,0021530235	0,071578383	0,20830573	0,0011111111	0,01	0,9995	0,89196561	+
4	3,0038638	3,0060995	0,0022356923	0,074427219	0,17878848	0,00037037037	0,001	0,86687473	0,99876478	+
5	3,0037792	3,0051422	0,0013630149	0,045376669	0,035765876	0,00012345679	0,0001	0,9995	0,9995	+
6	3,0036507	3,0044835	0,00083282573	0,027727117	0,0016544424	$4,1152263 \times 10^{-5}$	1×10^{-5}	0,9995	0,9995	+
7	3,0036459	3,0041279	0,00048207511	0,016049665	$2,9866947 \times 10^{-5}$	$1,3717421 \times 10^{-5}$	1×10^{-6}	0,9995	0,9995	+
8	3,0036458	3,0041279	0,00048208707	0,016050064	$1,4933597 \times 10^{-8}$	$4,5724737 \times 10^{-6}$	1×10^{-7}	0,9995	0,9995	–

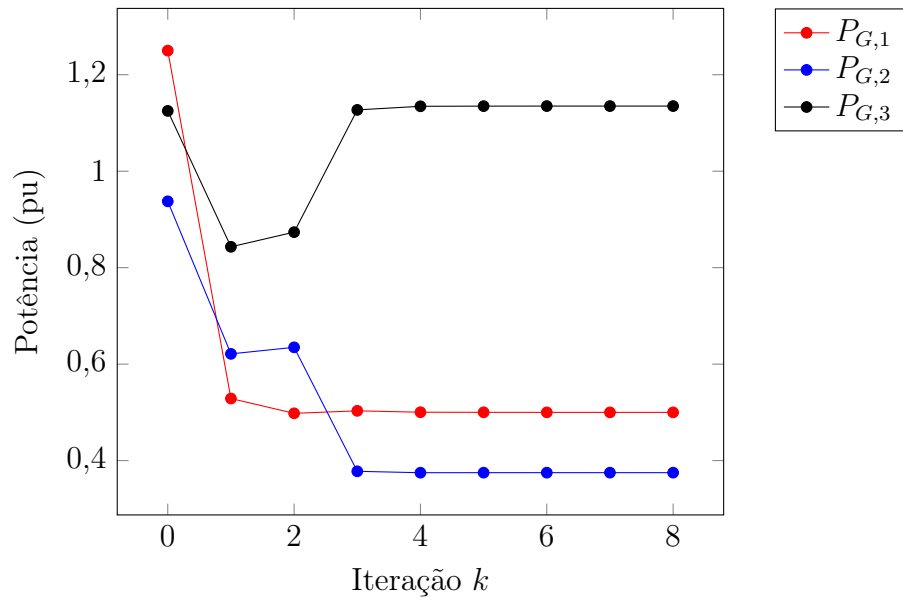


Figura 6.30: Sequência de potências geradas pelo método PDPCBLM para o PDE-PV-RT de 6 Barras (Parte I).

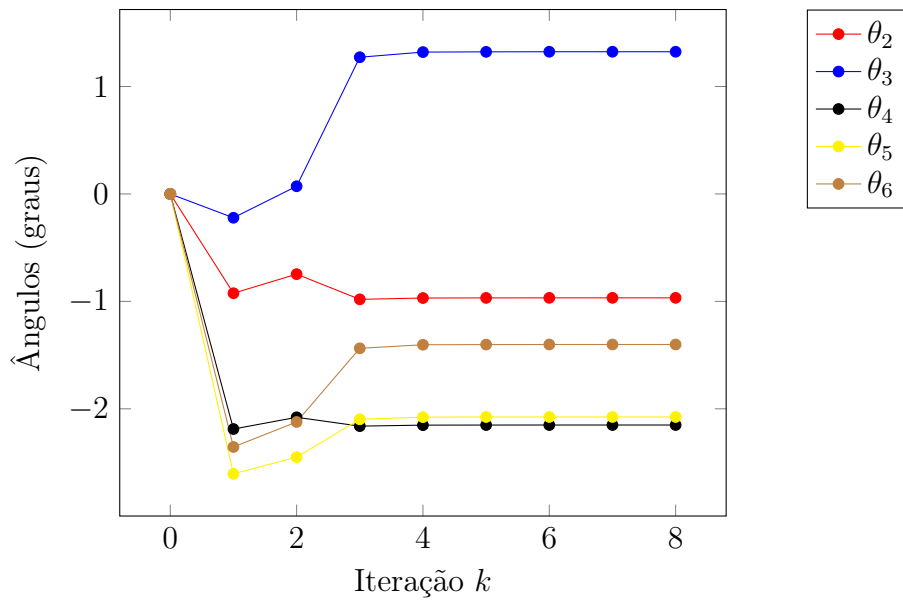


Figura 6.31: Sequência de ângulos de tensão nas barras (diferentes da *slack*) geradas pelo método PDPCBLM para o PDE-PV-RT de 6 Barras (Parte I).

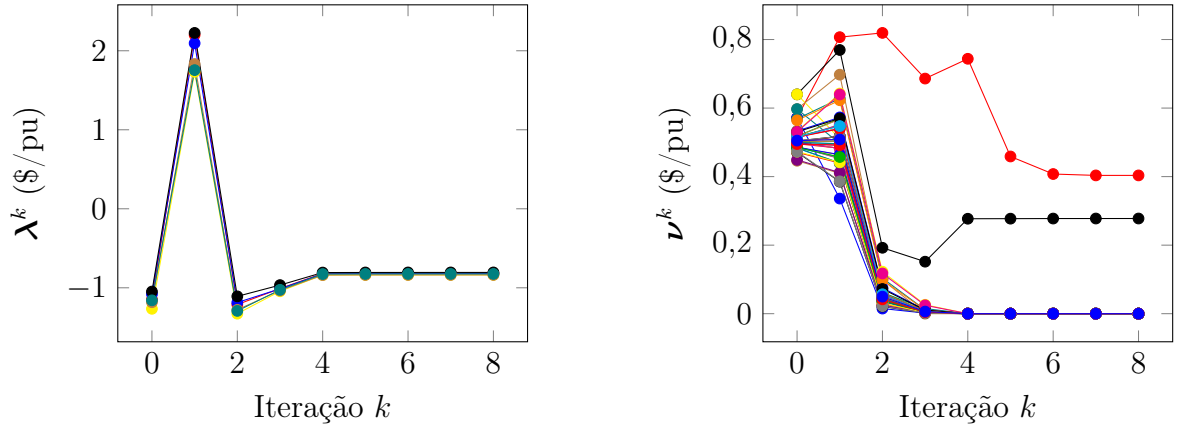


Figura 6.32: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PV-RT de 6 Barras (Parte I).

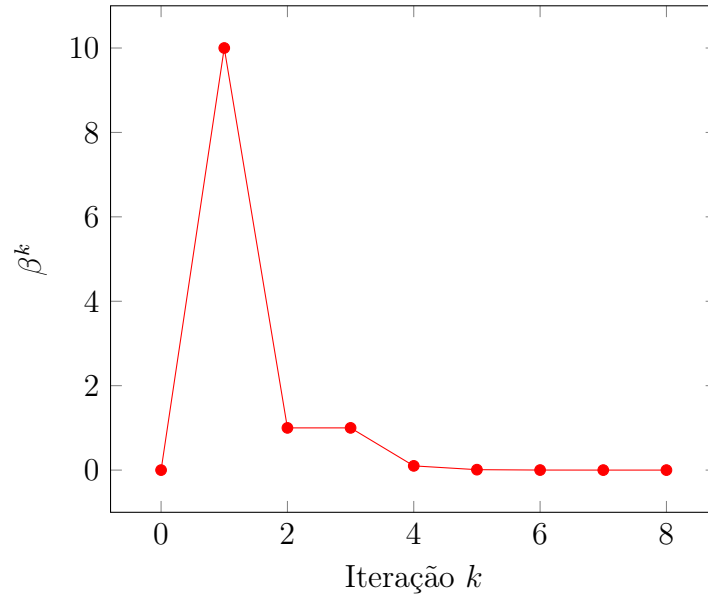


Figura 6.33: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PV-RT de 6 Barras (Parte I).

Tabela 6.38: Resultados – PDE-PV-RT: 6 Barras (Parte I)

Variável	PDE-PPV	PDE-PV-RT
$P_{G,1}$	50	50
$P_{G,2}$	37,5	37,5
$P_{G,3}$	113,5015	113,50149
θ_1	0	0
θ_2	-0,9670899	-0,9670899
θ_3	1,3237134	1,3237134
θ_4	-2,1507633	-2,1507633
θ_5	-2,075729	-2,0757289
θ_6	-1,401307	-1,401307
Função Objetivo (\$)	3003,64	3003,64

Tabela 6.39: Resultados (Iterações) – PDE-PV-RT: 6 Barras (Parte II)

Iteração k	$f(\mathbf{x}^k)$	$f_{\eta^k}(\mathbf{x}^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	4,8110313	4,8139361	0,0029047961	0,060377824	1,1275092	0,01	1	0	0	–
1	3,1228952	3,1243863	0,0014910532	0,047745859	3,8191543	0,01	1	0,995	0,995	–
2	3,1342174	3,1348548	0,00063741973	0,020337445	0,20137909	0,0033333333	0,1	0,995	0,995	–
3	3,1106495	3,1127869	0,0021374114	0,068712704	1,1747561	0,0011111111	0,01	0,74649851	0,84915438	–
4	3,0065985	3,0092504	0,0026519375	0,088203914	0,74479102	0,00037037037	0,001	0,32961486	0,47271086	+
5	3,0092696	3,0107022	0,0014325398	0,047604234	0,072342101	0,00012345679	0,00027878055	0,9995	0,9995	+
6	3,0090309	3,0099527	0,00092183039	0,030635458	0,015429277	4,1152263 $\times 10^{-5}$	2,7878055 $\times 10^{-5}$	0,9995	0,99949204	+
7	3,0089866	3,0095398	0,00055318475	0,018384421	0,0020799693	1,3717421 $\times 10^{-5}$	2,7878055 $\times 10^{-6}$	0,9995	0,99949539	+
8	3,0089833	3,0095394	0,00055615418	0,018483127	8,9085802 $\times 10^{-7}$	4,5724737 $\times 10^{-6}$	2,7878055 $\times 10^{-7}$	0,9995	0,9995	+

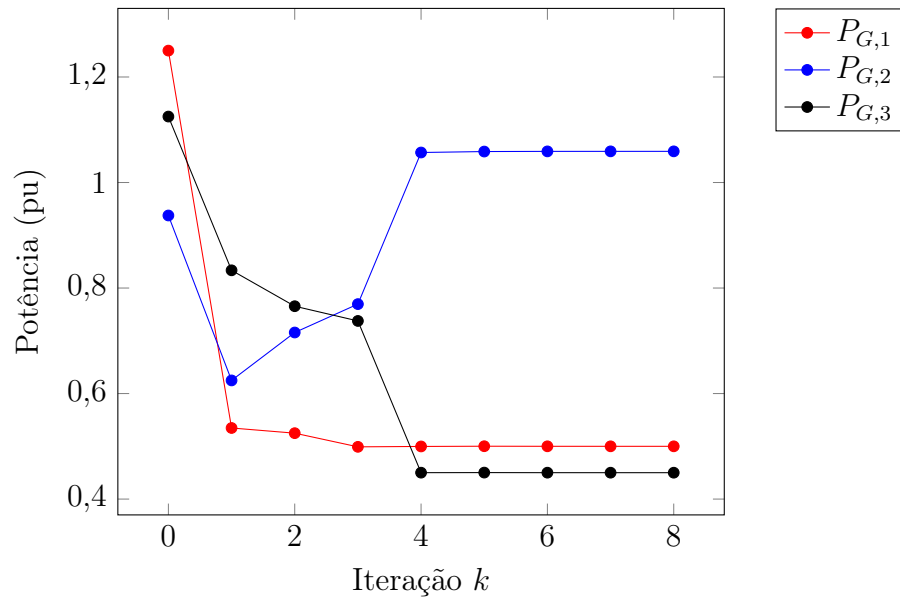


Figura 6.34: Sequência de potências geradas pelo método PDPCBLM para o PDE-PV-RT de 6 Barras (Parte II).

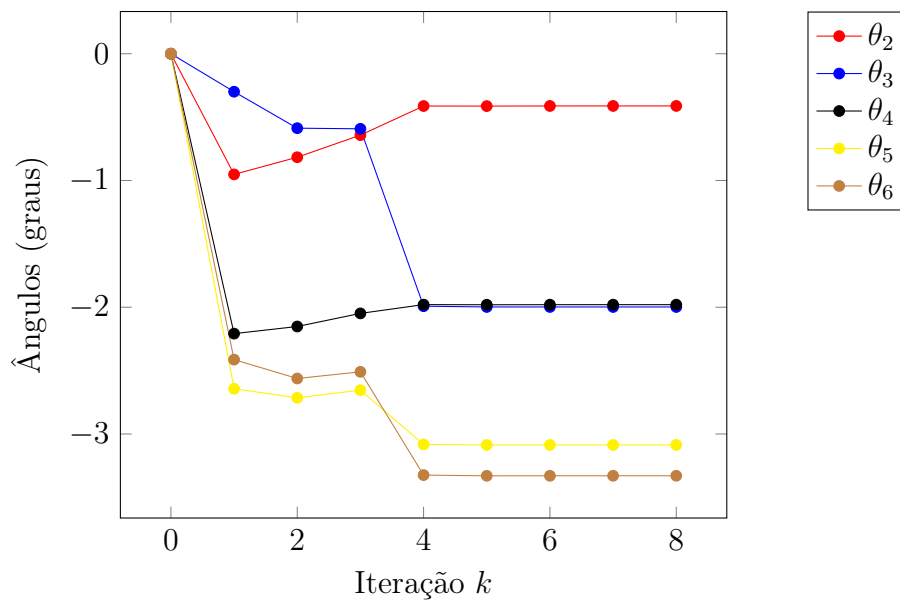


Figura 6.35: Sequência de ângulos de tensão nas barras (diferentes da *slack*) geradas pelo método PDPCBLM para o PDE-PV-RT de 6 Barras (Parte II).

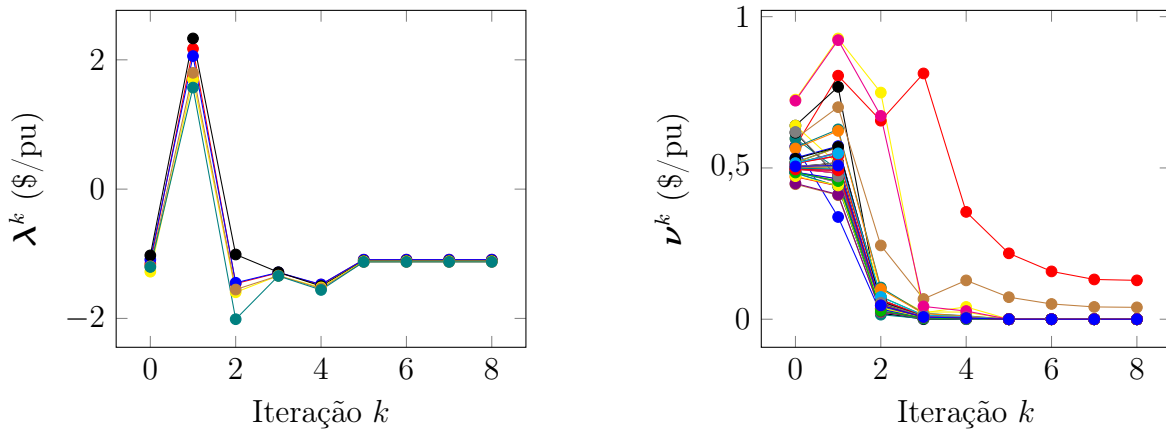


Figura 6.36: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PV-RT de 6 Barras (Parte II).

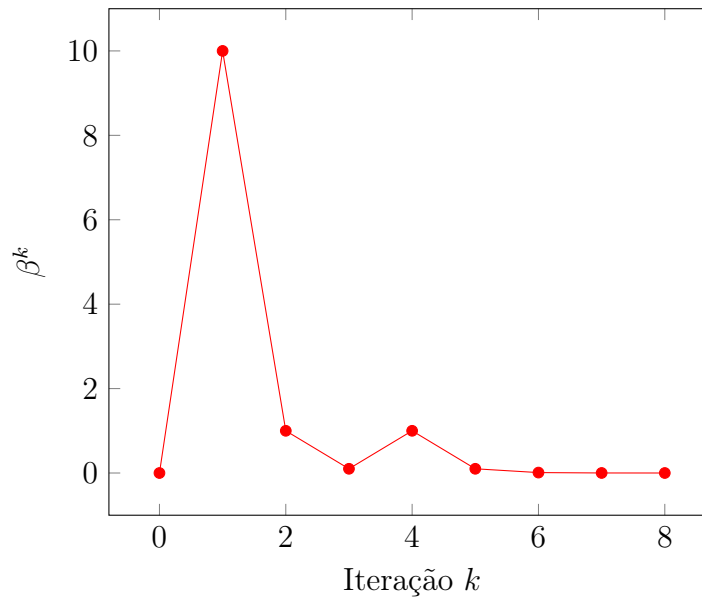


Figura 6.37: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PV-RT de 6 Barras (Parte II).

Tabela 6.40: Resultados – PDE-PV-RT: 6 Barras (Parte II)

Variável	PDE-PPV	PDE-PV-RT (Parte I)	PDE-PV-RT (Parte II)
$P_{G,1}$	50	50	50
$P_{G,2}$	37,5	37,5	105,90693
$P_{G,3}$	113,5015	113,50149	45
θ_1	0	0	0
θ_2	-0,9670899	-0,9670899	-0,41191119
θ_3	1,3237134	1,3237134	-1,9979328
θ_4	-2,1507633	-2,1507633	-1,9798558
θ_5	-2,075729	-2,0757289	-3,086489
θ_6	-1,401307	-1,401307	-3,329801
Função Objetivo (\$)	3003,64	3003,64	3008,98

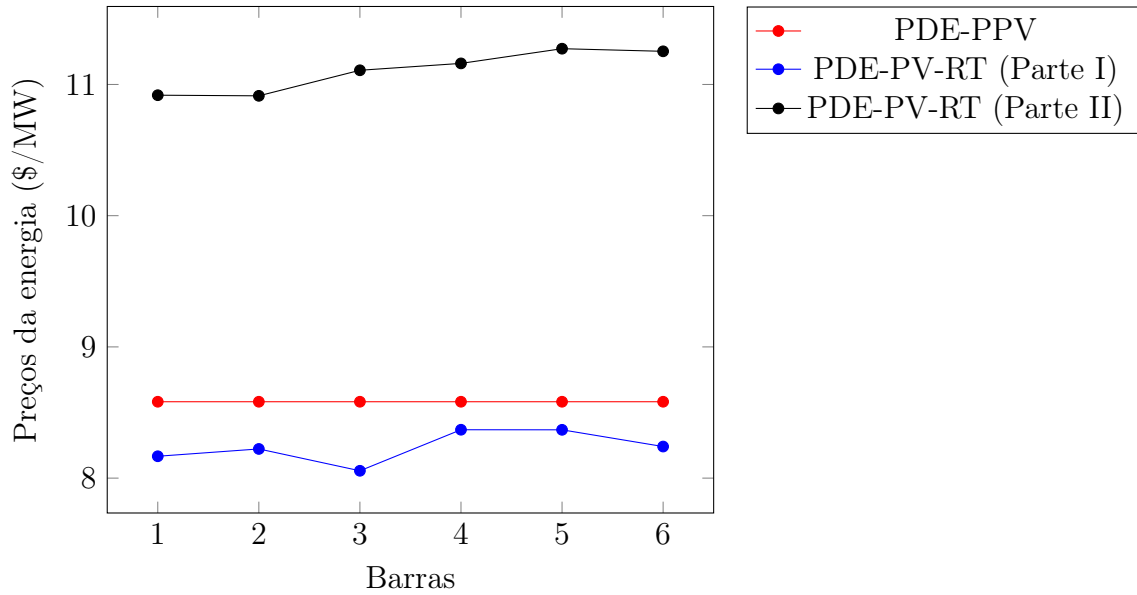


Figura 6.38: Comparação entre os preços de energia obtidos ao resolver o PDE-PPV, PDE-PV-RT (Parte I) e PDE-PV-RT (Parte II) para o sistema 6 Barras.

Nota-se que, no caso em que não há congestionamento nas linhas, os preços nodais de energia obtidos pelo PDE-PV-RT são semelhantes, porém inferiores, ao preço “médio” obtido pelo PDE-PPV. Quando as linhas encontram-se em uma situação de sobrecarga, os preços da energia aumentam em todas as barras. Apesar da solução do caso congestionado apresentar valor da função objetivo próximo ao do caso descongestionado, os preços variam significativamente, conforme indica a Figura 6.38. Assim, a solução do PDE-PV-RT se mostra mais conveniente do ponto de vista econômico, ao menos do lado da demanda, já que os preços podem ser menores quando não há sobrecargas nos ramos do sistema.

Apesar de possuir um número maior de restrições e variáveis, a convergência deste caso do PDE-PV-RT também se mostrou mais vantajosa, em termos de número de iterações. A isto, acrescenta-se o fato de que não é necessária a utilização de um fluxo de potência externo ao método de pontos interiores para calcular coeficientes-B, já que as perdas ficam implícitas no modelo. Em contrapartida, o PDE-PV-RT requer que as magnitudes de tensão sejam previamente conhecidas. Outra conclusão, já esperada, é o fato de que as barras que possuem gerador possuem preços de energia menores que as demais.

6.3.3 Sistema IEEE-14 Barras – Parte I

Nesta seção e na próxima é feita uma análise semelhante ao caso de 6 barras do PDE-PV-RT, para o sistema IEEE-14 barras. Começa-se investigando a convergência do método PDPCBLM no caso em que as linhas de transmissão estão descongestionadas. As magnitudes de tensão foram fixadas nos valores apresentadas na Tabela 6.30. Os coeficientes da

Tabela 6.41: Inicialização – PDE-PV-RT: 14 Barras (Parte I)

Fatores de aumento/redução da correção de inércia	$\kappa_1 = \kappa_2 = 10$	
Fatores de redução do parâmetro de suavização	$\rho_1 = 3$ e $\rho_2 = 0$	
Limitante da atualização do parâmetro de suavização	$\bar{\eta} = 10^{-7}$	
Parâmetro de barreira/relaxação inicial	$\mu^0 = 1$	
Fatores de atualização do parâmetro de barreira	$\tau_1 = \tau_2 = 0,4$	
Fator de controle da escolha entre passo previsor ou corretor	$\chi = 1$	
Fator de redução do passo	$\sigma = 0,995$	
Precisão de parada	$\varepsilon = 10^{-6}$	
	1	0,55
	2	0,35
Potência inicial das unidades geradoras (pu)	3	0,475
	6	0,275
	8	0,275
Vetor de ângulos nas barras		$\boldsymbol{\theta} = \mathbf{0}$
Estimadores dos multiplicadores de Lagrange		$\boldsymbol{\delta}^0 = 100 \cdot \mathbb{1}_p$

função custo estão apresentados na Tabela 6.26, e foram convertidos para pu. Novamente, a função objetivo foi reescalada por um fator de 1000.

Assume-se que o limite para fluxo nas linhas de transmissão é de 1 pu, visando evitar congestionamentos. Esta escolha, assim como no caso de 6 barras, foi realizada considerando o maior fluxo determinado ao resolver o PDE-PPV, que era de aproximadamente 0,6 pu, na linha 1–2. A formulação do modelo do PDE-PV-RT para este caso possui:

- 18 variáveis;
- 14 restrições de igualdade;
- 50 restrições de desigualdade.

Para inicialização do método, utilizou-se os parâmetros na Tabela 6.41. Nota-se que, ao invés de utilizar $\sigma = 0,9995$, utilizou-se $\sigma = 0,995$. Isto porque o método ficava próximo da solução com o fator de passo $\sigma = 0,9995$, mas apresentava dificuldades de convergência no final do processo iterativo (convergia com uma precisão $\bar{\varepsilon} = 10^{-5}$, mas não $\bar{\varepsilon} = 10^{-6}$). Desta forma, o método convergiu em 15 iterações, e os resultados estão resumidos na Tabela 6.42.

As Figuras 6.39, 6.40, 6.41 e 6.42 a seguir apresentam a trajetória das potências geradas, dos ângulos de tensão, dos multiplicadores de Lagrange e do parâmetro de correção de inércia, ao longo do processo iterativo. Novamente, observa-se que a solução obtida, convertida em MW, convergiu para uma solução próxima à obtida pelo modelo do PDE-PPV de 14 Barras, conforme indica a Tabela 6.43. Além disso, a diferença relativa entre a função objetivo original e aproximada, calculadas na solução determinada, foi

Tabela 6.42: Resultados (Iterações) – PDE-PV-RT: 14 Barras (Parte I)

Iteração k	$f(\mathbf{x}^k)$	$f_{\eta^k}(\mathbf{x}^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	0,68670723	0,69036127	0,003654043	0,53211075	8,2427568	0,01	1	0	0	–
1	1,287355	1,2889106	0,001555549	0,12083295	54,742734	0,01	1	0,995	0,995	–
2	1,276467	1,2767251	0,00025807845	0,020218184	18,412292	0,00333333333	0,4	0,995	0,995	–
3	1,1165525	1,1167165	0,00016394097	0,014682781	4,8681394	0,0011111111	0,16	0,995	0,995	+
4	1,119695	1,1197581	6,3156991 $\times 10^{-5}$	0,0056405534	2,4079386	0,00037037037	0,064	0,995	0,995	+
5	1,114352	1,1144227	7,0709868 $\times 10^{-5}$	0,0063453798	0,69219596	0,00012345679	0,0256	0,995	0,995	+
6	1,1127003	1,1129603	0,0002599894	0,023366485	0,62176063	4,1152263 $\times 10^{-5}$	0,01024	0,995	0,995	–
7	1,1074655	1,1075531	8,7570465 $\times 10^{-5}$	0,0079072857	0,089810394	1,3717421 $\times 10^{-5}$	0,004096	0,995	0,94639382	+
8	1,0771715	1,0772071	3,5553793 $\times 10^{-5}$	0,0033006621	0,11742307	4,5724737 $\times 10^{-6}$	0,0016384	0,995	0,97452811	+
9	1,0704833	1,0705045	2,114621 $\times 10^{-5}$	0,001975389	0,085431941	1,5241579 $\times 10^{-6}$	0,00065536	0,92610082	0,82005435	–
10	1,0705635	1,0705752	1,1728266 $\times 10^{-5}$	0,0010955227	0,083590234	5,0805263 $\times 10^{-7}$	0,000262144	0,995	0,995	+
11	1,0705642	1,0705713	7,1219951 $\times 10^{-6}$	0,00066525625	0,071136182	1,6935088 $\times 10^{-7}$	0,0001048576	0,995	0,995	+
12	1,0705612	1,070571	9,738113 $\times 10^{-6}$	0,0009096269	0,027894538	5,6450293 $\times 10^{-8}$	4,194304 $\times 10^{-5}$	0,995	0,995	–
13	1,0705619	1,0705709	9,0740255 $\times 10^{-6}$	0,0008475947	0,001842743	5,6450293 $\times 10^{-8}$	1,6777216 $\times 10^{-5}$	0,995	0,995	+
14	1,0705619	1,070571	9,0271011 $\times 10^{-6}$	0,00084321146	1,8178858 $\times 10^{-5}$	5,6450293 $\times 10^{-8}$	6,7108864 $\times 10^{-6}$	0,995	0,995	+
15	1,0705619	1,070571	9,0266356 $\times 10^{-6}$	0,00084316798	9,9596699 $\times 10^{-8}$	5,6450293 $\times 10^{-8}$	2,6843546 $\times 10^{-6}$	0,995	0,995	–

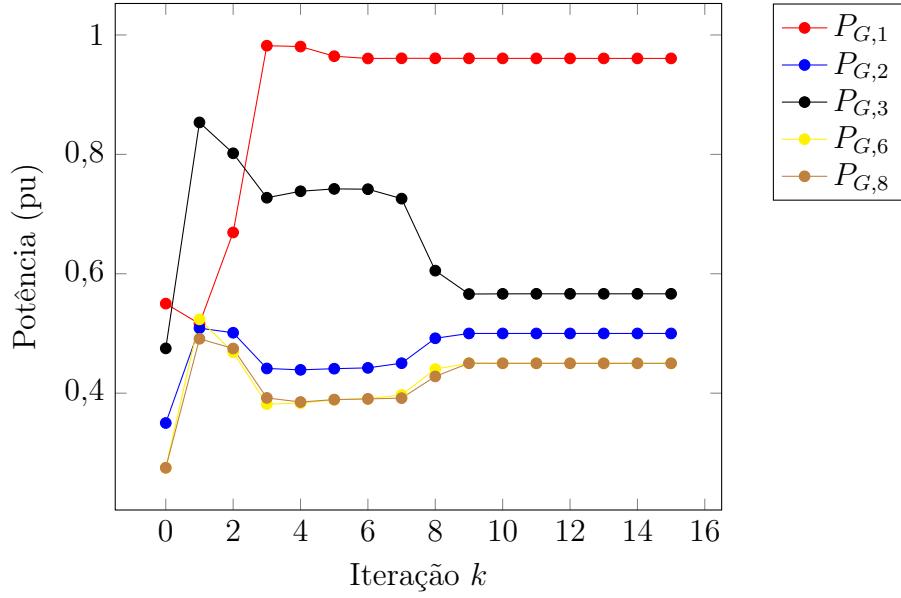


Figura 6.39: Sequência de potências geradas pelo método PDPCBLM para o PDE-PV-RT de 14 Barras (Parte I).

de 0,00084316798%. Outro ponto que merece atenção é o fato de que não houve correção de inércia, assim como no caso de 14 barras do PDE-PPV.

6.3.4 Sistema IEEE-14 Barras – Parte II

O objetivo desta seção é investigar o que ocorre quando há sobrecarga em alguma linha de transmissão, na resolução do PDE-PV-RT associado ao sistema IEEE-14 barras. Reduziu-se o fluxo máximo de potência ativa permitido na linha 1–2, de 1 pu para 0,45 pu, a qual era a linha com maior fluxo ao se executar o fluxo de potência na solução final do PDE-PPV. Para inicialização do método, utilizou-se os mesmos parâmetros da Parte I (Tabela 6.41). O método convergiu em 12 iterações, e os resultados estão resumidos na Tabela 6.44.

As Figuras 6.43, 6.44, 6.45 e 6.46 a seguir apresentam a trajetória das potências geradas, dos ângulos de tensão, dos multiplicadores de Lagrange e do parâmetro de correção de inércia, ao longo do processo iterativo. Neste caso, a diferença relativa entre a função original e aproximada, na solução ótima, foi de $2,5889845 \times 10^{-6}\%$. Assim como ocorreu no caso de 6 barras, apesar da redução do limite de fluxo máximo imposto à linha de transmissão 1–2, o método PDPCBLM determinou outra solução, em que nenhuma linha encontra-se em situação de sobrecarga, ainda que com um custo maior. A Tabela 6.45 apresenta a solução ótima determinada, em que os ângulos de tensão estão expressos em graus.

Neste caso, quando não há congestionamento nas linhas, os preços nodais de energia

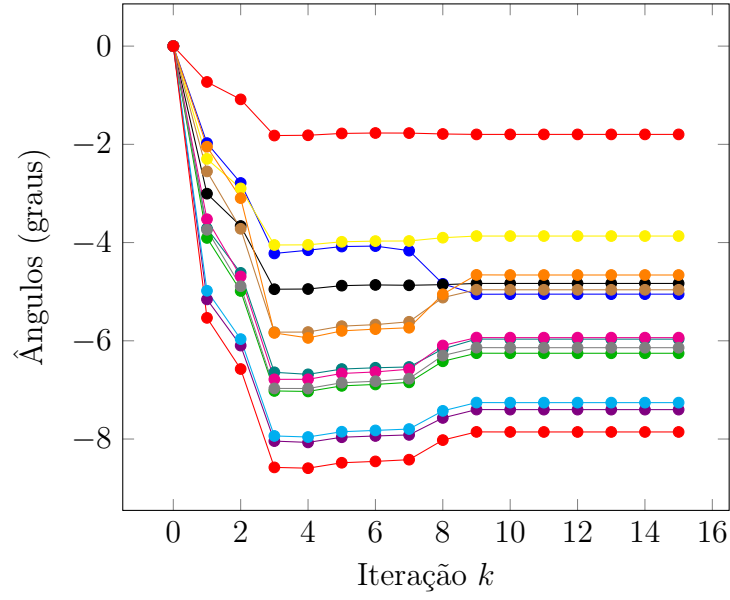


Figura 6.40: Sequência de ângulos de tensão nas barras (diferentes da *slack*) geradas pelo método PDPCBLM para o PDE-PV-RT de 14 Barras (Parte I).

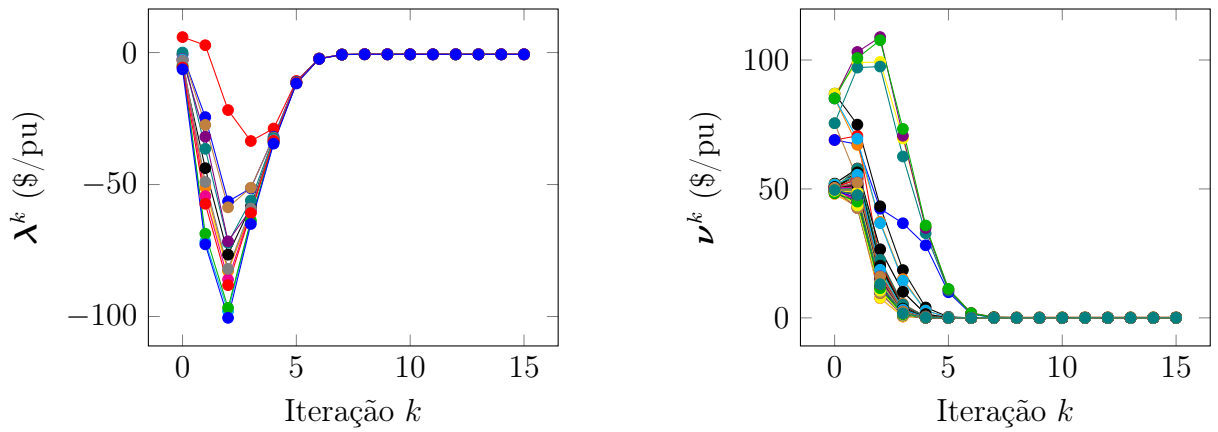


Figura 6.41: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PV-RT de 14 Barras (Parte I).

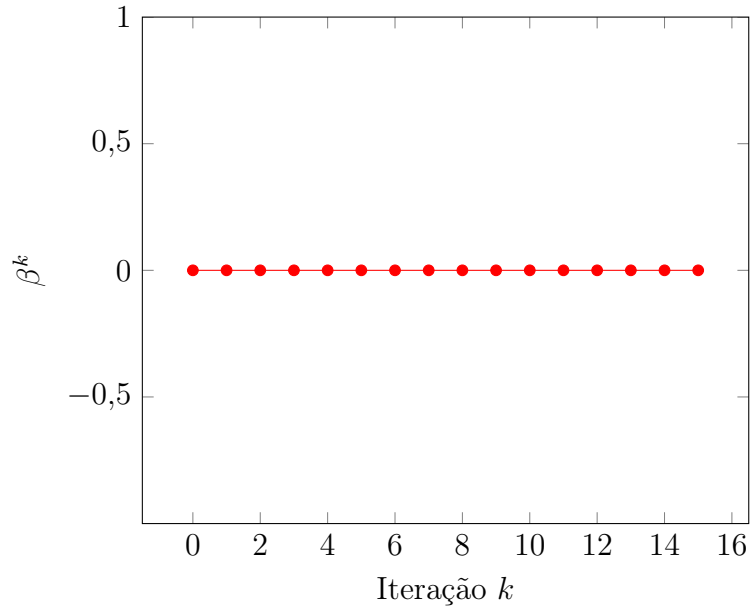


Figura 6.42: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PV-RT de 14 Barras (Parte I).

Tabela 6.43: Resultados – PDE-PV-RT: 14 Barras (Parte I)

Variável	PDE-PPV	PDE-PV-RT
$P_{G,1}$	96,0739	96,07369
$P_{G,2}$	50	50
$P_{G,3}$	56,6433	56,64359
$P_{G,6}$	45	45
$P_{G,8}$	45	45
θ_1	0	0
θ_2	-1,7976297	-1,7976232
θ_3	-5,0501767	-5,0501545
θ_4	-4,8321096	-4,8320998
θ_5	-3,8673193	-3,8673113
θ_6	-4,9614839	-4,9614756
θ_7	-5,9663423	-5,9663333
θ_8	-4,6603299	-4,6603206
θ_9	-7,4013548	-7,4013459
θ_{10}	-7,2589876	-7,2589787
θ_{11}	-6,2539031	-6,2538945
θ_{12}	-5,9358079	-5,9357995
θ_{13}	-6,1402754	-6,140267
θ_{14}	-7,8559872	-7,8559785
Função Objetivo (\$)	1070,56	1070,56

Tabela 6.44: Resultados (Iterações) – PDE-PV-RT: 14 Barras (Parte II)

Iteração k	$f(\mathbf{x}^k)$	$f_{\eta^k}(\mathbf{x}^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	0,68670723	0,69036127	0,003654043	0,53211075	10,400225	0,01	1	0	0	–
1	1,3624426	1,3629742	0,00053161653	0,039019372	54,938284	0,01	1	0,995	0,995	–
2	1,2753617	1,2759775	0,00061585543	0,048288689	17,939389	0,00333333333	0,4	0,995	0,995	–
3	1,2670356	1,2671287	9,3105162 $\times 10^{-5}$	0,007348267	4,2900029	0,00111111111	0,16	0,995	0,995	+
4	1,2633383	1,2633707	3,2414509 $\times 10^{-5}$	0,0025657821	1,2698682	0,00037037037	0,064	0,995	0,995	–
5	1,2663707	1,2663814	1,0729183 $\times 10^{-5}$	0,00084723876	0,48690102	0,00012345679	0,0256	0,995	0,995	+
6	1,2667643	1,2667678	3,5765181 $\times 10^{-6}$	0,00028233493	0,31136914	4,1152263 $\times 10^{-5}$	0,01024	0,995	0,995	+
7	1,2668438	1,266845	1,1922275 $\times 10^{-6}$	9,4110062 $\times 10^{-5}$	0,13978612	1,3717421 $\times 10^{-5}$	0,004096	0,995	0,995	+
8	1,2667299	1,2667303	3,9880626 $\times 10^{-7}$	3,1483132 $\times 10^{-5}$	0,036652694	4,5724737 $\times 10^{-6}$	0,0016384	0,995	0,995	+
9	1,2661839	1,266184	1,3468832 $\times 10^{-7}$	1,0637342 $\times 10^{-5}$	0,0064934151	1,5241579 $\times 10^{-6}$	0,00065536	0,995	0,995	–
10	1,2550462	1,2550463	6,716526 $\times 10^{-8}$	5,3516164 $\times 10^{-6}$	0,0090568419	5,0805263 $\times 10^{-7}$	0,000262144	0,995	0,33557099	–
11	1,2499105	1,2499105	3,2334679 $\times 10^{-8}$	2,5869596 $\times 10^{-6}$	0,00015583557	1,6935088 $\times 10^{-7}$	0,0001048576	0,995	0,995	+
12	1,2499079	1,2499079	3,235992 $\times 10^{-8}$	2,5889845 $\times 10^{-6}$	7,7947919 $\times 10^{-7}$	5,6450293 $\times 10^{-8}$	4,194304 $\times 10^{-5}$	0,995	0,995	+

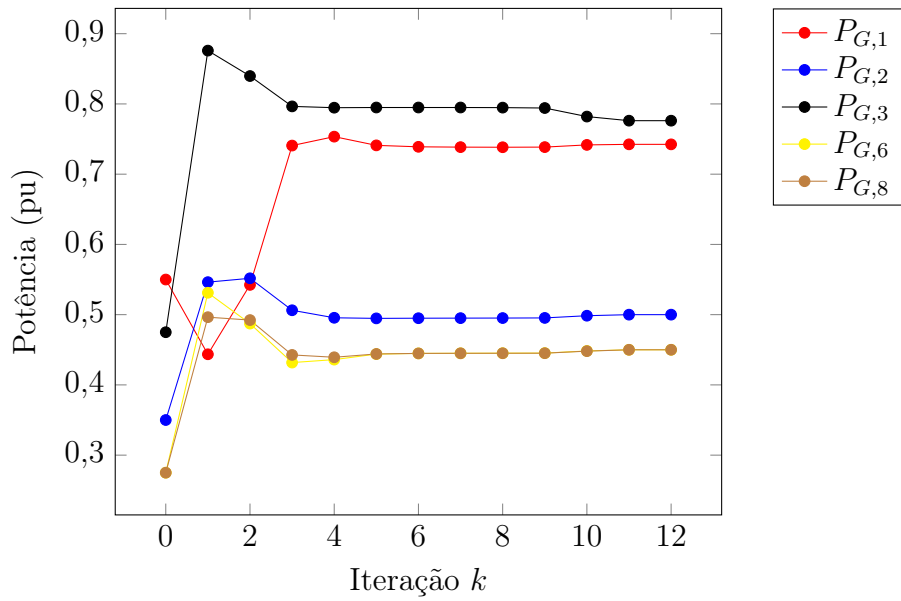


Figura 6.43: Sequência de potências geradas pelo método PDPCBLM para o PDE-PV-RT de 14 Barras (Parte II).

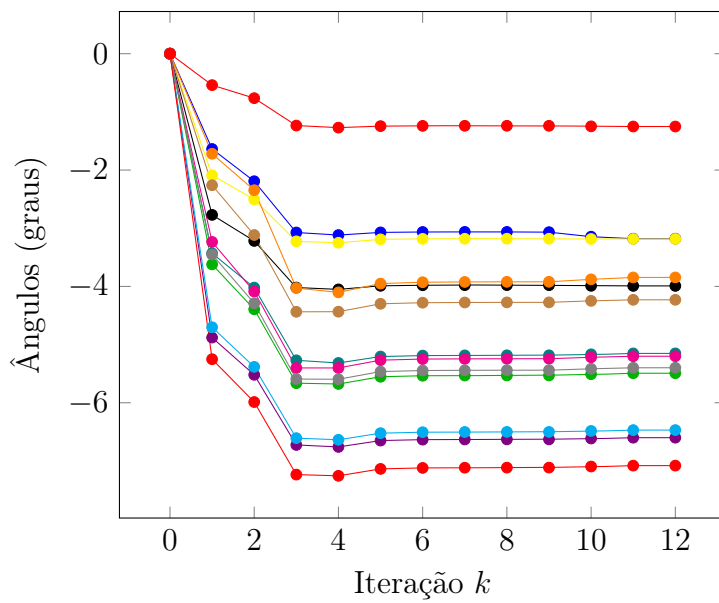


Figura 6.44: Sequência de ângulos de tensão nas barras (diferentes da *slack*) geradas pelo método PDPCBLM para o PDE-PV-RT de 14 Barras (Parte II).

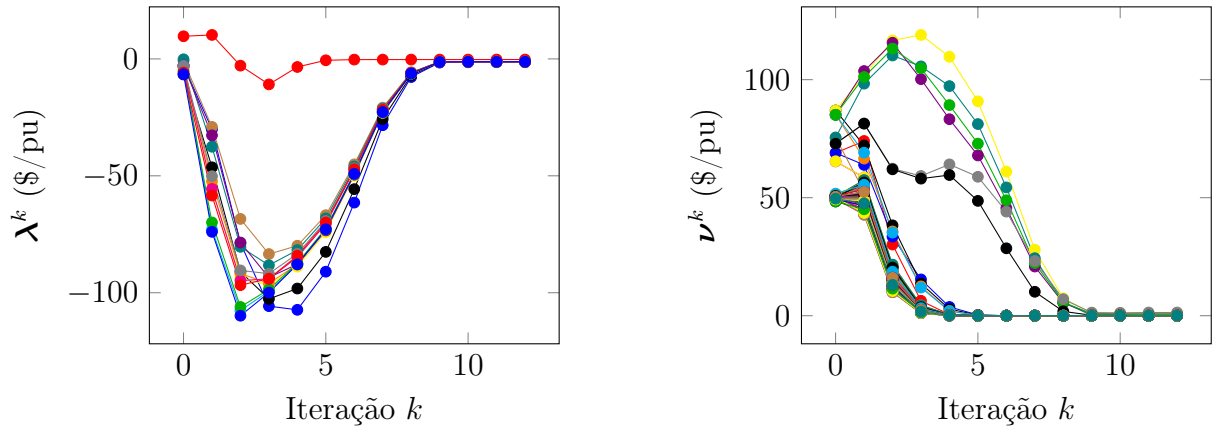


Figura 6.45: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PV-RT de 14 Barras (Parte II).

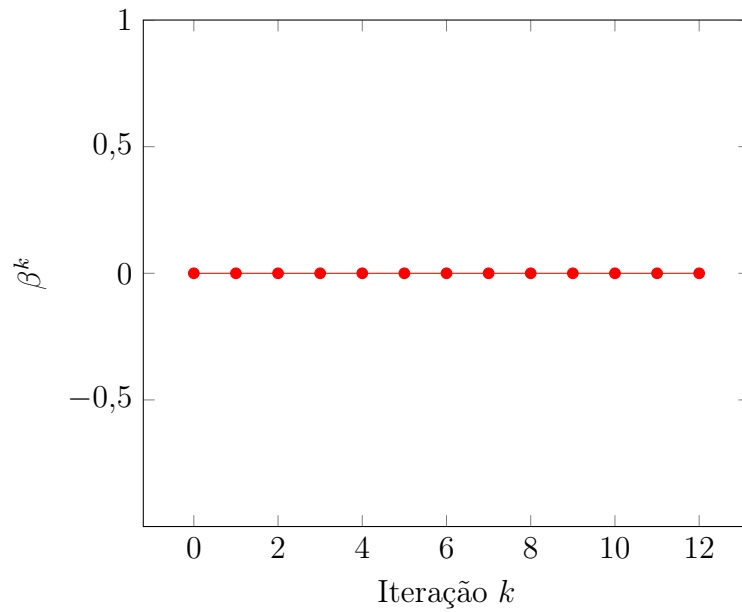


Figura 6.46: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PV-RT de 14 Barras (Parte II).

Tabela 6.45: Resultados – PDE-PV-RT: 14 Barras (Parte II)

Variável	PDE-PPV	PDE-PV-RT (Parte I)	PDE-PV-RT (Parte II)
$P_{G,1}$	96,0739	96,07369	74,24469
$P_{G,2}$	50	50	49,99999
$P_{G,3}$	56,6433	56,64359	77,61466
$P_{G,6}$	45	45	44,99999
$P_{G,8}$	45	45	44,99999
θ_1	0	0	0
θ_2	-1,7976297	-1,7976232	-1,2512581
θ_3	-5,0501767	-5,0501545	-3,1801169
θ_4	-4,8321096	-4,8320998	-3,9911707
θ_5	-3,8673193	-3,8673113	-3,1847377
θ_6	-4,9614839	-4,9614756	-4,2301301
θ_7	-5,9663423	-5,9663333	-5,1515849
θ_8	-4,6603299	-4,6603206	-3,8455727
θ_9	-7,4013548	-7,4013459	-6,6002806
θ_{10}	-7,2589876	-7,2589787	-6,4696827
θ_{11}	-6,2539031	-6,2538945	-5,4925272
θ_{12}	-5,9358079	-5,9357995	-5,2001275
θ_{13}	-6,1402754	-6,140267	-5,3984521
θ_{14}	-7,8559872	-7,8559785	-7,0803389
Função Objetivo (\$)	1070,56	1070,56	1249,21

obtidos pelo PDE-PV-RT ao preço “médio” obtido pelo PDE-PPV, sendo inferior em algumas barras e superior em outras. Na presença de sobrecarga na linha 1–2, além de um aumento de cerca de 20% no valor da função objetivo, notou-se que o preço na maioria das barras aumentou. Observa-se que o menor preço ocorre na barra 1, enquanto o maior ocorre na barra 2, exatamente aquelas relacionadas à linha de transmissão que teve limite máximo de fluxo de potência reduzido. A Figura 6.47 apresenta os preços nodais do PDE-PPV, PDE-PV-RT (Parte I) e PDE-PV-RT (Parte II), para o sistema de 14 barras. Testando outras situações de sobrecarga, notou-se que o método convergia para outras soluções, inclusive com preço negativo na barra 1. Isto pode ocorrer, segundo Palma-Benhke et al. (2013), e indica uma situação em que o custo pode ser reduzido pelo aumento do consumo na barra. Os mesmos autores apontam que preços negativos podem indicar uma possível perda de otimalidade, dificultando a convergência de algoritmos.

6.3.5 Sistema IEEE-30 Barras – Parte I

A discussão feita para os sistemas de 6 e 14 barras é, agora, estendida para o sistema IEEE-30 barras. Assume-se que o limite para fluxo nas linhas de transmissão é de 1 pu, exceto o da linha 1–2, que é de 1,5 pu visando evitar congestionamentos, já que ao resolver o PDE-PPV para este sistema, notou-se que o fluxo de potência na linha 1–2

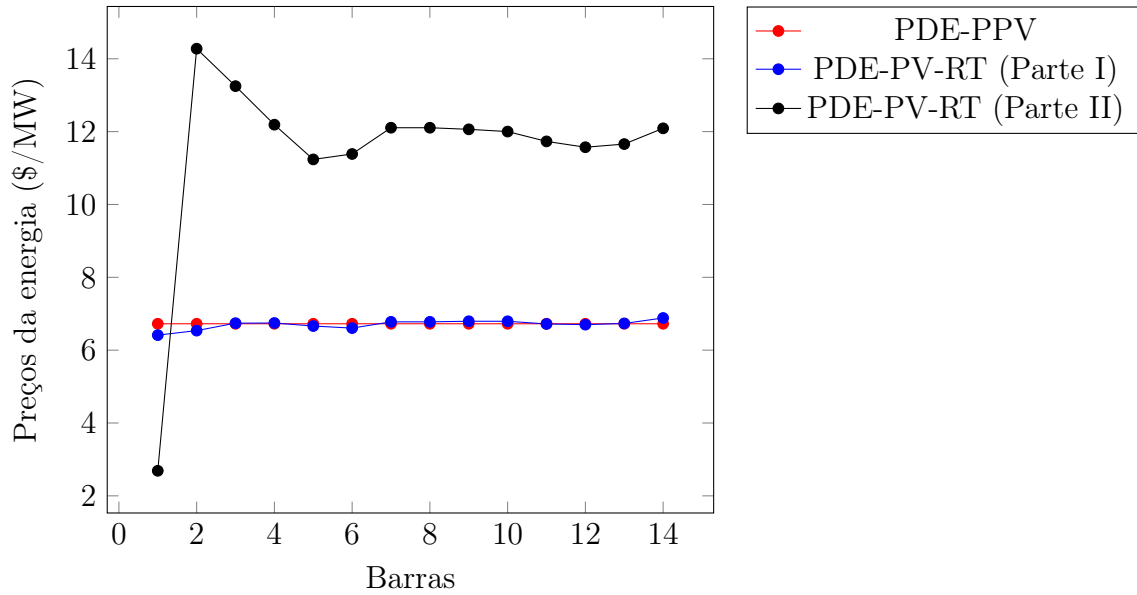


Figura 6.47: Comparação entre os preços de energia obtidos ao resolver o PDE-PPV, PDE-PV-RT (Parte I) e PDE-PV-RT (Parte II) para o sistema 14 Barras.

era de aproximadamente 1,2 pu. A formulação do modelo do PDE-PV-RT para este caso possui:

- 35 variáveis;
- 30 restrições de igualdade;
- 94 restrições de desigualdade.

Para inicialização do método, utilizou-se os parâmetros na Tabela 6.46. O método convergiu em 31 iterações, e os resultados estão resumidos na Tabela 6.47.

As Figuras 6.48, 6.49, 6.50 e 6.51 a seguir apresentam a trajetória das potências geradas, dos ângulos de tensão, dos multiplicadores de Lagrange e do parâmetro de correção de inércia, ao longo do processo iterativo. A solução obtida melhora, levemente, aquela obtida no caso do PDE-PPV, conforme indica a Tabela 6.48. Além disso, a diferença relativa entre a função objetivo original e aproximada, calculadas na solução determinada, foi de 0,036194785%. Nota-se que o método precisou realizar correções de inércia em diversas iterações, até que entre as iterações 20 e 25 ele consegue encontrar uma boa trajetória para convergência.

6.3.6 Sistema IEEE-30 Barras – Parte II

Analisa-se, nesta seção, a influência de sobrecarga na transmissão para o PDE-PV-RT associado ao sistema IEEE-30 barras. Reduziu-se o fluxo máximo de potência ativa per-

Tabela 6.46: Inicialização – PDE-PV-RT: 30 Barras (Parte I)

Fatores de aumento/redução da correção de inércia	$\kappa_1 = \kappa_2 = 10$	
Fatores de redução do parâmetro de suavização	$\rho_1 = 2$ e $\rho_2 = 0$	
Limitante da atualização do parâmetro de suavização	$\bar{\eta} = 10^{-5}$	
Parâmetro de barreira/relaxação inicial	$\mu^0 = 1$	
Fatores de atualização do parâmetro de barreira	$\tau_1 = \tau_2 = 0,6$	
Fator de controle da escolha entre passo previsor ou corretor	$\chi = 1$	
Fator de redução do passo	$\sigma = 0,95$	
Precisão de parada	$\varepsilon = 10^{-6}$	
Potência inicial das unidades geradoras (pu)	1	1,25
	2	0,50
	5	0,325
	8	0,225
	11	0,2
	13	0,26
Vetor de ângulos nas barras	$\theta = \mathbf{0}$	
Estimadores dos multiplicadores de Lagrange	$\delta^0 = \mathbb{1}_p$	

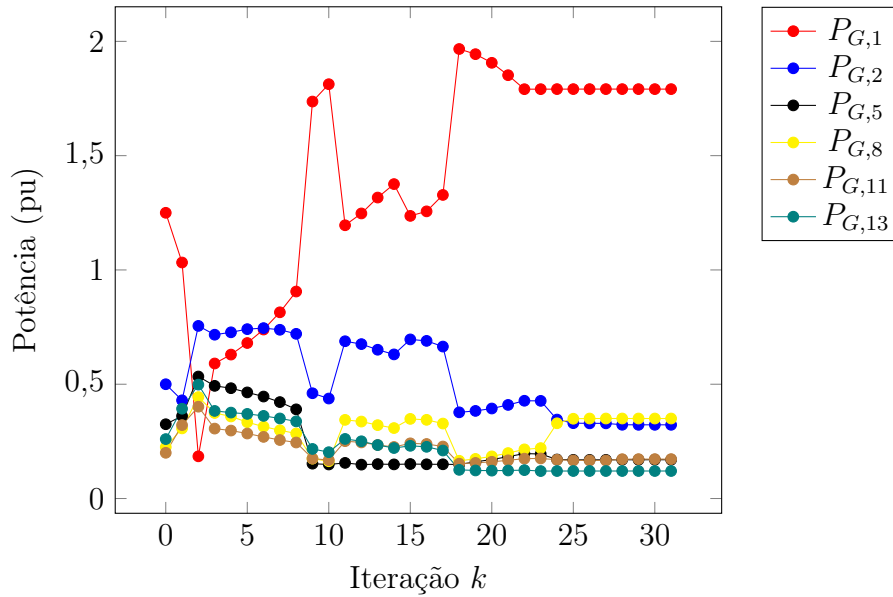


Figura 6.48: Sequência de potências geradas pelo método PDPCBLM para o PDE-PV-RT de 30 Barras (Parte I).

Tabela 6.47: Resultados (Iterações) – PDE-PV-RT: 30 Barras (Parte I)

Iteração k	$f(x^k)$	$f_{\eta^k}(x^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(x^k, s^k, \lambda^k, \nu^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	0,96477584	0,98679517	0,022019335	2,2823265	1,1114907	0,1	1	0	0	—
1	1,0545108	1,0645093	0,009985515	0,94816968	49,607053	0,1	1	0,95	0,95	—
2	1,2669632	1,2712106	0,0042474013	0,33524267	1,9005308	0,05	0,6	0,087831314	0,14020305	—
3	1,1587124	1,1611603	0,0024478409	0,21125526	0,99756433	0,025	0,36	0,95	0,37812469	—
4	1,1741561	1,1752508	0,0010947298	0,093235454	0,41879084	0,0125	0,216	0,95	0,95	—
5	1,1782871	1,1788094	0,00052233121	0,044329707	0,55664944	0,00625	0,1296	0,95	0,95	—
6	1,1644787	1,1647451	0,00026637633	0,022875156	0,6261027	0,003125	0,07776	0,95	0,95	—
7	1,1168403	1,1169922	0,00015196669	0,013606842	0,73946123	0,0015625	0,046656	0,95	0,95	+
8	1,0267029	1,0268795	0,0001765591	0,017196708	0,78727601	0,00078125	0,0279936	0,95	0,95	+
9	0,91109251	0,9113496	0,00025708702	0,028217444	0,59700331	0,000390625	0,01679616	0,70338763	0,20379727	—
10	0,88176703	0,88199832	0,00023128292	0,026229481	1,0674113	0,0001953125	0,010077696	0,95	0,95	+
11	1,0408044	1,0408367	3,2349541 $\times 10^{-5}$	0,0031081289	0,27030046	9,765625 $\times 10^{-5}$	0,0060466176	0,55951368	0,14239343	—
12	1,0127291	1,0127715	4,2397744 $\times 10^{-5}$	0,0041864842	0,52053795	4,8828125 $\times 10^{-5}$	0,0036279706	0,87106777	0,95	+
13	0,95170778	0,95175603	4,8257687 $\times 10^{-5}$	0,0050706413	0,6922542	2,4414063 $\times 10^{-5}$	0,0021767823	0,95	0,48401207	—
14	0,918855	0,91888194	2,6947862 $\times 10^{-5}$	0,0029327654	1,0811587	1,2207031 $\times 10^{-5}$	0,0013060694	0,76429704	0,95	+
15	1,0172353	1,0172544	1,9159655 $\times 10^{-5}$	0,0018835028	0,34471971	6,1035156 $\times 10^{-6}$	0,00078364164	0,16010383	0,0071548423	—
16	1,002787	1,0028358	4,8854064 $\times 10^{-5}$	0,0048718288	0,43965064	6,1035156 $\times 10^{-6}$	0,00047018498	0,27430698	0,12052252	—
17	0,93800208	0,9380516	4,9516872 $\times 10^{-5}$	0,0052789725	0,64078956	6,1035156 $\times 10^{-6}$	0,00028211099	0,95	0,95	+
18	0,95741651	0,9574369	2,0387078 $\times 10^{-5}$	0,0021293844	0,54507576	6,1035156 $\times 10^{-6}$	0,00036138293	0,013089226	0,0017081372	—
19	0,95365911	0,95366917	1,0064116 $\times 10^{-5}$	0,0010553159	0,65236252	6,1035156 $\times 10^{-6}$	0,0018105643	0,95	0,056808084	—
20	0,93760834	0,93762045	1,2108277 $\times 10^{-5}$	0,0012914003	0,39360087	6,1035156 $\times 10^{-6}$	0,0010863386	0,95	0,95	+
21	0,90371943	0,90372889	9,4642203 $\times 10^{-6}$	0,0010472521	0,5420678	6,1035156 $\times 10^{-6}$	0,00065180313	0,95	0,59129943	—
22	0,85731381	0,85745884	0,00014503046	0,016916846	0,87920747	6,1035156 $\times 10^{-6}$	0,00039108188	0,95	0,95	+
23	0,85676054	0,85700076	0,00024021985	0,028038155	0,52310927	6,1035156 $\times 10^{-6}$	0,00023464913	0,17179653	0,95	+
24	0,85317793	0,85345076	0,00027282857	0,031977921	0,14081046	6,1035156 $\times 10^{-6}$	0,00014078948	0,95	0,95	+
25	0,85285809	0,85313954	0,00028144929	0,033000717	0,11496249	6,1035156 $\times 10^{-6}$	8,4473686 $\times 10^{-5}$	0,21494488	0,95	+
26	0,85284467	0,85312653	0,00028185234	0,033048496	0,11450853	6,1035156 $\times 10^{-6}$	5,0684212 $\times 10^{-5}$	0,012572308	0,72547216	—
27	0,85284417	0,85312604	0,000281868	0,033050353	0,11460469	6,1035156 $\times 10^{-6}$	3,0410527 $\times 10^{-5}$	0,00049711765	0,73244427	—
28	0,85292105	0,8532298	0,00030875103	0,036199251	0,0034595934	6,1035156 $\times 10^{-6}$	1,8246316 $\times 10^{-5}$	0,95	0,81789046	—
29	0,85292816	0,853323689	0,00030872804	0,036196253	0,00017302819	6,1035156 $\times 10^{-6}$	0,00012794936	0,95	0,95	+
30	0,85292887	0,85323759	0,00030871646	0,036194865	8,6526135 $\times 10^{-6}$	6,1035156 $\times 10^{-6}$	7,6769619 $\times 10^{-5}$	0,95	0,95	+
31	0,85292892	0,85323763	0,00030871578	0,036194785	4,3263528 $\times 10^{-7}$	6,1035156 $\times 10^{-6}$	4,6061771 $\times 10^{-5}$	0,95	0,95	+

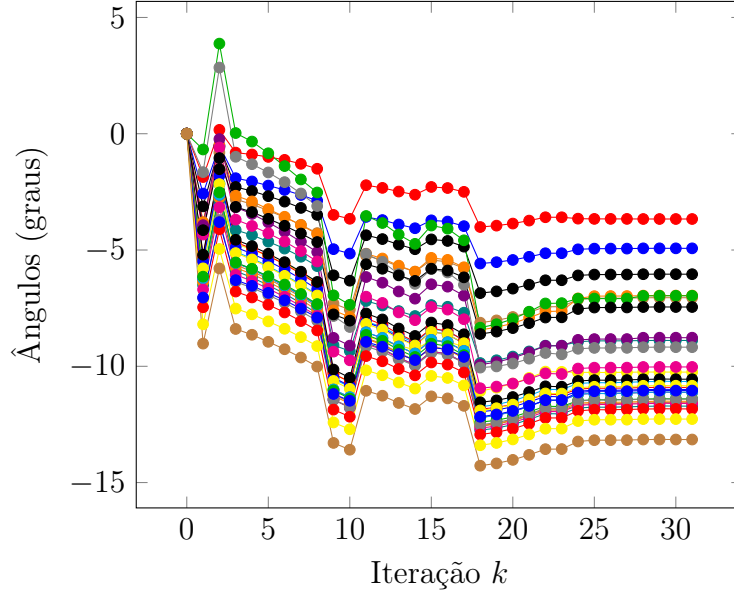


Figura 6.49: Sequência de ângulos de tensão nas barras (diferentes da *slack*) geradas pelo método PDPCBLM para o PDE-PV-RT de 30 Barras (Parte I).

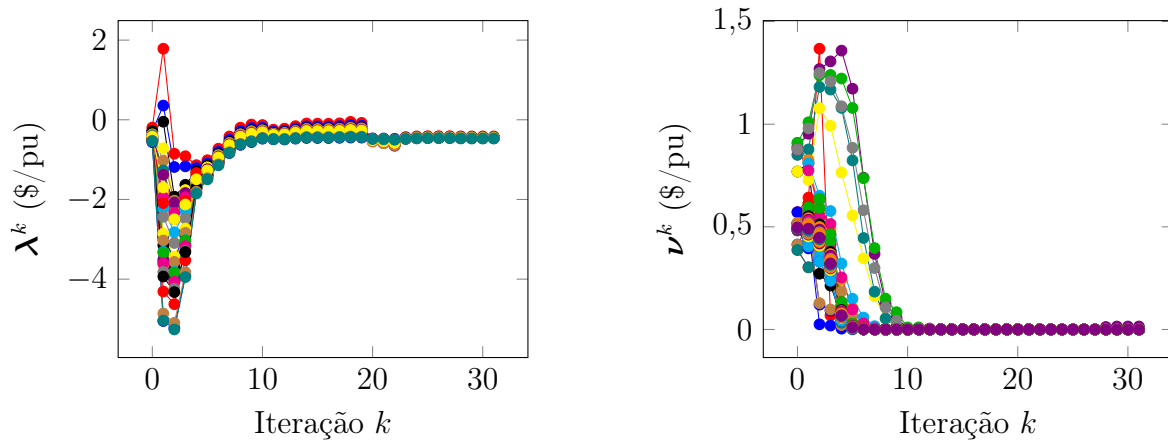


Figura 6.50: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PV-RT de 30 Barras (Parte I).

Tabela 6.48: Resultados – PDE-PV-RT: 30 Barras (Parte I)

Variável	PDE-PPV	PDE-PV-RT
$P_{G,1}$	179,1104	179,10979
$P_{G,2}$	31,082	32,24451
$P_{G,5}$	16,8028	16,95909
$P_{G,8}$	35	35
$P_{G,11}$	18,5226	17,24442
$P_{G,13}$	12,0462	12,03791
θ_1	0	0
θ_2	-3,6748744	-3,6665933
θ_3	-4,9045768	-4,9279395
θ_4	-6,0106255	-6,0396179
θ_5	-10,245512	-10,249558
θ_6	-7,0137254	-7,051465
θ_7	-8,8655974	-8,8896209
θ_8	-6,9217869	-6,9598469
θ_9	-8,6469167	-8,7703249
θ_{10}	-10,557223	-10,6525
θ_{11}	-6,7104989	-6,9675793
θ_{12}	-9,9686714	-10,028264
θ_{13}	-9,1176085	-9,1777904
θ_{14}	-10,846284	-10,908922
θ_{15}	-10,928985	-10,995959
θ_{16}	-10,479711	-10,555107
θ_{17}	-10,746999	-10,835932
θ_{18}	-11,488131	-11,565105
θ_{19}	-11,630031	-11,712966
θ_{20}	-11,418717	-11,504799
θ_{21}	-11,027654	-11,12003
θ_{22}	-11,023444	-11,11488
θ_{23}	-11,328567	-11,400224
θ_{24}	-11,519377	-11,59733
θ_{25}	-11,351736	-11,414391
θ_{26}	-11,769402	-11,832056
θ_{27}	-10,990686	-11,043988
θ_{28}	-7,4112508	-7,4508689
θ_{29}	-12,215962	-12,269264
θ_{30}	-13,09533	-13,148632
Função Objetivo (\$)	852,9911	852,9289

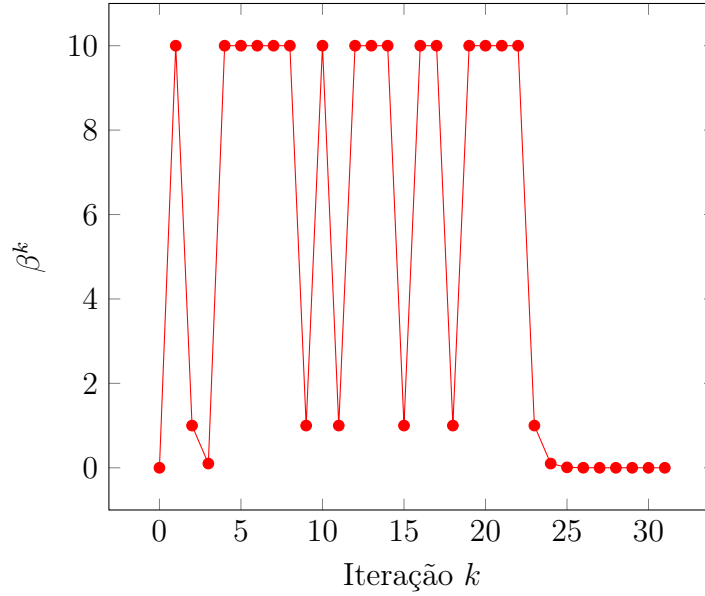


Figura 6.51: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PV-RT de 30 Barras (Parte I).

mitido na linha 1–2, de 1,5 pu para 0,8 pu, a qual era a linha com maior fluxo ao se executar o fluxo de potência na solução final do PDE-PPV. Para inicialização do método, utilizou-se os mesmos parâmetros da Parte I (Tabela 6.46). O método convergiu em 18 iterações, e os resultados estão resumidos na Tabela 6.49.

As Figuras 6.52, 6.53, 6.54 e 6.55 a seguir apresentam a trajetória das potências geradas, dos ângulos de tensão, dos multiplicadores de Lagrange e do parâmetro de correção de inércia, ao longo do processo iterativo. Neste caso, a diferença relativa entre a função original e aproximada, na solução ótima, foi de 0,022734966%. Mais uma vez, o método resolveu o problema com congestionamento na transmissão, determinando uma solução, de maior custo, em que nenhuma linha encontra-se em situação de sobrecarga. A Tabela 6.50 apresenta a solução ótima determinada, em que os ângulos de tensão estão expressos em graus.

A Figura 6.56 apresenta os preços nodais de energia, obtidos pelas soluções do PDE-PPV, PDE-PV-RT (Parte I) e PDE-PV-RT (Parte II). Com a análise deste gráfico, percebe-se que na situação de linhas descongestionadas, a solução determinada pelo PDE-PV-RT apresenta preços nodais de energia menores, na maioria das barras. O preço da energia na barra 1 é significativamente menor que nas demais, pois o gerador situado nesta barra produz grande parcela da potência ativa despachada. Na situação de sobrecarga, nota-se que os preços nodais da energia aumentaram em todas as barras.

A análise dos preços nodais de energia obtidos nos testes do PDE-PV-RT permite concluir que, em geral, os preços para o consumidor tendem a ficar maiores em situações de congestionamento das linhas de transmissão. Este fato mostra a importância da represen-

Tabela 6.49: Resultados (Iterações) – PDE-PV-RT: 30 Barras (Parte II)

Iteração k	$f(\mathbf{x}^k)$	$f_{\eta^k}(\mathbf{x}^k)$	Δf_{abs}^k	Δf_{rel}^k (%)	$\ \nabla \mathcal{L}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k)\ _{\infty}$	η_k	μ^k	α_P^k	α_D^k	Corretor
0	0,96477584	0,98679517	0,022019335	2,2823265	1,1114907	0,1	1	0	0	–
1	1,0441104	1,0549478	0,010837374	1,0379529	2,7418813	0,1	1	0,95	0,95	–
2	1,13084	1,1404411	0,0096010724	0,84902129	0,47423964	0,05	0,6	0,95	0,95	–
3	1,1406081	1,1434387	0,0028306673	0,24817178	1,5047079	0,025	0,36	0,95	0,95	–
4	0,94093515	0,94346162	0,0025264769	0,26850701	0,4640953	0,0125	0,216	0,95	0,28766318	–
5	1,1447598	1,1453301	0,00057026022	0,049814835	0,85521137	0,00625	0,1296	0,50756087	0,39445458	+
6	1,0974796	1,0978062	0,00032660033	0,029759124	0,93069022	0,003125	0,07776	0,95	0,95	+
7	0,99338574	0,9962902	0,0029044695	0,29238083	0,50986567	0,0015625	0,046656	0,95	0,95	+
8	0,98949671	0,99107222	0,0015755088	0,15922325	0,096361282	0,00078125	0,0279936	0,95	0,94728932	–
9	0,97645403	0,97759119	0,0011371615	0,11645828	0,086082238	0,000390625	0,01679616	0,95	0,95	+
10	0,97355346	0,9743499	0,0007964368	0,081807197	0,089731871	0,0001953125	0,010077696	0,95	0,95	+
11	0,97330536	0,97386734	0,00056197801	0,057739126	0,089018364	$9,765625 \times 10^{-5}$	0,0060466176	0,95	0,95	+
12	0,97320111	0,97359712	0,00039600489	0,040690962	0,089280075	$4,8828125 \times 10^{-5}$	0,0036279706	0,95	0,95	+
13	0,97314562	0,973425	0,00027937726	0,02870868	0,089269943	$2,4414063 \times 10^{-5}$	0,0021767823	0,95	0,95	+
14	0,97310841	0,97330375	0,00019722475	0,0202675	0,089274833	$1,2207031 \times 10^{-5}$	0,0013060694	0,95	0,95	–
15	0,97308225	0,97330375	0,00022150435	0,022763169	0,00099975143	$6,1035156 \times 10^{-6}$	0,00078364164	0,95	0,95	+
16	0,97308252	0,97330377	0,0002212438	0,022736386	$5,0470472 \times 10^{-5}$	$6,1035156 \times 10^{-6}$	0,00047018498	0,95	0,95	+
17	0,97308254	0,97330377	0,00022123064	0,022735034	$2,5247559 \times 10^{-6}$	$6,1035156 \times 10^{-6}$	0,00028211099	0,95	0,95	+
18	0,97308254	0,97330377	0,00022122998	0,022734966	$1,2624095 \times 10^{-7}$	$6,1035156 \times 10^{-6}$	0,00016926659	0,95	0,95	+

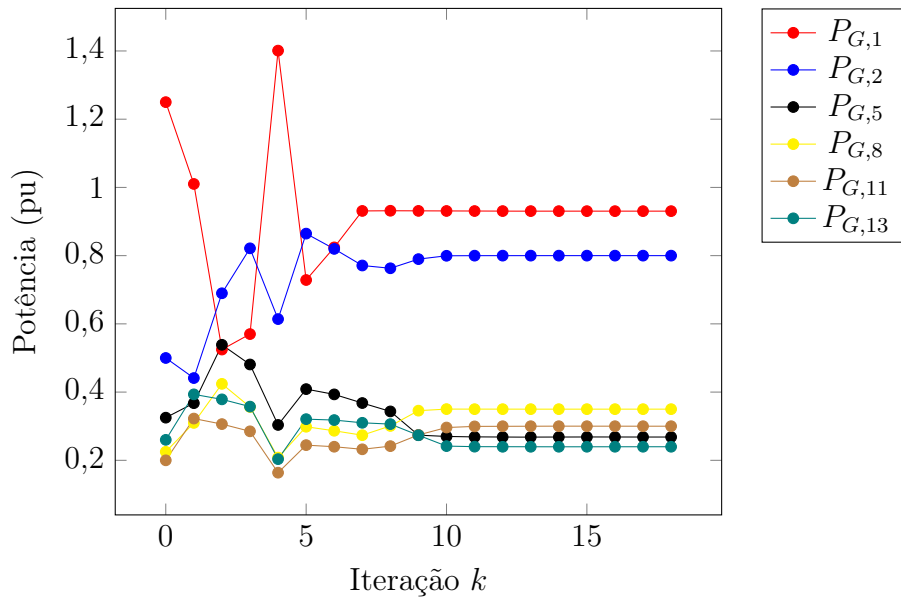


Figura 6.52: Sequência de potências geradas pelo método PDPCBLM para o PDE-PV-RT de 30 Barras (Parte II).

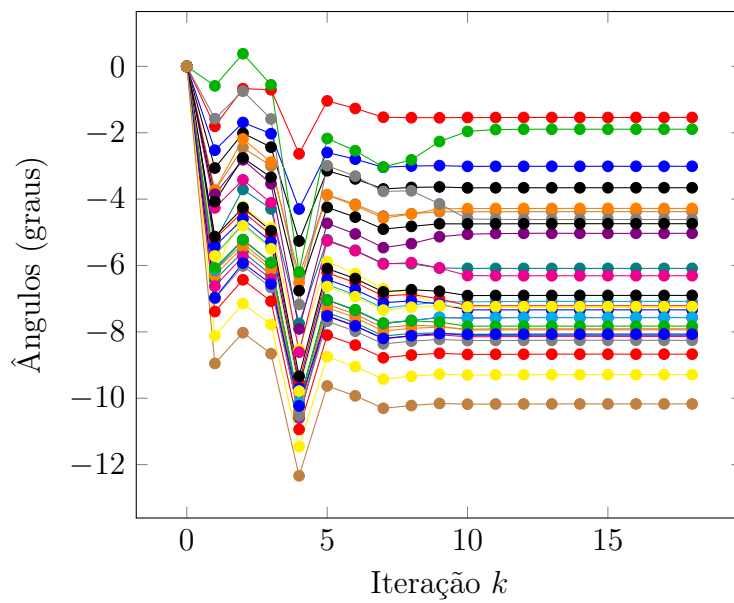


Figura 6.53: Sequência de ângulos de tensão nas barras (diferentes da *slack*) geradas pelo método PDPCBLM para o PDE-PV-RT de 30 Barras (Parte II).

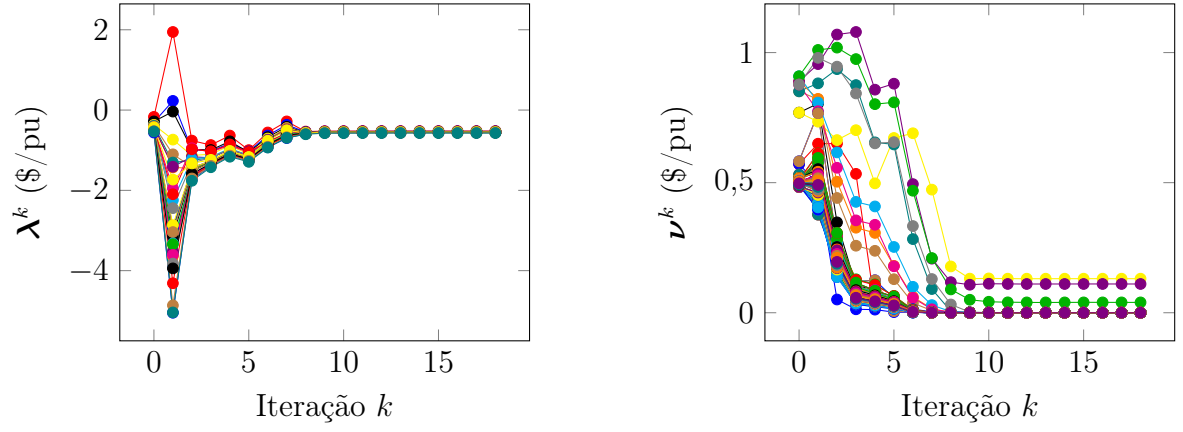


Figura 6.54: Evolução dos multiplicadores de Lagrange das restrições de igualdade, λ^k , e desigualdade, ν^k , no caso PDE-PV-RT de 30 Barras (Parte II).

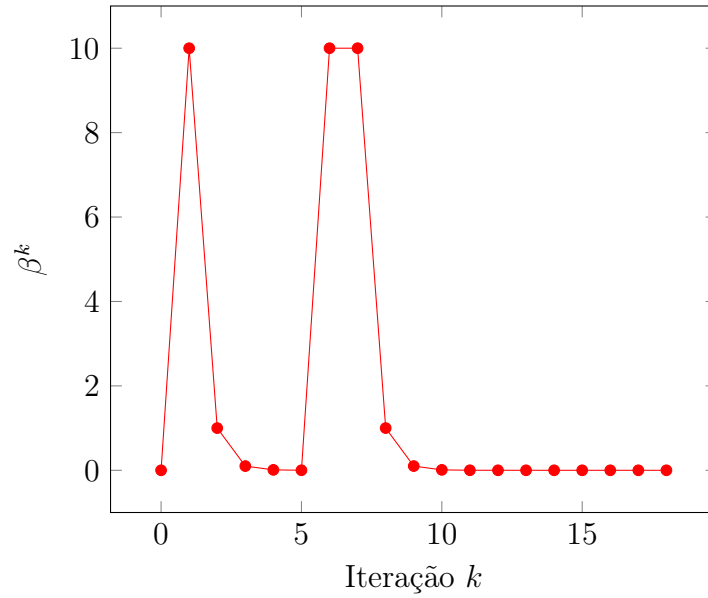


Figura 6.55: Sequência de parâmetros de correção de inércia, β^k , no método PDPCBLM para o PDE-PV-RT de 30 Barras (Parte II).

Tabela 6.50: Resultados – PDE-PV-RT: 30 Barras (Parte II)

Variável	PDE-PPV	PDE-PV-RT (Parte I)	PDE-PV-RT (Parte II)
$P_{G,1}$	179,1104	179,10979	93,045688
$P_{G,2}$	31,082	32,24451	80
$P_{G,5}$	16,8028	16,95909	26,825704
$P_{G,8}$	35	35	35
$P_{G,11}$	18,5226	17,24442	30
$P_{G,13}$	12,0462	12,03791	23,977385
θ_1	0	0	0
θ_2	-3,6748744	-3,6665933	-1,5400663
θ_3	-4,9045768	-4,9279395	-3,0109351
θ_4	-6,0106255	-6,0396179	-3,6573347
θ_5	-10,245512	-10,249558	-7,2561564
θ_6	-7,0137254	-7,051465	-4,3783301
θ_7	-8,8655974	-8,8896209	-6,0861951
θ_8	-6,9217869	-6,9598469	-4,2804439
θ_9	-8,6469167	-8,7703249	-5,0303464
θ_{10}	-10,557223	-10,6525	-7,0823376
θ_{11}	-6,7104989	-6,9675793	-1,8930735
θ_{12}	-9,9686714	-10,028264	-6,3082685
θ_{13}	-9,1176085	-9,1777904	-4,6140923
θ_{14}	-10,846284	-10,908922	-7,2138833
θ_{15}	-10,928985	-10,995959	-7,3364104
θ_{16}	-10,479711	-10,555107	-6,9014919
θ_{17}	-10,746999	-10,835932	-7,2391421
θ_{18}	-11,488131	-11,565105	-7,9371501
θ_{19}	-11,630031	-11,712966	-8,1038355
θ_{20}	-11,418717	-11,504799	-7,9056063
θ_{21}	-11,027654	-11,12003	-7,5682713
θ_{22}	-11,023444	-11,11488	-7,5690877
θ_{23}	-11,328567	-11,400224	-7,825725
θ_{24}	-11,519377	-11,59733	-8,1370401
θ_{25}	-11,351736	-11,414391	-8,2540827
θ_{26}	-11,769402	-11,832056	-8,6717483
θ_{27}	-10,990686	-11,043988	-8,0666479
θ_{28}	-7,4112508	-7,4508689	-4,741013
θ_{29}	-12,215962	-12,269264	-9,2919236
θ_{30}	-13,09533	-13,148632	-10,171292
Função Objetivo (\$)	852,9911	852,9289	973,0825

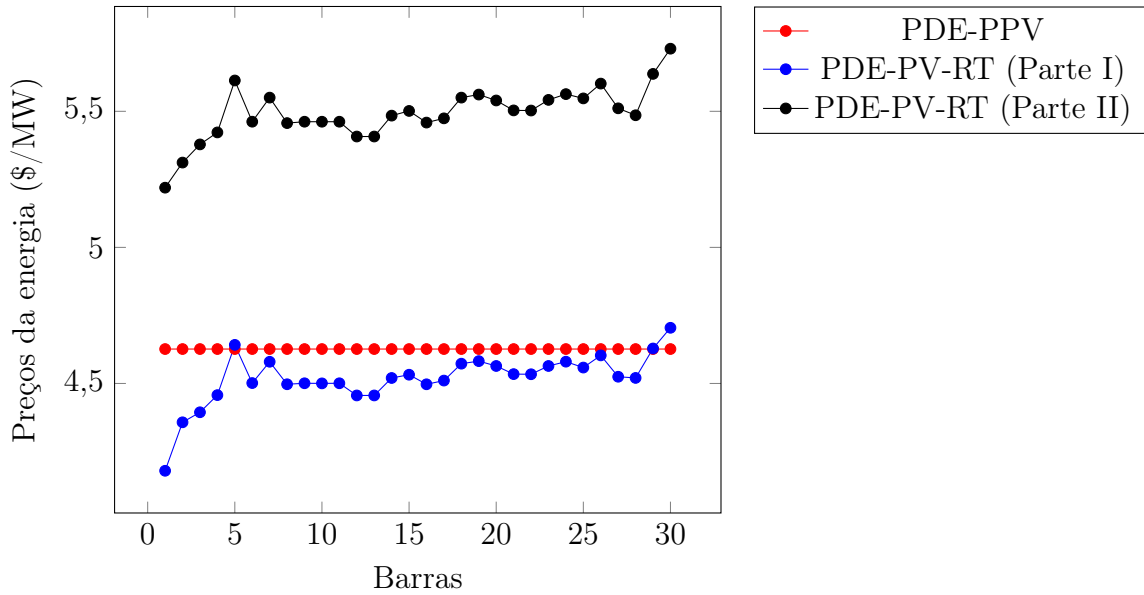


Figura 6.56: Comparação entre os preços de energia obtidos ao resolver o PDE-PPV, PDE-PV-RT (Parte I) e PDE-PV-RT (Parte II) para o sistema 30 Barras.

tação da transmissão de maneira explícita para a operação do sistema elétrico de maneira adequada.

Com a introdução dos mercados de energia do dia seguinte em diversos países, os modelos de despacho baseados em FPO passaram a ser substituídos por modelos de leilão, executados em base horária. Visando a simplicidade e transparência para o mercado, diversos modelos de leilão tendem a desconsiderar alguns aspectos, incluindo a transmissão, o que leva a soluções sub-ótimas devido à necessidade de correções posteriores. Os resultados obtidos nesta dissertação corroboram a noção de que os modelos de despacho devem incluir a representação da transmissão, a fim de evitar os desvios de otimalidade causados pelas correções posteriores e para que os preços pagos pelos consumidores sejam justos.

Capítulo 7

Conclusões e Perspectivas Futuras

O problema de despacho econômico com ponto de carregamento de válvula é modelado como um problema de programação não-linear, não-convexo e não-diferenciável, com restrições de igualdade e desigualdade. Por tais motivos, sua resolução na literatura usualmente é feita através de métodos heurísticos.

Neste trabalho, foi apresentado um método determinístico, a saber o método primal-dual previsor-corretor barreira logarítmica modificada (PDPCBLM), baseado no trabalho de Pinheiro (2012), associado a uma estratégia de correção de inércia e uma técnica de suavização hiperbólica, para a resolução de problemas de despacho econômico com ponto de carregamento de válvula, modelado de três maneiras distintas: sem representação das perdas (PDE-PV), com representação das perdas através dos coeficientes-B (PDE-PPV) e com representação da transmissão (PDE-PV-RT). A estratégia de correção de inércia, apresentada na Seção 5.1.2, previne a obtenção de pontos críticos que não são pontos de mínimos locais, como pontos de sela ou máximos locais, ao se aplicar as condições necessárias de primeira ordem à função Lagrangiana associada ao problema. Os resultados obtidos apontam que a estratégia utilizada permitiu a obtenção de mínimos locais, evitando a parada do método em pontos indesejados.

Além disso, tendo em vista a não-diferenciabilidade da função objetivo do PDE-PV, na Seção 5.2 apresentou-se uma estratégia de suavização de funções, baseada no trabalho de Junior (2010) e Xavier & Xavier (2011). Esta estratégia, denominada suavização hiperbólica, permite a obtenção de aproximações suaves para funções obtidas pela composição de funções diferenciáveis com a função modular, tal como a função custo do PDE-PV.

Os testes numéricos realizados e apresentados no Capítulo 6, através da linguagem de programação Matlab, apontam que o método PDPCBLM, associado às estratégias supracitadas, é capaz de resolver de modo eficiente e obter boas soluções para o PDE-PV e suas

variantes, PDE-PPV e PDE-PV-RT, quando comparado a outros trabalhos disponíveis na literatura. Destaca-se que os resultados obtidos para o PDE-PV-RT, os quais não são encontrados na literatura, mostraram o bom desempenho do método citado quando comparados ao do PDE-PPV. O método determinou soluções consistentes e semelhantes para ambos os problemas em destaque, nas situações em que não há congestionamento nas linhas de transmissão.

A principal dificuldade computacional do método proposto foi o controle do parâmetro de suavização, sendo que uma escolha inadequada do mesmo pode levar à divergência ou à instabilidade do método. Entretanto, os resultados teóricos e computacionais apresentados mostram que a suavização hiperbólica é uma técnica viável para a resolução de problemas não-diferenciáveis devido à existência de termos modulares. A técnica de suavização hiperbólica aqui apresentada pode ser empregada para a resolução de outros problemas de otimização, como problemas de FPO em que a função objetivo é a minimização dos desvios de tensão.

Além da resolução desse problema, como propostas futuras de continuidade deste trabalho, pode-se citar:

- estudar outras funções de rescalamento não-linear, além da função barreira modificada, para a resolução de problemas de programação não-linear. Por exemplo, os testes apresentados podem ser realizados utilizando alguma função de rescalamento apresentada no capítulo 4;
- estudar a implementação de algum critério para aceitação de passo (como função mérito ou método de filtros), visando melhorar a convergência e evitar oscilações nos resíduos, em decorrência dos baixos valores do parâmetro η , para os quais a função aproximante se aproxima da natureza não-diferenciável da função original;
- investigar a possibilidade de associar um método de região de confiança ao método de pontos interiores/exteriorizados que utiliza a função barreira modificada, visando torná-lo mais robusto em situações nas quais as não-convexidades do problema dificultam o processo de convergência;
- propor formas adaptativas para a atualização do parâmetro η da suavização hiperbólica, visando reduzir a possibilidade de divergência ou instabilidades numéricas;
- comparar o desempenho do método com sua versão sem passo corretor, bem como testar diferentes valores para o parâmetro χ , que controla a escolha entre a direção do previsor e do corretor;
- resolver de maneira direta o sistema de direções, visando aproveitar a decomposição LDL^t realizada, ao invés de utilizar a estratégia desenvolvida por Karmarkar (1984)

de decomposição das direções e solução através do método dos mínimos quadrados.

Capítulo 8

Trabalhos Publicados

Silva, D. N.; Balbo, A. R.; Nepomuceno, L.; Pinheiro, R. B. N. Resolvendo o problema de despacho econômico com ponto de válvula e restrição de segurança através de um método primal-dual de pontos interiores/exteriores. In: *XXI Simpósio de Engenharia de Produção – SIMPEP*, Bauru – SP, 2014.

Silva, D. N.; Balbo, A. R. Métodos de Rescalamento Não-Linear e Suavização de Funções em Problemas de Despacho Econômico com Ponto de Válvula. In: *IV Seminário da Pós-Graduação em Engenharia Elétrica*, Bauru – SP, 2014.

Bregadioli, G. F.; Gonçalves, E.; **Silva, D. N.**; Nepomuceno, L.; Baptista, E. C.; Balbo, A. R. Um estudo de caso sobre a influência da transmissão em modelos de leilão de mercados de energia. In: *20º Congresso Brasileiro de Automática – CBA 2014*, Belo Horizonte – MG, 2014.

Gonçalves, E.; **Silva, D. N.**; Balbo, A. R. Resolvendo o problema multiobjetivo de despacho econômico/ambiental através dos métodos da soma ponderada e de pontos interiores. In: *Congresso Nacional de Matemática Aplicada e Computacional – CNMAC 2014*, Natal – RN, 2014.

Silva, D. N.; Balbo, A. R.; Pinheiro, R. B. N. Método primal-dual previsor-corretor aplicado ao problema de despacho econômico com ponto de válvula e perdas na transmissão. In: *Congresso Nacional de Matemática Aplicada e Computacional – CNMAC 2014*, Natal – RN, 2014.

Silva, D. N.; Balbo, A. R.; Pinheiro, R. B. N. Uma estratégia para minimização de funções com termos modulares via métodos de pontos interiores. In: *Congresso de Matemática Aplicada e Computacional – CMAC Sul*, Curitiba – PR, 2014.

Silva, D. N.; Balbo, A. R. Um Método Primal-Dual de Pontos Interiores aplicado ao

Problema de Despacho Econômico com Ponto de Válvula. In: *III Seminário da Pós-Graduação em Engenharia Elétrica*, Bauru – SP, 2013.

Silva, D. N.; Balbo, A. R. Um Método Primal-Dual de Pontos Interiores Barreira Logarítmica Modificada com Aproximantes Spline aplicado ao Problema de Despacho Econômico com Ponto de Válvula. In: *Congresso de Matemática Aplicada e Computacional – CMAC Sudeste*, Bauru – SP, 2013.

Silva, D. N.; Balbo, A. R. Um Método Primal-Dual de Pontos Interiores com Estratégia de Convergência Global em Problemas de Minimização de Funções com Termos Modulares. In: *XXV Semana da Licenciatura em Matemática – SELMAT*, Bauru – SP, 2013.

Silva, D. N.; Balbo, A. R. Método Previsor-Corretor Primal-Dual de Pontos Interiores, com Estratégia de Barreira Modificada e Convergência Global, em Problemas de Despacho com Restrição de Segurança. In: *II Seminário da Pós-Graduação em Engenharia Elétrica*, Bauru – SP, 2012.

Referências Bibliográficas

- Adaryani, M. R. & Karami, A. (2013). Artificial bee colony algorithm for solving multi-objective optimal power flow problem. *International Journal of Electrical Power & Energy Systems*, 53(0), 219–230.
- Adler, I., Resende, M., Veiga, G., & Karmarkar, N. (1989). An implementation of karmarkar's algorithm for linear programming. *Mathematical Programming*, 44(1-3), 297–335.
- Afzalan, E. & Joorabian, M. (2013). Emission, reserve and economic load dispatch problem with non-smooth and non-convex cost functions using epsilon-multi-objective genetic algorithm variable. *International Journal of Electrical Power & Energy Systems*, 52(0), 55–67.
- ANEEL (2005). *Atlas de Energia Elétrica do Brasil*. Brasília, 2 edition. <http://www.aneel.gov.br/aplicacoes/atlas/pdf/atlas2005.zip>.
- ANEEL (2008). *Atlas de Energia Elétrica do Brasil*. Brasília, 3 edition. <http://www.aneel.gov.br/arquivos/PDF/atlas3ed.pdf>.
- Aoki, K. & Satoh, T. (1984). New algorithms for classic economic load dispatch. *IEEE Transactions on Power Apparatus and Systems*, PAS-103(6), 1423–1431.
- Arul, R., Ravi, D. G., & Velusami, D. S. (2011). Non-convex economic dispatch with heuristic load patterns using harmony search algorithm. *International Journal of Computer Applications*, 16(1), 26–33. Published by Foundation of Computer Science.
- Arul, R., Ravi, G., & Velusami, S. (2013). Non-convex economic dispatch with heuristic load patterns, valve point loading effect, prohibited operating zones, ramp-rate limits and spinning reserve constraints using harmony search algorithm. *Electrical Engineering*, 95(1), 53–61.
- Balamurugan, K., Muralisachithnndam, R., & Krishnan, S. R. (2014). Differential evolution based solution for combined economic and emission power dispatch with valve loading effect. *International Journal on Electrical Engineering and Informatics*, 6(1), 74–92.

- Balbo, A. R., Souza, M. A. S., Baptista, E. C., & Nepomuceno, L. (2012). Predictor-corrector primal-dual interior point method for solving economic dispatch problems: A postoptimization analysis. *Mathematical Problems in Engineering*, 2012, 1–26.
- Barnes, E. R. (1986). A variation on karmarkar’s algorithm for solving linear programming problems. *Mathematical Programming*, 36(2), 174–182.
- Basu, M. & Chowdhury, A. (2013). Cuckoo search algorithm for economic dispatch. *Energy*, 60(0), 99–108.
- Bazaraa, M. S., Sherali, H. D., & Shetty, C. M. (1979). *Nonlinear Programming – Theory and Algorithms*. New York: John Wiley & Sons.
- Benasla, L., Belmadani, A., & Rahli, M. (2014). Spiral optimization algorithm for solving combined economic and emission dispatch. *International Journal of Electrical Power & Energy Systems*, 62(0), 163–174.
- Benson, H. Y., Shanno, D. F., & Vanderbei, R. J. (2004). Interior-point methods for nonconvex nonlinear programming: Jamming and numerical testing. *Mathematical Programming*, 99(1), 35–48.
- Bertsekas, D. P. (1999). *Nonlinear Programming*. Massachusetts: Athena Scientific.
- Bhagwan Das, D. & Patvardhan, C. (1998). A new approach for security constrained economic emission dispatch in power systems. In *TENCON ’98. 1998 IEEE Region 10 International Conference on Global Connectivity in Energy, Computer, Communication and Control*, volume 2 (pp. 470–473 vol.2).
- Bhattacharjee, K., Bhattacharya, A., & nee Dey, S. H. (2014). Solution of economic emission load dispatch problems of power systems by real coded chemical reaction algorithm. *International Journal of Electrical Power & Energy Systems*, 59(0), 176–187.
- Bhattacharya, A. & Chattopadhyay, P. (2010). Hybrid differential evolution with biogeography-based optimization for solution of economic load dispatch. *IEEE Transactions on Power Systems*, 25(4), 1955–1964.
- Binetti, G., Davoudi, A., Naso, D., Turchiano, B., & Lewis, F. (2014). A distributed auction-based algorithm for the nonconvex economic dispatch problem. *IEEE Transactions on Industrial Informatics*, 10(2), 1124–1132.
- Borckmans, P. B., Selvan, S. E., Boumal, N., & Absil, P.-A. (2014). A riemannian subgradient algorithm for economic dispatch with valve-point effect. *Journal of Computational and Applied Mathematics*, 255(0), 848–866.

- Botta, A. C. (2009). *Um Algoritmo de Newton de Ponto Interior e Aplicações na Função Eletromagnética de Metais*. PhD thesis, Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia (COPPE) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.
- Brasil (2000). Relatório e Pareceres Prévios do TCU. <http://www.cgu.gov.br/publicacoes/prestacaocontaspresidente/RelatorioPareceresTCU/RPP2000.pdf>. Acessado em: 12 de julho de 2014.
- Brasil (2001). Relatório e Pareceres Prévios sobre as Contas do Governo da República. <http://www.cgu.gov.br/publicacoes/prestacaocontaspresidente/RelatorioPareceresTCU/RPP2001.pdf>. Acessado em: 12 de julho de 2014.
- Brasil (2007). *Plano Nacional de Energia 2030*. Empresa de Pesquisa Energética.
- Brasil (2012). *Anuário estatístico de energia elétrica 2012*. Empresa de Pesquisa Energética, Ministério de Minas e Energia.
- Brasil (2013). *Plano Decenal de Expansão de Energia 2021*. Empresa de Pesquisa Energética.
- Brasil (2014). *Plano Decenal de Expansão de Energia 2022*. Empresa de Pesquisa Energética.
- Breda, J. C. (2013). *Modelo de leilão multiperíodo para sistemas hidrotérmicos com representação da transmissão*. Master's thesis, Faculdade de Engenharia de Bauru – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Bauru, SP.
- Breitfeld, M. G. & Shanno, D. F. (1996). Computational experience with penalty-barrier methods for nonlinear programming. *Annals of Operations Research*, 62(1), 439–463.
- Bunch, J. R. & Kaufman, L. (1977). Some stable methods for calculating inertia and solving symmetric linear systems. *Mathematics of Computation*, 31(137), 163–179.
- Bunch, J. R. & Parlett, B. N. (1971). Direct methods for solving symmetric indefinite systems of linear equations. *SIAM Journal on Numerical Analysis*, 8(4), 639–655.
- Campanha, P. S. (2011). *Uma Adaptação do Método Barreira Penalidade Quasi-Newton ao Problema de Fluxo de Potência Ótimo*. Master's thesis, Faculdade de Engenharia de Bauru – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Bauru, SP.
- Carpentier, J. L. (1962). Contribution to the economic dispatch. *Bulletin de la Société Française des Electriciens*, 3(8), 431–447. In French.

- Carroll, C. W. & Fiacco, A. V. (1961). The created response surface technique for optimizing nonlinear restrained systems. *Operations Research*, 9(2), 169–185.
- Carvalho, S. M. S. & Oliveira, A. R. L. (2012). Interior-point methods applied to the pre-dispatch problem of a hydroelectric system with scheduled line manipulations. *American Journal of Operations Research*, 2(2), 266–271.
- Chen, C. & Mangasarian, O. (1996). A class of smoothing functions for nonlinear and mixed complementarity problems. *Computational Optimization and Applications*, 5(2), 97–138.
- Chen, X. (2012). Smoothing methods for nonsmooth, nonconvex minimization. *Mathematical Programming*, 134(1), 71–99.
- Coelho, L. S., Bora, T. C., & Mariani, V. C. (2014). Differential evolution based on truncated Lévy-type flights and population diversity measure to solve economic load dispatch problems. *International Journal of Electrical Power & Energy Systems*, 57(0), 178–188.
- Coelho, L. S. & Mariani, V. C. (2008). Particle swarm approach based on quantum mechanics and harmonic oscillator potential well for economic load dispatch with valve-point effects. *Energy Conversion and Management*, 49(11), 3080–3085. Special Issue 3rd International Conference on Thermal Engineering: Theory and Applications.
- Dikin, I. I. (1967). Iterative solution of problems of linear and quadratic programming. *Doklady Akademii Nauk*, 8, 747–748.
- Djukanovic, M., Calovic, M., Milosevic, B., & Sobajic, D. (1996). Neural-net based real-time economic dispatch for thermal power plants. *IEEE Transactions on Energy Conversion*, 11(4), 755–761.
- Dopazo, J. F., Klitin, O. A., Stagg, G. W., & Watson, M. (1967). An optimization technique for real and reactive power allocation. *Proceedings of the IEEE*, 55(11), 1877–1885.
- Druinsky, A. & Toledo, S. (2011). The growth-factor bound for the Bunch-Kaufman factorization is tight. *SIAM Journal on Matrix Analysis and Applications*, 32(3), 928–937.
- El-Gallad, A., El-Hawary, M., Sallam, A., & Kalas, A. (2002). Particle swarm optimizer for constrained economic dispatch with prohibited operating zones. In *Canadian Conference on Electrical and Computer Engineering, 2002. IEEE CCECE 2002.*, volume 1 (pp. 78–81 vol.1).

- Facchinei, F., Jiang, H., & Qi, L. (1999). A smoothing method for mathematical programs with equilibrium constraints. *Mathematical Programming*, 85(1), 107–134.
- Fahmideh-Vojdani, A. & Galiana, F. (1980). Economic dispatch with generation constraints. *IEEE Transactions on Automatic Control*, 25(2), 213–217.
- Fang, S. C. & Puthenpura, S. (1993). *Linear optimization and extensions: theory and algorithms*. New Jersey: Prentice-Hall.
- Ferreira, E. C. (2013). *Uma investigação da influência do parâmetro de barreira na solução do problema de fluxo de potência ótimo*. Master's thesis, Faculdade de Engenharia de Bauru – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Bauru, SP.
- Frisch, K. R. (1955). *The logarithmic potential method of convex programming*. Technical report, University Institute of Economics, Oslo, Norway.
- Furnas (2014). Sistema Furnas de geração e transmissão. <http://www.furnas.com.br/hotsites/sistemafurnas/>. Acessado em: 14 de julho de 2014.
- George, E. E. (1943). Intrasystem transmission losses. *Electrical Engineering*, 62(3), 153–158.
- Gonzaga, C. C. (1990). Polynomial affine algorithms for linear programming. *Mathematical Programming*, 49(1-3), 7–21.
- Gonçalves, R., Almeida, C., Goldbarg, M., Goldbarg, E., & Delgado, M. (2013). Improved cultural immune systems to solve the economic load dispatch problems. In *IEEE Congress on Evolutionary Computation (CEC) 2013* (pp. 621–628).
- Granville, S. (1994). Optimal reactive dispatch through interior point methods. *IEEE Transactions on Power Systems*, 9(1), 136–146.
- Griva, I. (2004). Numerical experiments with an interior-exterior point method for nonlinear programming. *Computational Optimization and Applications*, 29(2), 173–195.
- Griva, I., Nash, S. G., & Sofer, A. (2009). *Linear and Nonlinear Optimization*. SIAM.
- Griva, I. & Polyak, R. A. (2006). Primal-dual nonlinear rescaling method with dynamic scaling parameter update. *Mathematical Programming*, 106(2), 237–259.
- Happ, H. H. (1977). Optimal power dispatch – A comprehensive survey. *IEEE Transactions on Power Apparatus and Systems*, 96(3), 841–854.
- Hardiansyah (2013). Solving economic dispatch problem with valve-point effect using a modified ABC algorithm. *International Journal of Electrical and Computer Engineering*, 3(3), 377–385.

- He, D., Wang, F., & Mao, Z. (2008). A hybrid genetic algorithm approach based on differential evolution for economic dispatch with valve-point effect. *International Journal of Electrical Power & Energy Systems*, 30(1), 31–38.
- Hemamalini, S. & Simon, S. P. (2010). Dynamic economic dispatch with valve-point effect using Maclaurin series based lagrangian method. *International Journal of Computer Applications*, 1(17), 60–67. Published By Foundation of Computer Science.
- Hemamalini, S. & Simon, S. P. (2011). Dynamic economic dispatch using artificial bee colony algorithm for units with valve-point effect. *European Transactions on Electrical Power*, 21(1), 70–81.
- International Energy Agency (2013). *Key World Energy Statistics 2013*. OECD/IEA.
- Irisarri, G., Kimball, L., Clements, K., Bagchi, A., & Davis, P. (1998). Economic dispatch with network and ramping constraints via interior point methods. *IEEE Transactions on Power Systems*, 13(1), 236–242.
- Jasper, J. & Victoire, T. A. A. (2014). Dispatching a 19-unit Indian utility system using a refined differential evolution algorithm. *Mathematical Problems in Engineering*, 2014, 1–12.
- Jayabarathi, T., Yazdani, A., Ramesh, V., & Raghunathan, T. (2014). Combined heat and power economic dispatch problem using the invasive weed optimization algorithm. *Frontiers in Energy*, 8(1), 25–30.
- Jeddi, B. & Vahidinasab, V. (2014). A modified harmony search method for environmental/economic load dispatch of real-world power systems. *Energy Conversion and Management*, 78(0), 661–675.
- Jensen, D. L. & Polyak, R. (1994). The convergence of a modified barrier method for convex programming. *IBM Journal of Research and Development*, 38(3), 307–321.
- Jiang, S., Ji, Z., & Shen, Y. (2014). A novel hybrid particle swarm optimization and gravitational search algorithm for solving economic emission load dispatch problems with various practical constraints. *International Journal of Electrical Power & Energy Systems*, 55(0), 628–644.
- Junior, J. R. P. (2010). *Resolução do Problema de Empacotamento de Circunferências na Superfície de uma Esfera utilizando Suavização Hiperbólica*. Master’s thesis, Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia (COPPE) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.

- Karmarkar, N. (1984). A new polynomial time algorithm for linear programming. *Combinatorica*, 4(4), 373–395.
- Khamsawang, S. & Jiriwibhakorn, S. (2010). DSPSO–TSA for economic dispatch problem with nonsmooth and noncontinuous cost functions. *Energy Conversion and Management*, 51(2), 365–375.
- Kim, J., Kim, C. S., & Geem, Z. W. (2014). A memetic approach for improving minimum cost of economic load dispatch problems. *Mathematical Problems in Engineering*, 2014, 1–11.
- Klee, V. & Minty, G. J. (1972). How good is the simplex algorithm? In *Inequalities, III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969; dedicated to the memory of Theodore S. Motzkin)* (pp. 159–175). Academic Press, New York.
- Kojima, M., Megiddo, N., & Mizuno, S. (1993). A primal-dual infeasible-interior-point algorithm for linear programming. *Mathematical Programming*, 61(1-3), 263–280.
- Kojima, M., Mizuno, S., & Yoshise, A. (1989a). A polynomial-time algorithm for a class of linear complementary problems. *Mathematical Programming*, 44(1), 1–26.
- Kojima, M., Mizuno, S., & Yoshise, A. (1989b). A primal-dual interior point algorithm for linear programming. In *Progress in Mathematical Programming* (pp. 29–47).
- Kothari, D. P. & Dhillon, J. S. (2006). *Power System Optimization*. New Delhi: Prentice-Hall.
- Lage, G. G. (2013). *O Fluxo de Potência Ótimo Reativo com Variáveis de Controle Discretas e Restrições de Atuação de Dispositivos de Controle de Tensão*. PhD thesis, Escola de Engenharia de São Carlos – Universidade de São Paulo, São Carlos, SP.
- Lavaei, J. & Low, S. (2012). Zero duality gap in optimal power flow problem. *IEEE Transactions on Power Systems*, 27(1), 92–107.
- Lee, K., Ortiz, J. L., Mohtadi, M. A., & Park, Y. (1988). Optimal operation of large-scale power systems. *IEEE Transactions on Power Systems*, 3(2), 413–420.
- Lee, Y.-J. & Mangasarian, O. (2001). SSVM: A smooth support vector machine for classification. *Computational Optimization and Applications*, 20(1), 5–22.
- Lima, C. (2013). *Métodos híbridos de pontos interiores e de programação inteira 0-1 para problemas de custo de colheita da cana-de-açúcar e de custo de coleta e geração de energia relacionados à sua biomassa*. Master’s thesis, Faculdade de Engenharia de Bauru – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Bauru, SP.

- Lin, C., Chen, S., & Huang, C. (1992). A direct newton-raphson economic dispatch. *IEEE Transactions on Power Systems*, 7(3), 1149–1154.
- Lin, W.-M., Cheng, F.-S., & Tsay, M.-T. (2002). An improved tabu search for economic dispatch with multiple minima. *Power Systems, IEEE Transactions on*, 17(1), 108–112.
- Lu, P., Zhou, J., Zhang, H., Zhang, R., & Wang, C. (2014). Chaotic differential bee colony optimization algorithm for dynamic economic dispatch problem with valve-point effects. *International Journal of Electrical Power & Energy Systems*, 62(0), 130–143.
- Luenberger, D. G. & Ye, Y. (2008). *Linear and Nonlinear Programming*. Springer.
- Ma, C. (2008). A globally convergent Levenberg–Marquardt method for the least -norm solution of nonlinear inequalities. *Applied Mathematics and Computation*, 206(1), 133–140.
- Malik, T. N., Ul Asar, A., Wyne, M. F., & Akhtar, S. (2010). A new hybrid approach for the solution of nonconvex economic dispatch problem with valve-point effects. *Electric Power Systems Research*, 80(9), 1128–1136.
- Mandal, B., Roy, P. K., & Mandal, S. (2014). Economic load dispatch using krill herd algorithm. *International Journal of Electrical Power & Energy Systems*, 57(0), 1–10.
- Manteaw, E. D. & Odero, N. A. (2012). Combined economic and emission dispatch solution using ABC_PSO hybrid algorithm with valve point loading effect. *International Journal of Scientific and Research Publications*, 2(12), 1–9.
- Martínez, J. M. & Santos, S. A. (1998). Métodos computacionais de otimização. www.ime.unicamp.br/~martinez/mslivro.pdf. Acessado em: 12 de julho de 2014.
- Matsumura, E. H. (2003). *Ensaio sobre Liberalização, Regulação e Investimento em Sistemas Hidrotérmicos*. PhD thesis, Departamento de Economia – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, RJ.
- Megiddo, N. & Shub, M. (1989). Boundary behavior of interior point algorithms in linear programming. *Mathematics of Operations Research*, 14(1), 97–146.
- Mehrotra, S. (1992). On the implementation of a primal-dual interior point method. *SIAM Journal on Optimization*, 2(4), 575–601.
- Mehrotra, S. & Sun, J. (1990). An algorithm for convex quadratic programming that requires $\mathcal{O}(n^{3.5}L)$ arithmetic operations. *Mathematics of Operations Research*, 15(2), 342–363.

- Melman, A. & Polyak, R. (1996). The Newton modified barrier method for QP problems. *Annals of Operations Research*, 62(1), 465–519.
- Melo, T. M. M., Matias, J. L. H., & Monteiro, M. T. T. (2013). Numerical optimization experiments using the hyperbolic smoothing strategy to solve MPCC. In *Proceedings of the 13th International Conference on Computational and Mathematical Methods in Science and Engineering, CMMSE 2013*.
- Meng, K., Wang, H. G., Dong, Z., & Wong, K.-P. (2010). Quantum-inspired particle swarm optimization for valve-point economic load dispatch. *IEEE Transactions on Power Systems*, 25(1), 215–222.
- Miller, R. H. (1987). *Operação de Sistemas de Potência*. Rio de Janeiro: McGraw-Hill.
- Mizuno, S. (1994). Polynomiality of infeasible-interior-point algorithms for linear programming. *Mathematical Programming*, 67(1-3), 109–119.
- Monteiro, R. & Adler, I. (1989). Interior path following primal-dual algorithms. part i: Linear programming. *Mathematical Programming*, 44(1-3), 27–41.
- Monteiro, R. D. C., Adler, I., & Resende, M. C. (1990). A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension. *Mathematics of Operations Research*, 15(2), 191–214.
- Monticelli, A. J. (1983). *Fluxo de Carga em Redes de Energia Elétrica*. São Paulo: Edgard Blücher.
- Moradi-Dalvand, M., Mohammadi-Ivatloo, B., Najafi, A., & Rabiee, A. (2012). Continuous quick group search optimizer for solving non-convex economic dispatch problems. *Electric Power Systems Research*, 93(0), 93–105.
- Murray, W. (1971). Analytical expressions for the eigenvalues and eigenvectors of the hessian matrices of barrier and penalty functions. *Journal of Optimization Theory and Applications*, 7(3), 189–196.
- Nash, S. G., Polyak, R., & Sofer, A. (1994). A numerical comparison of barrier and modified barrier methods for large-scale bound-constrained optimization. In W. Hager, D. Hearn, & P. Pardalos (Eds.), *Large Scale Optimization* (pp. 319–338). Springer US.
- Niknam, T. (2010). A new fuzzy adaptive hybrid particle swarm optimization algorithm for non-linear, non-smooth and non-convex economic dispatch problem. *Applied Energy*, 87(1), 327–339.
- Nocedal, J. & Waltz, R. A. (2005). *Adaptive Barrier Update Strategies for Nonlinear Interior Methods*. Technical report.

- Nocedal, J. & Wright, S. J. (2006). *Numerical Optimization*. New York: Springer.
- Noman, N. & Iba, H. (2008). Differential evolution for economic load dispatch problems. *Electric Power Systems Research*, 78(8), 1322–1331.
- Oliveira, K. W. R. C. B., Nascimento, N. T., & Saavedra, O. R. (2008). Uma abordagem via estratégias evolutivas para o despacho econômico considerando restrições de geração. *IEEE Latin America Transactions*, 6(1), 42–50.
- Palma-Benhke, R., Philpott, A., Jofré, A., & Cortés-Carmona, M. (2013). Modelling network constrained economic dispatch problems. *Optimization and Engineering*, 14(3), 417–430.
- Park, J.-B., Lee, K.-S., Shin, J.-R., & Lee, K. Y. (2005). A particle swarm optimization for economic dispatch with nonsmooth cost functions. *IEEE Transactions on Power Systems*, 20(1), 34–42.
- Pinheiro, R. B. N. (2012). *Um método previsor-corretor primal-dual de pontos interiores barreira logarítmica modificada, com estratégias de global e de ajuste cúbico, para problemas de programação não-linear e não-convexa*. Master's thesis, Faculdade de Engenharia de Bauru – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Bauru, SP.
- Pinheiro, R. B. N., Balbo, A. R., Baptista, E. C., & Nepomuceno, L. (2014). Interior-exterior point method with global convergence strategy for solving the optimal reactive power flow problem. Aceito para publicação na revista *International Journal of Electrical Power & Energy Systems*.
- Polyak, R. (1992). Modified barrier functions (theory and methods). *Mathematical Programming*, 54(1-3), 177–222.
- Polyak, R. & Teboulle, M. (1997). Nonlinear rescaling and proximal-like methods in convex optimization. *Mathematical Programming*, 76(2), 265–284.
- Polyak, R. A. (2002). Nonlinear rescaling vs. smoothing technique in convex optimization. *Mathematical Programming*, 92(2), 197–235.
- Polyak, R. A. (2008). Primal–dual exterior point method for convex optimization. *Optimization Methods and Software*, 23(1), 141–160.
- Polyak, R. A. (2014). Lagrangian transformation and interior ellipsoid methods in convex optimization. *Journal of Optimization Theory and Applications*, (pp. 1–27).

- Pothiya, S., Ngamroo, I., & Kongprawechnon, W. (2010). Ant colony optimisation for economic dispatch problem with non-smooth cost functions. *International Journal of Electrical Power & Energy Systems*, 32(5), 478–487.
- Qi, L. & Sun, D. (2002). Smoothing functions and smoothing Newton method for complementarity and variational inequality problems. *Journal of Optimization Theory and Applications*, 113(1), 121–147.
- Ramesh, B., Chandra Jagan Mohan, V., & Veera Reddy, V. (2013). Application of bat algorithm for combined economic load and emission dispatch. *Journal of Electrical Engineering*, 13(2), 214–219.
- Ravi, G., Chakrabarti, R., & Choudhuri, S. (2006). Nonconvex economic dispatch with heuristic load patterns using improved fast evolutionary program. *Electric Power Components and Systems*, 34(1), 37–45.
- Renegar, J. (1988). A polynomial-time algorithm, based on newton’s method, for linear programming. *Mathematical Programming*, 40(1-3), 59–93.
- Rodrigues, N. M. (2007). *Um Algoritmo Cultural para Problemas de Despacho de Energia Elétrica*. Master’s thesis, Departamento de Informática – Universidade Estadual de Maringá, Maringá, PR.
- Sa-ngiamvibool, W., Pothiya, S., & Ngamroo, I. (2011). Multiple tabu search algorithm for economic dispatch problem considering valve-point effects. *International Journal of Electrical Power & Energy Systems*, 33(4), 846–854.
- Saadat, H. (1999). *Power System Analysis*. New York: McGraw-Hill.
- Sajjadi, S. M., Yazdankhah, A. S., & Ferdowsi, F. (2012). A new gumption approach for economic dispatch problem with losses effect based on valve-point active power. *Electric Power Systems Research*, 92(0), 81–86.
- Samed, M. M. A. (2004). *Um Algoritmo Genético Híbrido Co-Evolutivo para Resolver Problemas de Despacho*. PhD thesis, Departamento de Engenharia Química – Universidade Estadual de Maringá, Maringá, PR.
- Santos, A. B. A. (1997). *Problemas de Programação Não-Diferenciável: Uma Metodologia de Suavização*. Master’s thesis, Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia (COPPE) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.
- Schmidt, M., Fung, G., & Rosales, R. (2007). Fast optimization methods for L1 regularization: A comparative study and two new approaches. In J. Kok, J. Koronacki, R.

- Mantaras, S. Matwin, D. Mladenič, & A. Skowron (Eds.), *Machine Learning: ECML 2007*, volume 4701 of *Lecture Notes in Computer Science* (pp. 286–297). Springer Berlin Heidelberg.
- Selvakumar, A. I. (2011). Enhanced cross-entropy method for dynamic economic dispatch with valve-point effects. *International Journal of Electrical Power & Energy Systems*, 33(3), 783–790.
- Selvakumar, A. I. & Thanushkodi, K. (2009). Optimization using civilized swarm: Solution to economic dispatch with multiple minima. *Electric Power Systems Research*, 79(1), 8–16.
- Shanno, D. F. & Vanderbei, R. J. (2000). Interior-point methods for nonconvex nonlinear programming: orderings and higher-order methods. *Mathematical Programming*, 87(2), 303–316.
- Shaw, J. J., Gendron, R. F., & Bertsekas, D. (1985). Optimal scheduling of large hydrothermal power systems. *IEEE Transactions on Power Apparatus and Systems*, PAS-104(2), 286–294.
- Sinha, N., Chakrabarti, R., & Chattopadhyay, R. (2003). Evolutionary programming techniques for economic load dispatch. *IEEE Transactions on Evolutionary Computation*, 7(1), 83–94.
- Sinha, N., Hazarika, K., Paul, S., Shekhar, H., & Karmakar, A. (2010). Improved real quantum evolutionary algorithm for optimum economic load dispatch with non-convex loads. In B. K. Panigrahi, S. Das, P. N. Suganthan, & S. S. Dash (Eds.), *Swarm, Evolutionary, and Memetic Computing*, volume 6466 of *Lecture Notes in Computer Science* (pp. 689–700). Springer Berlin Heidelberg.
- Smil, V. (2004). *World history and energy*, volume 6 of *Encyclopedia of Energy*. Elsevier.
- Soler, E. M. (2011). *Resolução do problema de Fluxo de Potência Ótimo com variáveis de controle discretas*. PhD thesis, Escola de Engenharia de São Carlos – Universidade de São Paulo, São Carlos, SP.
- Soler, E. M., Asada, E. N., & Costa, G. R. M. (2013). Penalty-based nonlinear solver for optimal reactive power dispatch with discrete controls. *IEEE Transactions on Power Systems*, 28(3), 2174–2182.
- Somuah, C. & Khunaizi, N. (1990). Application of linear programming redispatch technique to dynamic generation allocation. *IEEE Transactions on Power Systems*, 5(1), 20–26.

- Song, Y., Chou, C., & Stonham, T. (1999). Combined heat and power economic dispatch by improved ant colony search algorithm. *Electric Power Systems Research*, 52(2), 115 – 121.
- Sousa, A. A. (2008). *Fluxo de Potência Ótimo Globalmente Convergente Utilizando Métodos de Pontos Interiores com Estratégias de Região de Confiança*. PhD thesis, Centro de Tecnologia e Geociências – Universidade Federal de Pernambuco, Recife, PE.
- Sousa, V. A. (2006). *Resolução do Problema de Fluxo de Potência Ótimo Reativo Via Método da Função Lagrangiana Barreira Modificada*. PhD thesis, Escola de Engenharia de São Carlos – Universidade de São Paulo, São Carlos, SP.
- Sousa, V. A., Baptista, E. C., & Costa, G. R. M. (2012). Optimal reactive power flow via the modified barrier lagrangian function approach. *Electric Power Systems Research*, 84(1), 159–164.
- Souza, M., Xavier, A. E., Lavor, C., & Maculan, N. (2011). Hyperbolic smoothing and penalty techniques applied to molecular structure determination. *Operations Research Letters*, 39(6), 461–465.
- Souza, M. F. (2010). *Suavização Hiperbólica Aplicada à Otimização de Geometria Molecular*. PhD thesis, Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia (COPPE) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.
- Stanzani, A. L. (2012). *Método predictor-corretor primal-dual de pontos interiores em problemas multiobjetivo de despacho econômico e ambiental*. Master’s thesis, Faculdade de Engenharia de Bauru – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Bauru, SP.
- Stanzani, A. L., Balbo, A. R., Nepomuceno, L., & Baptista, E. C. (2014). Solving the multiobjective environmental/economic dispatch problem using weighted sum and ϵ -constraint strategies and a predictor-corrector primal-dual interior point method. *Journal of Control, Automation and Electrical Systems*.
- Steinberg, M. J. & Smith, T. H. (1934). The theory of incremental rates and their practical application to load division – part i. *Transactions of the American Institute of Electrical Engineers*, 53(3), 432–445.
- Steinberg, M. J. & Smith, T. H. (1943). *Economy Loading of Power Plants and Electric Systems*. John Wiley & Sons.
- Tiwari, S., Kumar, A., Chaurasia, G. S., & Sirohi, G. S. (2013). Economic load dispatch using particle swarm optimization. *International Journal of Application or Innovation in Engineering & Management*, 2(4), 476–485.

- Torres, G. L. (1998). *Nonlinear Optimal Power Flow by Interior and Non-Interior Point Methods*. PhD thesis, University of Waterloo, Waterloo, Ontario, Canada.
- Vaidya, P. (1990). An algorithm for linear programming which requires $o(((m+n)n^2 + (m+n)^{1.5}n)l)$ arithmetic operations. *Mathematical Programming*, 47(1-3), 175–201.
- Vanderbei, R. J., Meketon, M. S., & Freedman, B. A. (1986). A modification of karmarkar’s linear programming algorithm. *Algorithmica*, 1(1-4), 395–407.
- Vanderbei, R. J. & Shanno, D. F. (1999). An interior-point algorithm for nonconvex nonlinear programming. *Computational Optimization and Applications*, 13(1-3), 231–252.
- Vergílio, A. H. B. (2011). *Um modelo de pré-despacho hidrotérmico para mercados de energia*. Master’s thesis, Faculdade de Engenharia de Bauru – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Bauru, SP.
- Viana, A. & Pedroso, J. P. (2013). A new MILP-based approach for unit commitment in power production planning. *International Journal of Electrical Power & Energy Systems*, 44(1), 997–1005.
- Victoire, T. A. A. & Jeyakumar, A. E. (2004). Hybrid PSO–SQP for economic dispatch with valve-point effect. *Electric Power Systems Research*, 71(1), 51–59.
- Victoire, T. A. A. & Jeyakumar, A. E. (2005). A modified hybrid EP–SQP approach for dynamic dispatch with valve-point effect. *International Journal of Electrical Power & Energy Systems*, 27(8), 594–601.
- Vishwakarma, K. K., Dubey, H. M., Pandit, M., & Panigrahi, B. K. (2012). Simulated annealing approach for solving economic load dispatch problems with valve point loading effects. *International Journal of Engineering, Science and Technology*, 4(4), 60–72.
- Walters, D. C. & Sheble, G. B. (1993). Genetic algorithm solution of economic dispatch with valve point loading. *IEEE Transactions on Power Systems*, 8(3), 1325–1332.
- Wood, A. J. & Wollenberg, B. F. (1996). *Power Generation, Operation and Control*. New York: John Wiley & Sons.
- Wright, M. (1994). Some properties of the hessian of the logarithmic barrier function. *Mathematical Programming*, 67(1-3), 265–295.
- Wright, M. H. (2004). The interior-point revolution in optimization: History, recent developments, and lasting consequences. *Bulletin of the American Mathematical Society*, 42(1), 39–56.

- Wu, Y.-C., Debs, A. S., & Marsten, R. E. (1994). A direct nonlinear predictor-corrector primal-dual interior point algorithm for optimal power flows. *IEEE Transactions on Power Systems*, 9(2), 876–883.
- Wächter, A. & Biegler, L. T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1), 25–57.
- Xavier, A. & Oliveira, A. (2005). Optimal covering of plane domains by circles via hyperbolic smoothing. *Journal of Global Optimization*, 31(3), 493–504.
- Xavier, A. E. (2010). The hyperbolic smoothing clustering method. *Pattern Recognition*, 43(3), 731–737.
- Xavier, A. E. & Xavier, V. L. (2011). Solving the minimum sum-of-squares clustering problem by hyperbolic smoothing and partition into boundary and gravitational regions. *Pattern Recognition*, 44(1), 70–77.
- Xiong, G., Li, Y., Chen, J., Shi, D., & Duan, X. (2014). Polyphyletic migration operator and orthogonal learning aided biogeography-based optimization for dynamic economic dispatch with valve-point effects. *Energy Conversion and Management*, 80(0), 457–468.
- Xiong, G., Shi, D., & Duan, X. (2013). Multi-strategy ensemble biogeography-based optimization for economic dispatch problems. *Applied Energy*, 111(0), 801–811.
- Xu, S. (2001). Smoothing method for minimax problems. *Computational Optimization and Applications*, 20(3), 267–279.
- Yang, L., Chen, Y., Tong, X., & Deng, C. (2011). A new smoothing Newton method for solving constrained nonlinear equations. *Applied Mathematics and Computation*, 217(24), 9855–9863.
- Yang, X.-S., Hosseini, S. S. S., & Gandomi, A. H. (2012). Firefly algorithm for solving non-convex economic dispatch problems with valve loading effect. *Applied Soft Computing*, 12(3), 1180–1186.
- Yasar, C. & Özyön, S. (2011). A new hybrid approach for nonconvex economic dispatch problem with valve-point effect. *Energy*, 36(10), 5838–5845.
- Ye, Y. (1991). An $o(n^3)$ potential reduction algorithm for linear programming. *Mathematical Programming*, 50(2), 239–258.
- Yuan, X., Wang, L., Zhang, Y., & Yuan, Y. (2009). A hybrid differential evolution method for dynamic economic dispatch with valve-point effects. *Expert Systems with Applications*, 36(2, Part 2), 4042–4048.

- Zare, K., Haque, M. T., & Davoodi, E. (2012). Solving non-convex economic dispatch problem with valve point effects using modified group search optimizer method. *Electric Power Systems Research*, 84(1), 83–89.
- Zhang, Y., Gong, D., Geng, N., & Sun, X. (2014). Hybrid bare-bones PSO for dynamic economic dispatch with valve-point effects. *Applied Soft Computing*, 18(0), 248–260.

Apêndice A

Cálculo dos Coeficientes-B

O método apresentado a seguir para cálculo dos coeficientes-B, extraído de Saadat (1999), baseia-se no trabalho de Kron e que foi, posteriormente, adotado por Kirchmayer.

A potência complexa total injetada em uma barra i , denotada por S_i , pode ser expressa por:

$$S_i = P_i + jQ_i = V_i I_i^* \quad (\text{A.1})$$

em que:

V_i é a tensão na barra i ;

I_i^* é o conjugado da corrente injetada na barra i .

O somatório das potências sobre todas as barras fornece o valor total das perdas no sistema, isto é,

$$P_{\text{loss}} + jQ_{\text{loss}} = \sum_{i=1}^n V_i I_i^* = V_{\text{bus}}^t I_{\text{bus}}^* \quad (\text{A.2})$$

em que P_{loss} e Q_{loss} denotam as perdas de potência ativa e reativa no sistema, respectivamente. V_{bus} é o vetor coluna das tensões nodais e I_{bus} é o vetor coluna das correntes injetadas nas barras. De acordo com Monticelli (1983), as injeções de corrente podem ser obtidas através da matriz admitância nodal (Y_{bus}) e do vetor de tensões nodais, através da expressão

$$I_{\text{bus}} = Y_{\text{bus}} V_{\text{bus}} \quad (\text{A.3})$$

de modo que:

$$V_{\text{bus}} = Y_{\text{bus}}^{-1} I_{\text{bus}} \text{ se, e somente se, } V_{\text{bus}} = Z_{\text{bus}} I_{\text{bus}} \quad (\text{A.4})$$

A inversa da matriz admitância, $Z_{\text{bus}} = Y_{\text{bus}}^{-1}$, é a matriz impedância. Substituindo V_{bus}

da equação (A.4) em (A.2), obtém-se:

$$P_{\text{loss}} + jQ_{\text{loss}} = (Z_{\text{bus}} I_{\text{bus}})^t I_{\text{bus}}^* = I_{\text{bus}}^t Z_{\text{bus}}^t I_{\text{bus}}^* \quad (\text{A.5})$$

Desde que Z_{bus} é uma matriz simétrica, isto é, $Z_{\text{bus}} = Z_{\text{bus}}^t$, então as perdas totais tornam-se:

$$P_{\text{loss}} + jQ_{\text{loss}} = I_{\text{bus}}^t Z_{\text{bus}} I_{\text{bus}}^* \quad (\text{A.6})$$

A expressão em (A.6) pode ser representada da seguinte maneira:

$$P_{\text{loss}} + jQ_{\text{loss}} = \sum_{i=1}^n \sum_{j=1}^n I_i Z_{ij} I_j^* \quad (\text{A.7})$$

Novamente, tendo em vista a simetria da matriz impedância, temos que:

$$P_{\text{loss}} + jQ_{\text{loss}} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n Z_{ij} (I_i I_j^* + I_j I_i^*) \quad (\text{A.8})$$

O termo $I_i I_j^* + I_j I_i^*$ é um número real, de modo que, considerando $Z_{\text{bus}} = R_{\text{bus}} + jX_{\text{bus}}$, as perdas podem ser separadas em suas componentes real e imaginária, ou seja:

$$P_{\text{loss}} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n R_{ij} (I_i I_j^* + I_j I_i^*) = I_{\text{bus}}^t R_{\text{bus}} I_{\text{bus}}^* \quad (\text{A.9})$$

$$Q_{\text{loss}} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n X_{ij} (I_i I_j^* + I_j I_i^*) = I_{\text{bus}}^t X_{\text{bus}} I_{\text{bus}}^* \quad (\text{A.10})$$

A fim de obter a fórmula geral para as perdas de potência no sistema, definimos a corrente de carga total como sendo a soma das correntes de carga individuais, isto é,

$$I_{L,1} + I_{L,2} + \dots + I_{L,n} = I_D \quad (\text{A.11})$$

em que n é o número de barras do sistema, $I_{L,k}$ é a corrente de carga da barra k e I_D é a corrente de carga total. Assume-se que as correntes individuais variam como uma fração complexa da corrente de carga total, ou seja,

$$I_{L,k} = \ell_k I_D \text{ se, e somente se, } \ell_k = \frac{I_{L,k}}{I_D}, \quad k = 1, \dots, n \quad (\text{A.12})$$

Por facilidade de notação, assume-se que a barra 1 é a barra *slack*. Expandindo a primeira

linha em (A.4), obtém-se:

$$\begin{aligned} V_1 &= \sum_{k=1}^n Z_{1k} I_k = \sum_{k=1}^n Z_{1k} (I_{G,k} + I_{L,k}) \\ &= \sum_{k=1}^n Z_{1k} I_{G,k} + \sum_{k=1}^n Z_{1k} I_{L,k} \end{aligned} \quad (\text{A.13})$$

em que $I_{G,k}$ denota a corrente de geração da barra k . Substituindo (A.12) em (A.13):

$$\begin{aligned} V_1 &= \sum_{k=1}^n Z_{1k} I_{G,k} + I_D \sum_{k=1}^n \ell_k Z_{1k} \\ &= \sum_{k=1}^n Z_{1k} I_{G,k} + I_D T \end{aligned} \quad (\text{A.14})$$

em que

$$T = \sum_{k=1}^n \ell_k Z_{1k} \quad (\text{A.15})$$

Define-se I_0 como sendo a corrente fluindo a partir da barra 1 (*slack*), com as demais correntes nulas. Esta corrente é denominada *corrente de carga nula*, e é tal que:

$$V_1 = -Z_{11} I_0 \quad (\text{A.16})$$

Substituindo-se (A.16) em (A.14), e isolando I_D , obtém-se:

$$I_D = -\frac{1}{T} \sum_{k=1}^n Z_{1k} I_{G,k} - \frac{1}{T} Z_{11} I_0 \quad (\text{A.17})$$

Substituindo-se (A.17) em (A.12), as correntes de carga tornam-se:

$$I_{L,k} = -\frac{\ell_k}{T} \sum_{k=1}^n Z_{1k} I_{G,k} - \frac{\ell_k}{T} Z_{11} I_0 \quad (\text{A.18})$$

Defina

$$\rho_k = -\frac{\ell_k}{T} \quad (\text{A.19})$$

Então

$$I_{L,k} = \rho_k \sum_{k=1}^n Z_{1k} I_{G,k} + \rho_k Z_{11} I_0 \quad (\text{A.20})$$

Como

$$I_{\text{bus}} = I_G + I_L \quad (\text{A.21})$$

e, além disso,

$$I_G = \begin{pmatrix} I_{G,1} \\ I_{G,2} \\ \vdots \\ I_{G,n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} I_{G,1} \\ I_{G,2} \\ \vdots \\ I_{G,n} \\ I_0 \end{pmatrix} \quad (\text{A.22})$$

e, por (A.20),

$$I_L = \begin{pmatrix} I_{L,1} \\ I_{L,2} \\ \vdots \\ I_{L,n} \end{pmatrix} = \begin{pmatrix} \rho_1 Z_{11} & \rho_1 Z_{12} & \cdots & \rho_1 Z_{1n} & \rho_1 Z_{11} \\ \rho_2 Z_{11} & \rho_2 Z_{12} & \cdots & \rho_2 Z_{1n} & \rho_2 Z_{11} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_n Z_{11} & \rho_n Z_{12} & \cdots & \rho_n Z_{1n} & \rho_n Z_{11} \end{pmatrix} \begin{pmatrix} I_{G,1} \\ I_{G,2} \\ \vdots \\ I_{G,n} \\ I_0 \end{pmatrix} \quad (\text{A.23})$$

segue de (A.21) que:

$$I_{\text{bus}} = \begin{pmatrix} 1 + \rho_1 Z_{11} & \rho_1 Z_{12} & \cdots & \rho_1 Z_{1n} & \rho_1 Z_{11} \\ \rho_2 Z_{11} & 1 + \rho_2 Z_{12} & \cdots & \rho_2 Z_{1n} & \rho_2 Z_{11} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_n Z_{11} & \rho_n Z_{12} & \cdots & 1 + \rho_n Z_{1n} & \rho_n Z_{11} \end{pmatrix} \begin{pmatrix} I_{G,1} \\ I_{G,2} \\ \vdots \\ I_{G,n} \\ I_0 \end{pmatrix} \quad (\text{A.24})$$

Seja

$$A = \begin{pmatrix} 1 + \rho_1 Z_{11} & \rho_1 Z_{12} & \cdots & \rho_1 Z_{1n} & \rho_1 Z_{11} \\ \rho_2 Z_{11} & 1 + \rho_2 Z_{12} & \cdots & \rho_2 Z_{1n} & \rho_2 Z_{11} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_n Z_{11} & \rho_n Z_{12} & \cdots & 1 + \rho_n Z_{1n} & \rho_n Z_{11} \end{pmatrix} = \begin{pmatrix} \mathbf{a}^1 & \mathbf{a}^2 & \cdots & \mathbf{a}^n & \mathbf{a}^{n+1} \end{pmatrix} \quad (\text{A.25})$$

em que

$$\mathbf{a}^k = \begin{pmatrix} \rho_1 Z_{1k} \\ \rho_2 Z_{1k} \\ \vdots \\ 1 + \rho_k Z_{1k} \\ \vdots \\ \rho_n Z_{1k} \end{pmatrix} \leftarrow k\text{-ésima linha}, \quad k = 1, \dots, n \quad (\text{A.26})$$

e

$$\mathbf{a}^{n+1} = \begin{pmatrix} \rho_1 Z_{11} \\ \rho_2 Z_{11} \\ \vdots \\ \rho_n Z_{11} \end{pmatrix} \quad (\text{A.27})$$

Dentre as barras do sistema, podem ocorrer barras que não possuem gerador. Se uma barra k não possui gerador, então $I_{G,k} = 0$, de modo que essa corrente de geração não é necessária para determinar I_{bus} pela equação (A.24). Seja $\mathcal{B} = \{1, 2, \dots, n\}$ o conjunto dos índices das barras do sistema e $\mathcal{G} = \{g_1, g_2, \dots, g_{|\mathcal{G}|}\}$ o conjunto dos índices das barras geradoras de modo ordenado, isto é, $g_1 < g_2 < \dots < g_{|\mathcal{G}|}$. Tendo em vista as igualdades (A.24) e (A.25), tem-se que:

$$\begin{aligned} I_{\text{bus}} &= \sum_{k=1}^n \mathbf{a}^k I_{G,k} + \mathbf{a}^{n+1} I_0 = \sum_{k \in \mathcal{G}} \mathbf{a}^k I_{G,k} + \mathbf{a}^{n+1} I_0 \\ &= \mathbf{a}^{g_1} I_{G,g_1} + \mathbf{a}^{g_2} I_{G,g_2} + \dots + \mathbf{a}^{g_{|\mathcal{G}|}} I_{G,g_{|\mathcal{G}|}} + \mathbf{a}^{n+1} I_0 \\ &= \begin{pmatrix} \mathbf{a}^{g_1} & \mathbf{a}^{g_2} & \dots & \mathbf{a}^{g_{|\mathcal{G}|}} & \mathbf{a}^{n+1} \end{pmatrix} \begin{pmatrix} I_{G,g_1} \\ I_{G,g_2} \\ \vdots \\ I_{G,g_{|\mathcal{G}|}} \\ I_0 \end{pmatrix} \end{aligned} \quad (\text{A.28})$$

Definindo

$$C = \begin{pmatrix} \mathbf{a}^{g_1} & \mathbf{a}^{g_2} & \dots & \mathbf{a}^{g_{|\mathcal{G}|}} & \mathbf{a}^{n+1} \end{pmatrix} \quad (\text{A.29})$$

e

$$I_{\text{new}} = \begin{pmatrix} I_{G,g_1} \\ I_{G,g_2} \\ \vdots \\ I_{G,g_{|\mathcal{G}|}} \\ I_0 \end{pmatrix} \quad (\text{A.30})$$

tem-se que

$$I_{\text{bus}} = C I_{\text{new}} \quad (\text{A.31})$$

Substituindo-se (A.31) em (A.9), obtém-se

$$P_{\text{loss}} = (C I_{\text{new}})^t R_{\text{bus}} (C I_{\text{new}})^* = I_{\text{new}}^t C^t R_{\text{bus}} C^* I_{\text{new}}^* \quad (\text{A.32})$$

Se S_{G,g_i} é a potência complexa gerada, em uma barra de geração $g_i \in \mathcal{G}$, e o fator de

potência é constante, a corrente de geração é:

$$I_{G,g_i} = \frac{S_{G,g_i}^*}{V_{g_i}^*} = \frac{P_{G,g_i} - jQ_{G,g_i}}{V_{g_i}^*} = \left(\frac{1 - j \frac{Q_{G,g_i}}{P_{G,g_i}}}{V_{g_i}^*} \right) P_{G,g_i} \quad (\text{A.33})$$

ou

$$I_{G,g_i} = \psi_i P_{G,g_i} \quad (\text{A.34})$$

em que

$$\psi_i = \frac{1 - j \frac{Q_{G,g_i}}{P_{G,g_i}}}{V_{g_i}^*} \quad (\text{A.35})$$

Dessa forma, pode-se escrever:

$$\begin{pmatrix} I_{G,g_1} \\ I_{G,g_2} \\ \vdots \\ I_{G,g_{|\mathcal{G}|}} \\ I_0 \end{pmatrix} = \begin{pmatrix} \psi_1 & 0 & \cdots & 0 & 0 \\ 0 & \psi_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \psi_{|\mathcal{G}|} & 0 \\ 0 & 0 & 0 & 0 & I_0 \end{pmatrix} \begin{pmatrix} P_{G,g_1} \\ P_{G,g_2} \\ \vdots \\ P_{G,g_{|\mathcal{G}|}} \\ 1 \end{pmatrix} \quad (\text{A.36})$$

ou de modo compacto

$$I_{\text{new}} = \Psi P_{G1} \quad (\text{A.37})$$

em que

$$P_{G1} = \begin{pmatrix} P_{G,g_1} \\ P_{G,g_2} \\ \vdots \\ P_{G,g_{|\mathcal{G}|}} \\ 1 \end{pmatrix} \quad (\text{A.38})$$

e

$$\Psi = \begin{pmatrix} \psi_1 & 0 & \cdots & 0 & 0 \\ 0 & \psi_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \psi_{|\mathcal{G}|} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \quad (\text{A.39})$$

Substituindo-se (A.37) em (A.32), a equação das perdas se torna:

$$\begin{aligned} P_{\text{loss}} &= (\Psi P_{G1})^t C^t R_{\text{bus}} C^* (\Psi P_{G1})^* \\ &= P_{G1}^t \Psi^t C^t R_{\text{bus}} C^* \Psi^* P_{G1}^* \\ &= P_{G1}^t H P_{G1}^* \end{aligned} \quad (\text{A.40})$$

em que

$$H = \Psi^t C^t R_{\text{bus}} C^* \Psi^* \quad (\text{A.41})$$

Calculando o complexo conjugado da equação (A.40), utilizando o fato de que $P_{\text{loss}} = P_{\text{loss}}^*$ e $P_{G1} = P_{G1}^*$ e somando a expressão resultante com (A.40), conclui-se que:

$$P_{\text{loss}} = P_{G1}^t \left(\frac{H + H^*}{2} \right) P_{G1} \quad (\text{A.42})$$

Pode-se verificar que a matriz H é uma matriz hermitiana e, além disso, $\Re(H) = \frac{H+H^*}{2}$, ou seja, $\frac{H+H^*}{2}$ é a parte real da matriz H , de modo que podemos escrever:

$$\Re(H) = \frac{H + H^*}{2} = \begin{pmatrix} B_{g1,g1} & B_{g1,g2} & \cdots & B_{g1,g|G|} & \frac{B_{0,g1}}{2} \\ B_{g2,g1} & B_{g2,g2} & \cdots & B_{g2,g|G|} & \frac{B_{0,g2}}{2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ B_{g|G|,g1} & B_{g|G|,g2} & \cdots & B_{g|G|,g|G|} & \frac{B_{0,g|G|}}{2} \\ \frac{B_{0,g1}}{2} & \frac{B_{0,g2}}{2} & \cdots & \frac{B_{0,g|G|}}{2} & B_{0,0} \end{pmatrix} \quad (\text{A.43})$$

Então

$$P_{\text{loss}} = \begin{pmatrix} P_{G,g1} & P_{G,g2} & \cdots & P_{G,g|G|} & 1 \end{pmatrix} \begin{pmatrix} B_{g1,g1} & B_{g1,g2} & \cdots & B_{g1,g|G|} & \frac{B_{0,g1}}{2} \\ B_{g2,g1} & B_{g2,g2} & \cdots & B_{g2,g|G|} & \frac{B_{0,g2}}{2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ B_{g|G|,g1} & B_{g|G|,g2} & \cdots & B_{g|G|,g|G|} & \frac{B_{0,g|G|}}{2} \\ \frac{B_{0,g1}}{2} & \frac{B_{0,g2}}{2} & \cdots & \frac{B_{0,g|G|}}{2} & B_{0,0} \end{pmatrix} \begin{pmatrix} P_{G,g1} \\ P_{G,g2} \\ \vdots \\ P_{G,g|G|} \\ 1 \end{pmatrix} \quad (\text{A.44})$$

ou, de modo equivalente,

$$P_{\text{loss}} = \begin{pmatrix} P_{G,g1} & P_{G,g2} & \cdots & P_{G,g|G|} \end{pmatrix} \begin{pmatrix} B_{g1,g1} & B_{g1,g2} & \cdots & B_{g1,g|G|} \\ B_{g2,g1} & B_{g2,g2} & \cdots & B_{g2,g|G|} \\ \vdots & \vdots & \ddots & \vdots \\ B_{g|G|,g1} & B_{g|G|,g2} & \cdots & B_{g|G|,g|G|} \end{pmatrix} \begin{pmatrix} P_{G,g1} \\ P_{G,g2} \\ \vdots \\ P_{G,g|G|} \end{pmatrix} \quad (\text{A.45})$$

$$+ \begin{pmatrix} P_{G,g1} & P_{G,g2} & \cdots & P_{G,g|G|} \end{pmatrix} \begin{pmatrix} B_{0,g1} \\ B_{0,g2} \\ \vdots \\ B_{0,g|G|} \end{pmatrix} + B_{0,0}$$

Dessa forma, pode-se escrever:

$$P_{\text{loss}} = \sum_{i \in G} \sum_{j \in G} P_{G,i} B_{i,j} P_{G,j} + \sum_{i \in G} B_{0,i} P_{G,i} + B_{0,0} \quad (\text{A.46})$$

Portanto, o cálculo dos coeficientes-B pode ser sintetizado através do seguinte algoritmo:

Algoritmo 9: Coeficientes-B

- 1 Obtenha um estado de operação inicial factível (por exemplo, através de um fluxo de carga).
 - 2 A partir dos resultados, calcule as correntes de carga individuais, $I_{L,k}$, a corrente de carga total, I_D e as frações complexas de cada corrente, ℓ_k .
 - 3 Calcule a matriz Z_{bus} .
 - 4 Calcule as matrizes de transformação C , Ψ e H .
 - 5 Obtenha a parte real da matriz H , $\Re(H)$.
 - 6 Utilize a igualdade apresentada em (A.43) para determinar os coeficientes-B.
-

Cabe ressaltar, novamente, que os coeficientes-B são função do estado de operação do sistema. Se uma nova programação ou despacho de geração for realizado, de modo que não mude drasticamente a situação de operação inicial, pode-se assumir que tais coeficientes não variam de modo significativo e, portanto, ainda podem ser utilizados para calcular, de forma aproximada, as perdas para a nova situação de operação.

Apêndice B

Revisão da Teoria de Otimização

B.1 Otimização irrestrita

Em problemas de otimização, é de fundamental importância a capacidade de reconhecer se determinado ponto é ou não um mínimo local. A partir da definição de mínimo local, para tanto seria necessário investigar todos os pontos em alguma vizinhança ao redor do candidato a mínimo, a fim de garantir que nenhum deles possui valor menor para a função objetivo. Naturalmente, isto é inviável para a maioria dos problemas.

Contudo, quando a hipótese de diferenciabilidade da função é acrescentada, existem resultados muito mais práticos e eficazes para caracterização de mínimos locais. Estes resultados são, usualmente, denominados *condições necessárias e/ou suficientes de otimalidade*. As condições necessárias são aquelas que obrigatoriamente são satisfeitas por todos minimizadores locais, ao passo que as condições suficientes são aquelas cuja veracidade assegura que um ponto é mínimo local. Nesta seção, são apresentados os conceitos fundamentais de otimização irrestrita.

B.1.1 Conceitos básicos

De acordo com Bertsekas (1999), modelos matemáticos de otimização geralmente podem ser representados por um *conjunto de restrições*, $\mathcal{X} \subseteq \mathbb{R}^n$, e uma *função objetivo* ou *função custo*, $f : \mathcal{X} \rightarrow \mathbb{R}$. O conjunto $\mathcal{X}$ consiste de todas as variáveis de decisão $\mathbf{x}$ disponíveis e o valor $f(\mathbf{x})$ é uma medida escalar do quão desejável é a escolha de $\mathbf{x}$. Dessa forma, um problema de otimização pode ser escrito da seguinte maneira:

$$\begin{aligned} &\text{minimizar} && f(\mathbf{x}) \\ &\text{sujeito a} && \mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^n \end{aligned} \tag{B.1}$$

Através do problema (B.1), podemos definir o que são pontos e conjuntos factíveis.

Definição B.1. Seja um problema de otimização na forma (B.1). Dizemos que $\mathbf{x} \in \mathbb{R}^n$ é um *ponto factível* quando $\mathbf{x} \in \mathcal{X}$. O conjunto $\mathcal{X}$, que é a coleção de todos os pontos factíveis, é denominado *conjunto factível*.

No contexto dos problemas de otimização, são de fundamental importância as definições de mínimo local e mínimo global, apresentadas a seguir.

Definição B.2. Um ponto $\mathbf{x}^* \in \mathcal{X}$ é um *mínimo local* de f sobre o conjunto $\mathcal{X}$ se existe $\varepsilon > 0$ e uma bola aberta $B(\mathbf{x}^*, \varepsilon)$ tal que

$$f(\mathbf{x}^*) \leq f(\mathbf{x}), \quad \forall \mathbf{x} \in B(\mathbf{x}^*, \varepsilon) \cap \mathcal{X}.$$

Quando

$$f(\mathbf{x}^*) < f(\mathbf{x}), \quad \forall \mathbf{x} \in (B(\mathbf{x}^*, \varepsilon) - \{\mathbf{x}^*\}) \cap \mathcal{X}$$

dizemos que $\mathbf{x}^*$ é um *mínimo local estrito*.

Definição B.3. Um ponto $\mathbf{x}^* \in \mathcal{X}$ é um *mínimo global* de f sobre o conjunto $\mathcal{X}$ se

$$f(\mathbf{x}^*) \leq f(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{X}.$$

Quando

$$f(\mathbf{x}^*) < f(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{X} - \{\mathbf{x}^*\}$$

dizemos que $\mathbf{x}^*$ é um *mínimo global estrito*.

Em geral, é difícil afirmar se existe ou não um mínimo global para uma função f sobre um conjunto $\mathcal{X}$. Por exemplo, a função $f(x) = e^x$ não possui mínimo global nem local, quando definida sobre $\mathbb{R}$. Contudo, sob determinadas hipóteses acerca da função objetivo e da região factível (ou domínio da função objetivo) do problema de otimização, pode-se garantir a existência de um ponto de mínimo global.

Definição B.4. Seja $\mathcal{X} \subseteq \mathbb{R}$ e $f : \mathcal{X} \rightarrow \mathbb{R}$ uma função real. Dizemos que f é *coerciva* se

$$\lim_{k \rightarrow \infty} f(\mathbf{x}_k) = \infty$$

para toda sequência $\{\mathbf{x}_k\}$ de elementos de $\mathcal{X}$ tal que $\|\mathbf{x}_k\| \rightarrow \infty$, para alguma norma $\|\cdot\|^1$.

O teorema a seguir apresenta algumas condições sob as quais é possível garantir a existência de pelo menos um mínimo global e, portanto, é de grande importância na teoria de otimização.

Teorema B.1. (Teorema de Weierstrass) Sejam $\emptyset \neq \mathcal{X} \subseteq \mathbb{R}^n$ e $f : \mathcal{X} \rightarrow \mathbb{R}$ contínua sobre $\mathcal{X}$. Assuma que uma das três condições abaixo é válida:

- (i) $\mathcal{X}$ é um conjunto compacto.
- (ii) $\mathcal{X}$ é um conjunto fechado e f é coerciva.
- (iii) Existe um escalar γ tal que o conjunto de nível

$$\{\mathbf{x} \in \mathcal{X} : f(\mathbf{x}) \leq \gamma\}$$

é não-vazio e compacto.

Então existe um vetor $\mathbf{x} \in \mathcal{X}$ tal que $f(\mathbf{x}) = \inf_{\mathbf{z} \in \mathcal{X}} f(\mathbf{z})$.

B.1.2 Convexidade e suas implicações

Funções convexas definidas sobre conjuntos convexos possuem importantes propriedades relevantes para a teoria de otimização. Em especial, a garantia de que mínimos locais são mínimos globais para este tipo de problema faz das funções convexas um tipo de função ideal para problemas de otimização.

Definição B.5. O conjunto $K \subseteq \mathbb{R}^n$ é denominado *conjunto convexo* se

$$\alpha \mathbf{x} + (1 - \alpha) \mathbf{y} \in K, \quad \forall \mathbf{x}, \mathbf{y} \in K, \forall \alpha \in [0, 1].$$

Definição B.6. Seja $K \subseteq \mathbb{R}^n$ um conjunto convexo. Dizemos que a função $f : K \rightarrow \mathbb{R}$ é uma *função convexa* se

$$f(\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha) f(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in K, \forall \alpha \in [0, 1].$$

A função f é dita *estritamente convexa* se a desigualdade acima é estrita para todo $\mathbf{x}, \mathbf{y} \in K$ com $\mathbf{x} \neq \mathbf{y}$ e todo $\alpha \in (0, 1)$. Dizemos que f é uma *função côncava* se $-f$ for

¹As normas em $\mathbb{R}^n$ são equivalentes, de modo que a norma utilizada não altera a definição.

uma função convexa, e que f é *estritamente côncava* se $-f$ for uma função estritamente convexa.

Sob a hipótese de diferenciabilidade de uma função sobre um conjunto convexo, a convexidade da mesma também pode ser caracterizada de maneira alternativa, através do teorema a seguir.

Teorema B.2. Sejam $K \subseteq \mathbb{R}^n$ um conjunto convexo e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável sobre K . A função f é convexa sobre K se, e somente se,

$$f(\mathbf{y}) \geq f(\mathbf{x}) + (\mathbf{y} - \mathbf{x})^t \nabla f(\mathbf{x}), \quad \forall \mathbf{x}, \mathbf{y} \in K.$$

O teorema anterior pode ser interpretado geometricamente da seguinte maneira: o hiperplano tangente a uma função convexa f sobre um conjunto convexo K , em um ponto $\mathbf{x} \in K$, assume valores inferiores ao da função sobre o conjunto K . Para funções de classe C^2 , a caracterização de convexidade pode ser feita através de matrizes hessianas, conforme o próximo teorema.

Teorema B.3. Seja $K \subseteq \mathbb{R}^n$ aberto e convexo, $f : K \rightarrow \mathbb{R}$ e $f \in C^2(K)$. Então f é convexa se, e somente se, $\nabla^2 f(\mathbf{x}) \geq 0$ para todo $\mathbf{x} \in K$.

As matrizes hessianas generalizam, de certo modo, o conceito de curvatura de funções. O teorema B.3 afirma que uma função é convexa sobre um conjunto aberto convexo se, e somente se, sua matriz hessiana é semidefinida positiva sobre o referido conjunto. Isto significa que, do ponto de vista geométrico, uma função convexa possui curvatura não-negativa. O próximo teorema é um forte resultado acerca de mínimos de funções convexas.

Teorema B.4. Se $K \subseteq \mathbb{R}^n$ é um conjunto convexo e $f : K \rightarrow \mathbb{R}$ é uma função convexa, então um mínimo local de f também é um mínimo global. Se, em adição, K é aberto e f é diferenciável sobre K , então qualquer ponto estacionário² de f é um mínimo global.

B.1.3 Condições de otimalidade

Segundo Martínez & Santos (1998), as condições de otimalidade são relações entre as derivadas da função objetivo e as derivadas das funções que definem as restrições. São elas que dizem se as derivadas envolvidas permitirão a redução do valor da função objetivo.

²Um ponto $\mathbf{x} \in \mathcal{X}$ é dito *ponto estacionário* de uma função diferenciável $f : \mathcal{X} \rightarrow \mathbb{R}$ se $\nabla f(\mathbf{x}) = \mathbf{0}$.

Nesta seção, consideraremos o problema da forma (B.1). De acordo com os autores, as curvas definidas no conjunto $\mathcal{X}$ são fundamentais na obtenção das condições práticas de otimalidade. Estas condições de otimalidade a partir de curvas serão apresentadas a seguir, com base no trabalho de Martínez & Santos (1998), e permitirão a obtenção de resultados bastante conhecidos em otimização. Para tanto, assume-se que a função f possui derivadas primeiras contínuas em um conjunto aberto que contém o conjunto $\mathcal{X}$.

Definição B.7. Seja $\bar{\mathbf{x}} \in \mathcal{X}$. Chamamos de *curva ou caminho em $\mathcal{X}$ de classe C^k partindo de $\bar{\mathbf{x}}$* a uma função $\gamma : [0, \varepsilon] \rightarrow \mathcal{X}$ tal que $\varepsilon > 0$, $\gamma(0) = \bar{\mathbf{x}}$ e $\gamma \in C^k([0, \varepsilon])$.

Observação: quando as derivadas nos extremos do caminho forem mencionadas estaremos, naturalmente, considerando as derivadas laterais.

O teorema a seguir é a versão baseada em curvas das condições necessárias de primeira ordem.

Teorema B.5. Sejam $\mathbf{x}^*$ mínimo local para o problema (B.1) e γ uma curva em $\mathcal{X}$ de classe C^1 partindo de $\mathbf{x}^*$. Então $\nabla f(\mathbf{x}^*)^t \gamma'(0) \geq 0$.

Demonstração. Defina $\phi : [0, \varepsilon] \rightarrow \mathbb{R}$ tal que $\phi(t) = f(\gamma(t))$. Como $\mathbf{x}^*$ é mínimo local de f , existe $\bar{\varepsilon} \in (0, \varepsilon)$ tal que $\phi(t) \geq \phi(0)$ para todo $t \in (0, \bar{\varepsilon})$. Dessa forma, $\frac{\phi(t) - \phi(0)}{t} \geq 0$, para todo $t \in (0, \bar{\varepsilon})$, de modo que a derivada quando $t \downarrow 0$ é tal que $\phi'(0) \geq 0$. Aplicando a regra da cadeia,

$$\phi'(t) = \nabla f(\gamma(t))^t \gamma'(t)$$

e, portanto, $\phi'(0) = \nabla f(\gamma(0))^t \gamma'(0) = \nabla f(\mathbf{x}^*)^t \gamma'(0) \geq 0$. □

Geometricamente, o teorema B.5 indica que, em um mínimo local, o vetor tangente à uma curva partindo do ponto de mínimo, na extremidade do mínimo, deve formar um ângulo menor ou igual a 90° com o gradiente da função calculado no ponto de mínimo.

O próximo teorema é a versão baseada em curvas das condições necessárias de segunda ordem.

Teorema B.6. Sejam $\mathbf{x}^*$ um mínimo local para o problema (B.1) e $f \in C^2(\mathcal{X})$.

- (a) Para toda curva γ em $\mathcal{X}$ de classe C^2 partindo de $\mathbf{x}^*$, $\phi'(0) = \nabla f(\mathbf{x}^*)^t \gamma'(0) \geq 0$, onde $\phi(t) = f(\gamma(t))$.
- (b) Se $\phi'(0) = 0$, então $\phi''(0) \geq 0$.

Demonstração. A prova do item (a) se encontra no teorema B.5. Já para o item (b), assumindo que $\phi'(0) = 0$ e aplicando a Fórmula de Taylor com resto infinitesimal, em

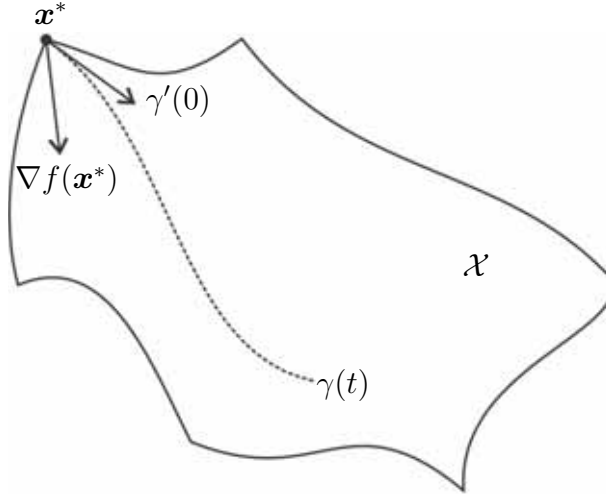


Figura B.1: Representação geométrica do teorema B.5.

torno do valor $t = 0$, tem-se que $\phi(t) = \phi(0) + \frac{1}{2}\phi''(0)t^2 + r(0, t^2)$ em que $\lim_{t \rightarrow 0} \frac{r(0, t^2)}{t^2} = 0$. Logo,

$$\lim_{t \rightarrow 0} \frac{\phi(t) - \phi(0)}{t^2} = \frac{1}{2}\phi''(0).$$

Como $\mathbf{x}^*$ é mínimo local, então $\phi(t) \geq \phi(0)$ para t suficientemente pequeno. Assim, $\phi''(0) \geq 0$. $\square$

Assim como se pode definir uma curva partindo de um ponto, pode-se definir o que é uma curva passando por um ponto.

Definição B.8. Seja $\bar{\mathbf{x}} \in \mathcal{X}$. Chamamos de *curva em $\mathcal{X}$ de classe C^k passando por $\bar{\mathbf{x}}$* a uma função $\gamma : [-\varepsilon, \varepsilon] \rightarrow \mathcal{X}$ tal que $\varepsilon > 0$, $\gamma(0) = \bar{\mathbf{x}}$ e $\gamma \in C^k([-\varepsilon, \varepsilon])$.

Lema B.1. Seja $\mathbf{x}^*$ um mínimo local para o problema (B.1) e γ uma curva em $\mathcal{X}$ de classe C^1 passando por $\mathbf{x}^*$. Então $\nabla f(\mathbf{x}^*)^t \gamma'(0) = 0$.

Demonstração. A partir da curva γ , defina $\gamma_1 : [0, \varepsilon] \rightarrow \mathcal{X}$ com $\gamma_1(t) = \gamma(t)$ e $\gamma_2 : [0, \varepsilon] \rightarrow \mathcal{X}$ com $\gamma_2(t) = \gamma(-t)$. Pelo teorema B.5,

$$\nabla f(\mathbf{x}^*)^t \gamma_1'(0) \geq 0 \quad \text{e} \quad \nabla f(\mathbf{x}^*)^t \gamma_2'(0) \geq 0.$$

Como $\gamma_1'(0) = \gamma'(0)$ e $\gamma_2'(0) = -\gamma'(0)$, segue que $\nabla f(\mathbf{x}^*)^t \gamma'(0) = 0$. $\square$

Corolário B.1. Seja $\mathbf{x}^*$ um ponto interior de $\mathcal{X}$ e mínimo local para o problema (B.1). Então $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

Demonstração. Como $\mathbf{x}^*$ é um ponto interior de $\mathcal{X}$, existe uma bola aberta de raio $\varepsilon > 0$ tal que $B(\mathbf{x}^*, \varepsilon) \subseteq \mathcal{X}$. Seja $\mathbf{d} \in \mathbb{R}^n$, com $\mathbf{d} \neq \mathbf{0}$. Defina a curva $\gamma : [-\varepsilon, \varepsilon] \rightarrow \mathcal{X}$ por

$\gamma(t) = \mathbf{x}^* + t \frac{\mathbf{d}}{2\|\mathbf{d}\|}$. Claramente a curva γ está contida na bola $B(\mathbf{x}^*, \varepsilon)$ e, portanto, está bem definida. Pelo lema B.1,

$$\nabla f(\mathbf{x}^*)^t \frac{\mathbf{d}}{2\|\mathbf{d}\|} = \nabla f(\mathbf{x}^*)^t \gamma'(0) = 0 \Leftrightarrow \nabla f(\mathbf{x}^*)^t \mathbf{d} = 0.$$

Como $\mathbf{d}$ é arbitrário, então $\nabla f(\mathbf{x}^*) = \mathbf{0}$. □

Corolário B.2. Seja $\mathbf{x}^*$ um ponto interior de $\mathcal{X}$ e mínimo local para o problema (B.1). Se f tem derivadas parciais de segunda ordem contínuas numa vizinhança de $\mathbf{x}^*$, então $\nabla f(\mathbf{x}^*) = \mathbf{0}$ e $\nabla^2 f(\mathbf{x}^*) \geq 0$.

Demonstração. De fato, $\nabla f(\mathbf{x}^*) = \mathbf{0}$ pelo corolário B.1. Utilizando a mesma curva γ do corolário B.1, defina a função $\phi : [-\varepsilon, \varepsilon] \rightarrow \mathbb{R}$ tal que $\phi(t) = f(\gamma(t))$. Então $\phi'(0) = \nabla f(\mathbf{x}^*)^t \gamma'(0) = 0$ e, pelo teorema B.6,

$$0 \leq \phi''(0) = \gamma'(0)^t \nabla^2 f(\mathbf{x}^*) \gamma'(0) = \frac{\mathbf{d}^t}{2\|\mathbf{d}\|} \nabla^2 f(\mathbf{x}^*) \frac{\mathbf{d}}{2\|\mathbf{d}\|}.$$

Dessa forma, $\mathbf{d}^t \nabla^2 f(\mathbf{x}^*) \mathbf{d} \geq 0$, para qualquer $\mathbf{d}$. Assim, $\nabla^2 f(\mathbf{x}^*) \geq 0$. □

Os resultados expressos pelos corolários B.1 e B.2 são usualmente denominados *condições necessárias de otimalidade de primeira e segunda ordem*, respectivamente, uma vez que estabelecem relações com as derivadas parciais de primeira ordem que compõem o vetor gradiente e com as derivadas parciais de segunda ordem, presentes na matriz hessiana. É importante ressaltar que o fato de um ponto satisfazer as condições necessárias não garante que o mesmo seja um mínimo local.

Exemplo B.1. Seja a função $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(x) = -x^2$. Claramente o ponto $x = 0$ satisfaz a tese do corolário B.1, pois $f'(0) = 0$, mas este ponto não é um mínimo local. Na realidade, é um máximo local/global.

Exemplo B.2. Seja a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = x^2 - y^4$. O ponto $(x, y) = (0, 0)$ satisfaz a tese do corolário B.2, pois $\nabla f(0, 0) = (0, 0)^t$, e $\nabla^2 f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ é semidefinida positiva. Entretanto, este ponto não é um ponto de mínimo.

O próximo teorema apresenta condições suficientes de segunda ordem, as quais garantem que um ponto que as satisfaz é mínimo local.

Teorema B.7. Sejam $f \in C^2(\mathcal{X})$ e $\mathbf{x}^*$ um ponto interior de $\mathcal{X} \subseteq \mathbb{R}^n$ tal que $\nabla f(\mathbf{x}^*) = \mathbf{0}$ e $\nabla^2 f(\mathbf{x}^*) > 0$. Então $\mathbf{x}^*$ é um mínimo local estrito para o problema (B.1).

Demonstração. Defina a função $\xi : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por $\xi(\mathbf{d}) = \mathbf{d}^t \nabla^2 f(\mathbf{x}^*) \mathbf{d}$. Como

$\nabla^2 f(\mathbf{x}^*) > 0$, então $\xi(\mathbf{d}) > 0$, para todo $\mathbf{d} \in \mathbb{R}^n$, com $\mathbf{d} \neq \mathbf{0}$. Claramente a função ξ é contínua. Seja

$$S^{n-1} = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}$$

a esfera unitária em $\mathbb{R}^n$. A esfera unitária S^{n-1} é um conjunto compacto, de modo que, pelo Teorema de Weierstrass (B.1), existe $\mathbf{v} \in S^{n-1}$ tal que $\xi(\mathbf{v}) = \inf_{\mathbf{d} \in S^{n-1}} \xi(\mathbf{d})$. Como $\mathbf{v} \in S^{n-1}$, então $\mathbf{v} \neq \mathbf{0}$ e, portanto, $\xi(\mathbf{v}) = a > 0$. Como $\nabla f(\mathbf{x}^*) = \mathbf{0}$, aplicando a Fórmula de Taylor com resto infinitesimal, tem-se que $f(\mathbf{x}^* + \mathbf{d}) = f(\mathbf{x}^*) + \frac{1}{2} \mathbf{d}^t \nabla^2 f(\mathbf{x}^*) \mathbf{d} + r(\mathbf{x}^*, \mathbf{d})$, com $\lim_{\mathbf{d} \rightarrow \mathbf{0}} \frac{r(\mathbf{x}^*, \mathbf{d})}{\|\mathbf{d}\|^2} = 0$.

Pela definição de limite, se $\lim_{\mathbf{d} \rightarrow \mathbf{0}} \frac{r(\mathbf{x}^*, \mathbf{d})}{\|\mathbf{d}\|^2} = 0$, existe $\delta > 0$ tal que, se $0 < \|\mathbf{d}\| < \delta$, então $\left| \frac{r(\mathbf{x}^*, \mathbf{d})}{\|\mathbf{d}\|^2} - 0 \right| < \frac{a}{4} \Leftrightarrow -\frac{a}{4} < \frac{r(\mathbf{x}^*, \mathbf{d})}{\|\mathbf{d}\|^2} < \frac{a}{4}$. Assim, para $0 < \|\mathbf{d}\| < \delta$, tem-se que

$$\begin{aligned} f(\mathbf{x}^* + \mathbf{d}) &= f(\mathbf{x}^*) + \frac{1}{2} \mathbf{d}^t \nabla^2 f(\mathbf{x}^*) \mathbf{d} + r(\mathbf{x}^*, \mathbf{d}) \\ &= f(\mathbf{x}^*) + \frac{\|\mathbf{d}\|^2}{2} \left[\left(\frac{\mathbf{d}}{\|\mathbf{d}\|} \right)^t \nabla^2 f(\mathbf{x}^*) \left(\frac{\mathbf{d}}{\|\mathbf{d}\|} \right) + \frac{2r(\mathbf{x}^*, \mathbf{d})}{\|\mathbf{d}\|^2} \right] \\ &> f(\mathbf{x}^*) + \frac{\|\mathbf{d}\|^2}{2} \left[a + 2 \left(-\frac{a}{4} \right) \right] = f(\mathbf{x}^*) + \frac{\|\mathbf{d}\|^2}{2} \frac{a}{2} \\ &> f(\mathbf{x}^*) + \frac{\|\mathbf{d}\|^2}{2} \cdot 0 = f(\mathbf{x}^*) \end{aligned}$$

Assim, $f(\mathbf{x}^* + \mathbf{d}) > f(\mathbf{x}^*)$, para todo $\mathbf{d}$ tal que $0 < \|\mathbf{d}\| < \delta$, ou seja, $\mathbf{x}^*$ é mínimo local estrito. $\square$

Portanto, pode-se garantir que um ponto estacionário de um problema irrestrito é mínimo local quando a matriz hessiana no referido ponto é definida positiva. Na próxima seção serão discutidos alguns aspectos de otimização restrita, quando a região factível é determinada por um conjunto de equações e inequações algébricas.

B.2 Otimização restrita

A fim de discutir os resultados teóricos fundamentais de otimização restrita, começaremos com o problema de minimização de funções com restrições de igualdade.

B.2.1 Restrições de igualdade

Seja um problema de otimização da forma

$$\begin{aligned} & \text{minimizar} && f(\mathbf{x}) \\ & \text{sujeito a} && \mathbf{g}(\mathbf{x}) = \mathbf{0} \end{aligned} \tag{B.2}$$

em que $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $\mathbf{x} \in \Omega \subseteq \mathbb{R}^n$, com Ω um conjunto aberto. Neste caso, o conjunto factível do problema é dado por $\mathcal{X} = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$.

No contexto da otimização restrita, os conceitos de conjunto tangente e de regularidade são de grande importância e possibilitam a demonstração do teorema relacionado aos multiplicadores de Lagrange.

Definição B.9. Seja $\mathbf{x} \in \mathcal{X}$. Chamamos de *conjunto tangente a $\mathcal{X}$ em $\mathbf{x}$* , que será denotado por $M(\mathbf{x})$, ao conjunto dos vetores tangentes às curvas em $\mathcal{X}$ passando por $\mathbf{x}$. Assim,

$$M(\mathbf{x}) = \{\mathbf{v} \in \mathbb{R}^n : \text{existe uma curva } \gamma \text{ em } \mathcal{X} \text{ passando por } \mathbf{x} \text{ com } \mathbf{v} = \gamma'(0)\}. \tag{B.3}$$

A matriz jacobiana da função vetorial $\mathbf{g}$ no ponto $\mathbf{x}$ será denotada por

$$J\mathbf{g}(\mathbf{x}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial g_1}{\partial x_n}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial g_m}{\partial x_n}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \nabla g_1(\mathbf{x})^t \\ \vdots \\ \nabla g_m(\mathbf{x})^t \end{pmatrix}. \tag{B.4}$$

Definição B.10. Seja $A \in \mathbb{R}^{m \times n}$ uma matriz real. O *núcleo* da matriz A é o conjunto

$$\mathcal{N}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}.$$

A *imagem* da matriz A é o conjunto

$$\mathcal{R}(A) = \{\mathbf{b} \in \mathbb{R}^m : A\mathbf{x} = \mathbf{b}, \text{ para algum } \mathbf{x} \in \mathbb{R}^n\}.$$

Definição B.11. Seja S um subconjunto não vazio de um espaço vetorial $\mathbb{V}$ munido de produto interno. O *complemento ortogonal* de S , denotado por $S^\perp$, é o conjunto dos vetores de $\mathbb{V}$ que são ortogonais a todos os vetores de S . Assim,

$$S^\perp = \{\mathbf{v} \in \mathbb{V} : \mathbf{v} \perp \mathbf{w} \text{ para todo } \mathbf{w} \in S\}.$$

Pode-se verificar sem muita dificuldade que $S^\perp$ é um subespaço vetorial de $\mathbb{V}$. Podemos

relacionar o núcleo e a imagem de uma matriz através do teorema a seguir, que pode ser encontrado em diversos livros de Álgebra Linear.

Teorema B.8. Seja $A \in \mathbb{R}^{m \times n}$ uma matriz real. Então $\mathcal{N}(A)^\perp = \mathcal{R}(A^t)$.

O próximo lema estabelece uma relação de inclusão entre o conjunto tangente $M(\mathbf{x})$ e o núcleo do jacobiano, $\mathcal{N}(J\mathbf{g}(\mathbf{x}))$.

Lema B.2. Para todo $\mathbf{x} \in \mathcal{X}$, $M(\mathbf{x}) \subseteq \mathcal{N}(J\mathbf{g}(\mathbf{x}))$.

Demonstração. Seja $\mathbf{v} \in M(\mathbf{x})$. Então existe uma curva $\gamma : [-\varepsilon, \varepsilon] \rightarrow \mathcal{X}$, com $\gamma(0) = \mathbf{x}$ e $\gamma'(0) = \mathbf{v}$. Defina $\Gamma : [-\varepsilon, \varepsilon] \rightarrow \mathbb{R}^m$ por $\Gamma(t) = \mathbf{g}(\gamma(t))$. Como γ é uma curva em $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$, então $\Gamma(t) = \mathbf{0}$ para todo $t \in [-\varepsilon, \varepsilon]$. Assim $\Gamma'(t) = \mathbf{0}$, para todo $t \in (-\varepsilon, \varepsilon)$. Pela regra da cadeia, temos que $\Gamma'(t) = J\mathbf{g}(\gamma(t))\gamma'(t)$. Logo,

$$\Gamma'(t) = J\mathbf{g}(\gamma(t))\gamma'(t) = \mathbf{0}, \quad \forall t \in (-\varepsilon, \varepsilon).$$

Em particular, $\mathbf{0} = J\mathbf{g}(\gamma(0))\gamma'(0) = J\mathbf{g}(\mathbf{x})\mathbf{v}$, isto é, $\mathbf{v} \in \mathcal{N}(J\mathbf{g}(\mathbf{x}))$. Portanto, $M(\mathbf{x}) \subseteq \mathcal{N}(J\mathbf{g}(\mathbf{x}))$. $\square$

É lícito o questionamento natural se a recíproca do lema B.2 é válida. Como mostra o exemplo a seguir, em geral, a volta do lema não é válida.

Exemplo B.3. Considere um problema de otimização em que a restrição de igualdade é dada por $g(x_1, x_2) = x_1^2 - x_2(x_2 - 1)^2$. Considere o ponto $\mathbf{x} = (0, 1)^t$. Então $Jg(0, 1) = \nabla g(0, 1)^t = (0, 0)$ e, portanto, $\mathcal{N}(Jg(0, 1)) = \mathbb{R}^2$. Por outro lado, pode-se verificar facilmente que o conjunto factível determinado por g pode ser parametrizado pela curva $\gamma(t) = (t^3 - t, t^2)^t$, cujos vetores tangentes em $\mathbf{x} = (0, 1)^t$ são $\gamma'(-1) = (2, -2)^t$ e $\gamma'(1) = (2, 2)^t$. Assim, conclui-se que $M(\mathbf{x}) = \{k(1, 1)^t : k \in \mathbb{R}\} \cup \{k(1, -1)^t : k \in \mathbb{R}\}$, ou seja, $\mathcal{N}(Jg(0, 1)) \not\subseteq M(\mathbf{x})$. A Figura B.2 apresenta a região factível do problema, bem como os vetores tangentes em $(0, 1)^t$.

A fim de estabelecer quando $\mathcal{N}(J\mathbf{g}(\mathbf{x})) \subseteq M(\mathbf{x})$, introduz-se a definição de ponto regular.

Definição B.12. Seja $\mathbf{x} \in \mathcal{X} = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$. Dizemos que $\mathbf{x}$ é um *ponto regular* se o posto de $J\mathbf{g}(\mathbf{x})$ é igual a m . De modo equivalente, isto quer dizer que o conjunto $\{\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})\}$ é linearmente independente. É comum dizer que, neste caso, o Requisito de Independência Linear das Restrições é satisfeito no ponto $\mathbf{x}$.

O teorema a seguir, cuja demonstração foge do escopo deste trabalho, será utilizado para

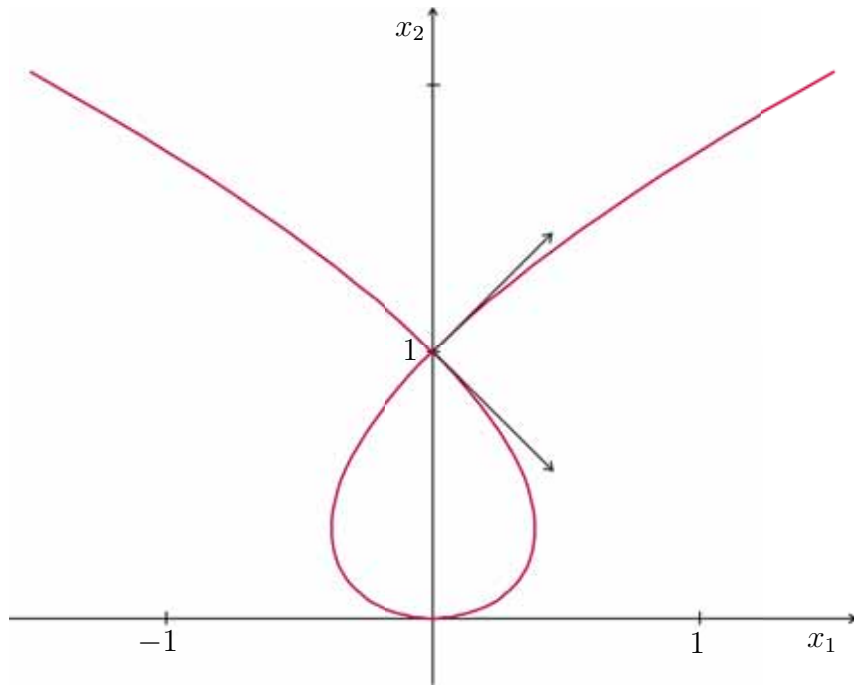


Figura B.2: Curva da região factível e vetores tangentes em $(0, 1)^t$, no exemplo B.3.

demonstrar que sob a hipótese de regularidade em um ponto, o conjunto tangente e o núcleo do jacobiano no referido ponto são iguais.

Teorema B.9. (Teorema da Função Implícita) Seja $\mathbf{f} = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ uma função vetorial de classe C^k ($k \geq 1$) no aberto $U \subseteq \mathbb{R}^{m+n}$. Suponha que no ponto $\mathbf{p} = (\mathbf{x}_0, \mathbf{y}_0) \in U$ (com $\mathbf{x}_0 \in \mathbb{R}^m$ e $\mathbf{y}_0 \in \mathbb{R}^n$), $\mathbf{f}(\mathbf{p}) = \mathbf{c}$ e a matriz jacobiana $n \times n$, em relação as variáveis $\mathbf{y}$,

$$J_{\mathbf{y}}\mathbf{f}(\mathbf{p}) = \left(\frac{\partial f_i}{\partial y_j}(\mathbf{p}) \right) \quad (i, j = 1, \dots, n),$$

seja não-singular. Então existem $Z \subseteq U$, aberto contendo $\mathbf{p}$, $V \subseteq \mathbb{R}^m$, aberto contendo $\mathbf{x}_0$, e $\xi : V \rightarrow \mathbb{R}^n$ de classe C^k , com $\xi(\mathbf{x}_0) = \mathbf{y}_0$, com a seguinte propriedade:

$$(\mathbf{x}, \mathbf{y}) \in Z \text{ e } \mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{c} \Leftrightarrow \mathbf{x} \in V \text{ e } \mathbf{y} = \xi(\mathbf{x}).$$

Teorema B.10. Sejam $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$, $\mathbf{g} \in C^k$ ($k \geq 1$) e $\mathbf{x} \in \mathcal{X}$ um ponto regular. Então para todo $\mathbf{v} \in \mathcal{N}(J\mathbf{g}(\mathbf{x}))$, existe uma curva γ de classe C^k passando por $\mathbf{x}$, tal que $\gamma'(0) = \mathbf{v}$. Isto é, $M(\mathbf{x}) = \mathcal{N}(J\mathbf{g}(\mathbf{x}))$.

Demonstração. Tomemos $\mathbf{v} \in \mathcal{N}(J\mathbf{g}(\mathbf{x}))$. Isto significa que $J\mathbf{g}(\mathbf{x})\mathbf{v} = \mathbf{0}$. Devemos determinar uma curva γ em $\mathcal{X}$, passando por $\mathbf{x}$, com $\gamma'(0) = \mathbf{v}$. Fixe $\mathbf{x}$ e $\mathbf{v}$, e considere

o seguinte sistema de equações nas variáveis $t \in \mathbb{R}$ e $\mathbf{u} \in \mathbb{R}^m$:

$$\mathbf{F}(t, \mathbf{u}) = \mathbf{g}(\mathbf{x} + t\mathbf{v} + J\mathbf{g}(\mathbf{x})^t \mathbf{u}) = \mathbf{0} \quad (\text{B.5})$$

O sistema apresentado possui m equações e $m + 1$ variáveis. Claramente, este sistema tem uma solução particular para $t = 0$ e $\mathbf{u} = \mathbf{0}$. A matriz jacobiana de $\mathbf{F}$ em relação às variáveis de $\mathbf{u}$ é uma matriz $m \times m$ dada por $J_{\mathbf{u}}\mathbf{F}(t, \mathbf{u}) = J\mathbf{g}(\mathbf{x} + t\mathbf{v} + J\mathbf{g}(\mathbf{x})^t \mathbf{u})J\mathbf{g}(\mathbf{x})^t$. Assim, para $t = 0$ e $\mathbf{u} = \mathbf{0}$, $J_{\mathbf{u}}\mathbf{F}(0, \mathbf{0}) = J\mathbf{g}(\mathbf{x})J\mathbf{g}(\mathbf{x})^t$, que é uma matriz não-singular devido à regularidade de $\mathbf{x}$. Pelo Teorema da Função Implícita (B.9), existe uma função $\bar{\gamma} : [-\varepsilon, \varepsilon] \rightarrow \mathbb{R}^m$ de classe C^k (uma curva), com $\bar{\gamma}(0) = \mathbf{0}$, tal que a igualdade (B.5) é satisfeita se, e somente se, $\mathbf{u} = \bar{\gamma}(t)$, para todo $t \in [-\varepsilon, \varepsilon]$ (a rigor, para um intervalo aberto em torno de 0, o qual contém um intervalo fechado da forma $[-\varepsilon, \varepsilon]$). Assim,

$$\mathbf{g}(\mathbf{x} + t\mathbf{v} + J\mathbf{g}(\mathbf{x})^t \bar{\gamma}(t)) = \mathbf{0} \text{ para todo } t \in [-\varepsilon, \varepsilon]. \quad (\text{B.6})$$

Derivando (B.6) em relação a t , e calculando em $t = 0$, obtém-se $J\mathbf{g}(\mathbf{x})(\mathbf{v} + J\mathbf{g}(\mathbf{x})^t \bar{\gamma}'(0)) = \mathbf{0}$. Como $J\mathbf{g}(\mathbf{x})\mathbf{v} = \mathbf{0}$, então $J\mathbf{g}(\mathbf{x})J\mathbf{g}(\mathbf{x})^t \bar{\gamma}'(0) = \mathbf{0}$. Mas $J\mathbf{g}(\mathbf{x})J\mathbf{g}(\mathbf{x})^t$ é não-singular, do que pode-se concluir que $\bar{\gamma}'(0) = \mathbf{0}$.

Defina $\gamma : [-\varepsilon, \varepsilon] \rightarrow \mathcal{X}$ por $\gamma(t) = \mathbf{x} + t\mathbf{v} + J\mathbf{g}(\mathbf{x})^t \bar{\gamma}(t)$. Então,

$$\gamma'(0) = \mathbf{v} + J\mathbf{g}(\mathbf{x})^t \bar{\gamma}'(0) = \mathbf{v} + J\mathbf{g}(\mathbf{x})^t \mathbf{0} = \mathbf{v}.$$

Assim, γ é a curva procurada. Como $\mathbf{v}$ é arbitrário, segue que $\mathcal{N}(J\mathbf{g}(\mathbf{x})) \subseteq M(\mathbf{x})$. Combinando este resultado com o do lema B.2, conclui-se que $\mathcal{N}(J\mathbf{g}(\mathbf{x})) = M(\mathbf{x})$. $\square$

Corolário B.3. Se $\mathbf{x}^*$ é um ponto de mínimo local regular para o problema (B.2), então $\nabla f(\mathbf{x}^*) \perp \mathcal{N}(J\mathbf{g}(\mathbf{x}^*))$.

Demonstração. Seja $\mathbf{v} \in \mathcal{N}(J\mathbf{g}(\mathbf{x}^*))$. Como $\mathbf{x}^*$ é regular, então $\mathcal{N}(J\mathbf{g}(\mathbf{x}^*)) = M(\mathbf{x}^*)$, de modo que $\mathbf{v} \in M(\mathbf{x}^*)$. Assim, existe uma curva γ em $\mathcal{X}$ passando por $\mathbf{x}^*$ tal que $\gamma'(0) = \mathbf{v}$. Pelo lema B.1, $\nabla f(\mathbf{x}^*)^t \gamma'(0) = \nabla f(\mathbf{x}^*)^t \mathbf{v} = 0$. Logo, $\nabla f(\mathbf{x}^*) \perp \mathcal{N}(J\mathbf{g}(\mathbf{x}^*))$. $\square$

B.2.1.1 Condições de otimalidade para restrições de igualdade

O seguinte teorema é de valor inestimável para a teoria de otimização com restrições de igualdade, e possui uma versão estendida, para restrições de desigualdade (as condições de Karush-Kuhn-Tucker, que serão apresentadas posteriormente).

Teorema B.11. (Teorema dos Multiplicadores de Lagrange) Seja $\mathbf{x}^*$ um mínimo local regular para o problema (B.2). Então existem únicos $\lambda_1^*, \dots, \lambda_m^*$ números reais, de-

nominados *multiplicadores de Lagrange*, tais que $\nabla f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{x}^*) = \mathbf{0}$.

Demonstração. Pelo corolário B.3, $\nabla f(\mathbf{x}^*) \perp \mathcal{N}(J\mathbf{g}(\mathbf{x}^*))$. Assim, $\nabla f(\mathbf{x}^*) \in \mathcal{N}(J\mathbf{g}(\mathbf{x}^*))^\perp$. Como, pelo teorema B.8, $\mathcal{N}(J\mathbf{g}(\mathbf{x}^*))^\perp = \mathcal{R}(J\mathbf{g}(\mathbf{x}^*)^t)$, então conclui-se que $\nabla f(\mathbf{x}^*)$ pode ser gerado pelos vetores coluna de $J\mathbf{g}(\mathbf{x}^*)^t$, que são os gradientes $\nabla g_1(\mathbf{x}^*), \dots, \nabla g_m(\mathbf{x}^*)$. Assim, existe $\boldsymbol{\lambda}^* = (\lambda_1^*, \dots, \lambda_m^*)^t$ tal que $\nabla f(\mathbf{x}^*) + J\mathbf{g}(\mathbf{x}^*)^t \boldsymbol{\lambda}^* = \mathbf{0}$. A unicidade de $\boldsymbol{\lambda}^*$ decorre do fato de que $\mathbf{x}^*$ é ponto regular, de modo que a matriz jacobiana $J\mathbf{g}(\mathbf{x}^*)$ possui posto completo. $\square$

O Teorema dos Multiplicadores de Lagrange (B.11) é o equivalente das condições necessárias de primeira ordem, considerando restrições de igualdade. A seguir, apresenta-se o equivalente das condições necessárias de segunda ordem.

Teorema B.12. Suponha que $f, \mathbf{g} \in C^2$, $\mathbf{x}^*$ é mínimo local regular para o problema B.2, e que $\boldsymbol{\lambda}^*$ é o vetor de multiplicadores de Lagrange, cuja existência é garantida pelo teorema B.11. Então $\mathbf{v}^t \left(\nabla^2 f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^*) \right) \mathbf{v} \geq 0$, para todo $\mathbf{v} \in \mathcal{N}(J\mathbf{g}(\mathbf{x}^*))$.

Demonstração. Sabe-se, do teorema dos multiplicadores de Lagrange, que

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{x}^*) = \mathbf{0}. \quad (\text{B.7})$$

Seja $\mathbf{v} \in \mathcal{N}(J\mathbf{g}(\mathbf{x}^*))$. Aplicando o teorema B.10, existe uma curva γ de classe C^2 em $\mathcal{X}$, passando por $\mathbf{x}^*$, tal que $\mathbf{v} = \gamma'(0) \in \mathcal{N}(J\mathbf{g}(\mathbf{x}^*))$. Defina $\phi(t) = f(\gamma(t))$. Pelo lema B.1, segue que $\phi'(0) = \nabla f(\mathbf{x}^*)^t \gamma'(0) = 0$. Aplicando o teorema B.6, temos que

$$\phi''(0) = \gamma'(0)^t \nabla^2 f(\mathbf{x}^*) \gamma'(0) + \nabla f(\mathbf{x}^*)^t \gamma''(0) \geq 0. \quad (\text{B.8})$$

Define-se as funções $\Phi_i(t) = \lambda_i^* g_i(\gamma(t))$, para $i = 1, \dots, m$. Observe que $\Phi_i(t) = 0$ para todo t no domínio $[-\varepsilon, \varepsilon]$ da curva γ , pois γ é uma curva em $\mathcal{X}$. Assim, $\Phi'_i(t) = 0$, para todo $t \in (-\varepsilon, \varepsilon)$. Logo

$$\Phi''_i(0) = \gamma'(0)^t \lambda_i^* \nabla^2 g_i(\mathbf{x}^*) \gamma'(0) + \lambda_i^* \nabla g_i(\mathbf{x}^*)^t \gamma''(0) = 0.$$

Portanto,

$$\sum_{i=1}^m \Phi''_i(0) = \gamma'(0)^t \left(\sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^*) \right) \gamma'(0) + \left(\sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{x}^*)^t \right) \gamma''(0) = 0. \quad (\text{B.9})$$

Somando (B.8) e (B.9), e considerando (B.7), temos que

$$\mathbf{v}^t \left(\nabla^2 f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^*) \right) \mathbf{v} = \gamma'(0)^t \left(\nabla^2 f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^*) \right) \gamma'(0) \geq 0.$$

Como tomou-se $\mathbf{v}$ arbitrário, segue que $\mathbf{v}^t \left(\nabla^2 f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^*) \right) \mathbf{v} \geq 0$, para todo $\mathbf{v} \in \mathcal{N}(J\mathbf{g}(\mathbf{x}^*))$. $\square$

Assim como no caso irrestrito, também pode-se introduzir condições suficientes de otimalidade para o problema com restrições de igualdade. A fim de demonstrá-lo, utiliza-se o seguinte resultado de Análise Real:

Teorema B.13. (Teorema de Bolzano-Weierstrass) Toda sequência limitada em $\mathbb{R}^n$ possui subsequência convergente.

Teorema B.14. Sejam $f, \mathbf{g} \in C^2$, $\mathbf{x}^* \in \mathcal{X}$ e $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ tais que

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{x}^*) = \mathbf{0}.$$

Suponha também que a matriz $\nabla^2 f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^*)$ é definida positiva em $\mathcal{N}(J\mathbf{g}(\mathbf{x}^*))$. Então $\mathbf{x}^*$ é um mínimo local estrito para o problema (B.2).

Demonstração. Suponha, por absurdo, que $\mathbf{x}^*$ não seja um mínimo local estrito. Então existe uma sequência de pontos factíveis $\{\mathbf{y}_k\}$, com $\mathbf{y}_k \neq \mathbf{x}^*$ para todo k , convergindo para $\mathbf{x}^*$, tal que para cada k , $f(\mathbf{y}_k) \leq f(\mathbf{x}^*)$ (isto é, toda vizinhança de $\mathbf{x}^*$ possui pelo menos algum ponto factível diferente de $\mathbf{x}^*$ com valor igual ou inferior ao de $\mathbf{x}^*$ pela função objetivo). Pode-se escrever cada $\mathbf{y}_k$ como $\mathbf{y}_k = \mathbf{x}^* + \delta_k \mathbf{s}_k$, em que $\mathbf{s}_k \in S^{n-1} = \{\mathbf{v} \in \mathbb{R}^n : \|\mathbf{v}\| = 1\}$ e $\delta_k > 0$ para cada k . Como $\mathbf{y}_k \rightarrow \mathbf{x}^*$ e $\|\mathbf{s}_k\| = 1$, devemos ter $\delta_k \rightarrow 0$. Como $\{\mathbf{s}_k\}$ é uma sequência limitada (sobre a esfera S^{n-1}), pelo teorema de Bolzano-Weierstrass, ela possui uma subsequência $\{\mathbf{s}_k\}_{k \in \mathcal{K}}$ que converge para algum $\mathbf{s}^*$. Além disso, $\mathbf{g}(\mathbf{y}_k) - \mathbf{g}(\mathbf{x}^*) = \mathbf{0}$, de modo que $\frac{\mathbf{g}(\mathbf{y}_k) - \mathbf{g}(\mathbf{x}^*)}{\delta_k} = \mathbf{0}$. Observemos que, devido à diferenciabilidade de $\mathbf{g}$, tem-se que $\mathbf{0} = \lim_{\substack{k \rightarrow \infty \\ k \in \mathcal{K}}} \frac{\mathbf{g}(\mathbf{y}_k) - \mathbf{g}(\mathbf{x}^*)}{\delta_k} = \lim_{\substack{k \rightarrow \infty \\ k \in \mathcal{K}}} \frac{\mathbf{g}(\mathbf{x}^* + \delta_k \mathbf{s}_k) - \mathbf{g}(\mathbf{x}^*)}{\delta_k} = J\mathbf{g}(\mathbf{x}^*)\mathbf{s}^*$. Assim, $J\mathbf{g}(\mathbf{x}^*)\mathbf{s}^* = \mathbf{0}$, isto é, $\mathbf{s}^* \in \mathcal{N}(J\mathbf{g}(\mathbf{x}^*))$. Observe, ainda, que sendo $\|\mathbf{s}_k\| = 1$, para todo k , deve-se ter que $\|\mathbf{s}^*\| = 1$. Pela fórmula de Taylor com resto de Lagrange, temos que, para cada $i = 1, \dots, m$, existe $\theta_i \in (0, 1)$ tal que

$$0 = g_i(\mathbf{y}_k) = g_i(\mathbf{x}^* + \delta_k \mathbf{s}_k) = g_i(\mathbf{x}^*) + \delta_k \nabla g_i(\mathbf{x}^*)^t \mathbf{s}_k + \frac{\delta_k^2}{2} \mathbf{s}_k^t \nabla^2 g_i(\mathbf{x}^* + \theta_i \delta_k \mathbf{s}_k) \mathbf{s}_k \quad (\text{B.10})$$

e, como $f(\mathbf{y}_k) \leq f(\mathbf{x}^*)$, existe $\theta_0 \in (0, 1)$ tal que

$$0 \geq f(\mathbf{y}_k) - f(\mathbf{x}^*) = \delta_k \nabla f(\mathbf{x}^*)^t \mathbf{s}_k + \frac{\delta_k^2}{2} \mathbf{s}_k^t \nabla^2 f(\mathbf{x}^* + \theta_0 \delta_k \mathbf{s}_k) \mathbf{s}_k. \quad (\text{B.11})$$

Multiplicando, para cada i , a equação (B.10) por λ_i^* , somando com (B.11) e utilizando a hipótese do teorema, de que $\nabla f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{x}^*) = \mathbf{0}$ e $\mathbf{g}(\mathbf{x}^*) = 0$, temos que

$$0 \geq \frac{\delta_k^2}{2} \mathbf{s}_k^t \left(\nabla^2 f(\mathbf{x}^* + \theta_0 \delta_k \mathbf{s}_k) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^* + \theta_i \delta_k \mathbf{s}_k) \right) \mathbf{s}_k.$$

Pela continuidade de $\nabla^2 f$ e $\nabla^2 g_i$, aplicando limite para $k \rightarrow \infty$, com $k \in \mathcal{K}$, conclui-se que:

$$0 \geq (\mathbf{s}_k^*)^t \left(\nabla^2 f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^*) \right) \mathbf{s}_k^*,$$

o que é um absurdo pois $\mathbf{s}^* \in \mathcal{N}(J\mathbf{g}(\mathbf{x}^*))$, $\mathbf{s}^* \neq \mathbf{0}$ (já que $\|\mathbf{s}^*\| = 1$) e $\nabla^2 f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^*)$ é definida positiva em $\mathcal{N}(J\mathbf{g}(\mathbf{x}^*))$. $\square$

Cabe destacar que a demonstração do teorema B.14 não requer a regularidade no ponto $\mathbf{x}^*$, diferentemente das condições necessárias. Pode-se notar a similaridade do teorema anterior com o das condições suficientes de segunda ordem para problemas irrestritos: ambos requerem a positividade de alguma matriz relacionada ao problema.

B.2.2 Restrições de desigualdade

Agora trataremos, brevemente, o problema de minimização com restrições de desigualdade, isto é, o problema da forma:

$$\begin{aligned} & \text{minimizar} && f(\mathbf{x}) \\ & \text{sujeito a} && \mathbf{h}(\mathbf{x}) \leq \mathbf{0} \end{aligned} \quad (\text{B.12})$$

em que $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ e $\mathbf{x} \in \Omega \subseteq \mathbb{R}^n$, com Ω um conjunto aberto. Neste caso, conjunto factível é expresso por $\mathcal{X} = \{\mathbf{x} \in \Omega : \mathbf{h}(\mathbf{x}) \leq \mathbf{0}\}$. Segundo Luenberger & Ye (2008), existem diversos estudos sobre condições de qualificação e de regularidade para problemas com restrições de desigualdade. Entretanto, de acordo com o mesmo autor, os resultados que advêm destas teorias são, normalmente, de interesse teórico isoladamente, exercendo pouca influência e contribuição no desenvolvimento de algoritmos.

Nesta seção, aborda-se este problema de forma sucinta. Isto porque a teoria para otimização com restrições de desigualdade necessita de um número bem maior de resultados

para atingir as condições de otimalidade.

Definição B.13. Seja $\mathbf{x} \in \mathcal{X}$. O *cone de direções factíveis em $\mathbf{x}$* é

$$\mathcal{C}(\mathbf{x}) = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{d} \neq \mathbf{0} \text{ e } \mathbf{x} + \alpha\mathbf{d} \in \mathcal{X} \text{ para todo } \alpha \in (0, \delta) \text{ para algum } \delta > 0\}.$$

Nem sempre a noção de direção factível representada pelo conjunto $\mathcal{C}(\mathbf{x})$ faz sentido. Por exemplo, se a região factível for apenas uma circunferência, a partir de um ponto sobre a mesma não existe uma direção factível: qualquer passo suficientemente pequeno ainda levará a um ponto infactível, que não se situa sobre a circunferência. Por isso, utilizar-se-á o seguinte conjunto, cujos elementos são as derivadas laterais sobre as curvas factíveis a partir de $\mathbf{x}$:

$$\mathcal{D}(\mathbf{x}) = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{d} = \gamma'(0), \mathbf{d} \neq \mathbf{0}, \text{ para alguma curva } \gamma \in C^1 \text{ em } \mathcal{X} \text{ a partir de } \mathbf{x}\}.$$

Pode-se verificar, sem dificuldade, que o conjunto $\mathcal{C}(\mathbf{x}) \subseteq \mathcal{D}(\mathbf{x})$ (basta considerar a curva como sendo um segmento de reta). Pode-se mostrar, também, que o conjunto $\mathcal{D}(\mathbf{x})$ apresentado está contido no conjunto de *direções limite de sequências factíveis*, definidas em Nocedal & Wright (2006).

Definição B.14. Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função real. O conjunto

$$\mathcal{F}(\mathbf{x}) = \{\mathbf{d} \in \mathbb{R}^n : f(\mathbf{x} + \alpha\mathbf{d}) < f(\mathbf{x}) \text{ para todo } \alpha \in (0, \delta) \text{ para algum } \delta > 0\}$$

é denominado *cone de direções de descida em $\mathbf{x}$* . Um elemento do conjunto $\mathcal{F}(\mathbf{x})$ é chamado *direção de descida em $\mathbf{x}$* .

Teorema B.15. Considere o problema (B.1). Suponha que f é diferenciável em $\mathbf{x}^* \in \mathcal{X}$ e que $\mathbf{x}^*$ é um mínimo local. Então $\mathcal{F}_0(\mathbf{x}^*) \cap \mathcal{D}(\mathbf{x}^*) = \emptyset$, em que $\mathcal{F}_0(\mathbf{x}^*) = \{\mathbf{d} \in \mathbb{R}^n : \nabla f(\mathbf{x}^*)^t \mathbf{d} < 0\}$.

Demonstração. Suponha, por absurdo, que exista $\mathbf{d} \in \mathcal{F}_0(\mathbf{x}^*) \cap \mathcal{D}(\mathbf{x}^*)$. Como $\mathbf{d} \in \mathcal{D}(\mathbf{x}^*)$, existe uma curva $\gamma : [0, \varepsilon] \rightarrow \mathcal{X}$, tal que $\mathbf{d} = \gamma'(0)$. Também, $\mathbf{d} \in \mathcal{F}_0(\mathbf{x}^*)$, de modo que $\nabla f(\mathbf{x}^*)^t \gamma'(0) = \nabla f(\mathbf{x}^*)^t \mathbf{d} < 0$. Mas isto contradiz o teorema B.5, tendo em vista que $\mathbf{x}^*$ é mínimo local. Logo, $\mathcal{F}_0(\mathbf{x}^*) \cap \mathcal{D}(\mathbf{x}^*) = \emptyset$. $\square$

Lema B.3. Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável. Considere os conjuntos $\mathcal{F}(\mathbf{x})$, $\mathcal{F}_0(\mathbf{x})$ (conforme o teorema B.15) e $\mathcal{F}'_0(\mathbf{x}) = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{d} \neq \mathbf{0} \text{ e } \nabla f(\mathbf{x})^t \mathbf{d} \leq 0\}$. Então $\mathcal{F}_0(\mathbf{x}) \subseteq \mathcal{F}(\mathbf{x}) \subseteq \mathcal{F}'_0(\mathbf{x})$.

Demonstração. Seja $\mathbf{d} \in \mathcal{F}_0(\mathbf{x})$. Então $\nabla f(\mathbf{x})^t \mathbf{d} < 0$. Pela fórmula de Taylor com resto

infinitesimal,

$$f(\mathbf{x} + \alpha \mathbf{d}) = f(\mathbf{x}) + \alpha \nabla f(\mathbf{x})^t \mathbf{d} + r(\mathbf{x}, \alpha \mathbf{d})$$

com $\lim_{\alpha \mathbf{d} \rightarrow \mathbf{0}} \frac{r(\mathbf{x}, \alpha \mathbf{d})}{\|\alpha \mathbf{d}\|} = 0$. Pela definição de limite, existe $\delta_1 > 0$ tal que se $0 < \|\alpha \mathbf{d}\| < \delta_1 \Leftrightarrow 0 < |\alpha| < \frac{\delta_1}{\|\mathbf{d}\|}$, então $\left| \frac{r(\mathbf{x}, \alpha \mathbf{d})}{\|\alpha \mathbf{d}\|} - 0 \right| < -\frac{1}{\|\mathbf{d}\|} \nabla f(\mathbf{x})^t \mathbf{d} \Rightarrow |r(\mathbf{x}, \alpha \mathbf{d})| < -|\alpha| \nabla f(\mathbf{x})^t \mathbf{d}$. Isto é, para $0 < |\alpha| < \frac{\delta_1}{\|\mathbf{d}\|}$, tem-se que $|\alpha| \nabla f(\mathbf{x})^t \mathbf{d} < r(\mathbf{x}, \alpha \mathbf{d}) < -|\alpha| \nabla f(\mathbf{x})^t \mathbf{d}$. Portanto, para $0 < \alpha < \frac{\delta_1}{\|\mathbf{d}\|}$,

$$r(\mathbf{x}, \alpha \mathbf{d}) < -|\alpha| \nabla f(\mathbf{x})^t \mathbf{d} = -\alpha \nabla f(\mathbf{x})^t \mathbf{d} \Leftrightarrow \alpha \nabla f(\mathbf{x})^t \mathbf{d} + r(\mathbf{x}, \alpha \mathbf{d}) < 0,$$

de modo que

$$f(\mathbf{x} + \alpha \mathbf{d}) = f(\mathbf{x}) + \alpha \nabla f(\mathbf{x})^t \mathbf{d} + r(\mathbf{x}, \alpha \mathbf{d}) < f(\mathbf{x}) + 0 = f(\mathbf{x}).$$

Portanto, $\mathbf{d} \in \mathcal{F}(\mathbf{x})$, ou seja, $\mathcal{F}_0(\mathbf{x}) \subseteq \mathcal{F}(\mathbf{x})$. Analogamente, prova-se que $\mathcal{F}(\mathbf{x}) \subseteq \mathcal{F}'_0(\mathbf{x})$. $\square$

Consideremos o problema do formato (B.12), isto é, quando $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}$.

Definição B.15. Uma restrição de desigualdade h_j é dita *restrição ativa em $\mathbf{x}$* se $h_j(\mathbf{x}) = 0$. Caso contrário, isto é, $h_j(\mathbf{x}) < 0$, dizemos que h_j é uma *restrição inativa em $\mathbf{x}$* . O conjunto $\mathcal{A}(\mathbf{x}) = \{j : h_j(\mathbf{x}) = 0\}$ é o conjunto dos índices das restrições ativas em $\mathbf{x}$.

Lema B.4. Considere a região factível $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{h}(\mathbf{x}) \leq \mathbf{0}\}$, com $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^p$. Dado um ponto factível $\mathbf{x} \in \mathcal{X}$, assumamos que h_j é diferenciável em $\mathbf{x}$ se $j \in \mathcal{A}(\mathbf{x})$ e que h_j é contínua em $\mathbf{x}$ se $j \notin \mathcal{A}(\mathbf{x})$. Defina os conjuntos

$$\mathcal{H}_0(\mathbf{x}) = \{\mathbf{d} \in \mathbb{R}^n : \nabla h_j(\mathbf{x})^t \mathbf{d} < 0 \text{ para cada } j \in \mathcal{A}(\mathbf{x})\}$$

$$\mathcal{H}'_0(\mathbf{x}) = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{d} \neq \mathbf{0} \text{ e } \nabla h_j(\mathbf{x})^t \mathbf{d} \leq 0 \text{ para cada } j \in \mathcal{A}(\mathbf{x})\}.$$

Então $\mathcal{H}_0(\mathbf{x}) \subseteq \mathcal{C}(\mathbf{x}) \subseteq \mathcal{D}(\mathbf{x}) \subseteq \mathcal{H}'_0(\mathbf{x})$.

Demonstração. Mostra-se que $\mathcal{H}_0(\mathbf{x}) \subseteq \mathcal{C}(\mathbf{x})$. Seja $\mathbf{d} \in \mathcal{H}_0(\mathbf{x})$. Se $j \notin \mathcal{A}(\mathbf{x})$, então $h_j(\mathbf{x}) < 0$ e, pela continuidade de h_j , tem-se que existe $\delta_1 > 0$ tal que $h_j(\mathbf{x} + \alpha \mathbf{d}) < 0$, para todo $\alpha \in (0, \delta_1)^3$. Por outro lado, se $j \in \mathcal{A}(\mathbf{x})$, então tem-se que $\nabla h_j(\mathbf{x})^t \mathbf{d} < 0$. Utilizando o mesmo procedimento da demonstração no lema B.3, mostra-se que existe $\delta_2 > 0$ tal que $h_j(\mathbf{x} + \alpha \mathbf{d}) < h_j(\mathbf{x}) = 0$ para todo $\alpha \in (0, \delta_2)$. Por fim, como Ω é aberto, existe $\delta_3 > 0$ tal que $\mathbf{x} + \alpha \mathbf{d} \in \Omega$ para todo $\alpha \in (0, \delta_3)$. Tomando $\delta = \min\{\delta_1, \delta_2, \delta_3\}$, segue que $h_j(\mathbf{x} + \alpha \mathbf{d}) < 0$, para todo $j = 1, \dots, p$ e $\alpha \in (0, \delta)$, isto é $\mathbf{x} + \alpha \mathbf{d} \in \mathcal{X}$, para todo $\alpha \in (0, \delta)$.

³Este resultado é conhecido como Teorema da Conservação do Sinal.

Mostra-se, agora, que $\mathcal{C}(\mathbf{x}) \subseteq \mathcal{D}(\mathbf{x})$. Seja $\mathbf{d} \in \mathcal{C}(\mathbf{x})$. Então existe $\delta > 0$ tal que $\mathbf{x} + \alpha \mathbf{d} \in \mathcal{X}$ para todo $\alpha \in (0, \delta)$. Defina a curva $\gamma : [0, \frac{\delta}{2}] \rightarrow \mathcal{X}$ por $\gamma(t) = \mathbf{x} + t\mathbf{d}$. Claramente esta é uma curva factível, com $\gamma(0) = \mathbf{x}$ e $\gamma'(0) = \mathbf{d}$. Assim, $\mathbf{d} \in \mathcal{D}(\mathbf{x})$ ($\mathbf{d} \neq \mathbf{0}$ pela definição de $\mathcal{C}(\mathbf{x})$).

Finalmente, mostra-se que $\mathcal{D}(\mathbf{x}) \subseteq \mathcal{H}'_0(\mathbf{x})$. Seja $\mathbf{d} \in \mathcal{D}(\mathbf{x})$. Então existe uma curva de classe C^1 em $\mathcal{X}$, $\gamma : [0, \varepsilon] \rightarrow \mathcal{X}$, com $\gamma(0) = \mathbf{x}$ e $\gamma'(0) = \mathbf{d}$. Suponha, por absurdo, que $\mathbf{d} \notin \mathcal{H}'_0(\mathbf{x})$. Então, existe $j \in \mathcal{A}(\mathbf{x})$ tal que $\nabla h_j(\mathbf{x})^t \mathbf{d} > 0$. Defina $g_j : [0, \varepsilon] \rightarrow \mathbb{R}$ por $g_j(t) = h_j(\gamma(t))$. Então a derivada lateral em 0 é $g'_j(0) = \nabla h_j(\gamma(0))^t \gamma'(0) = \nabla h_j(\mathbf{x})^t \mathbf{d} > 0$. Logo, existe δ , com $0 < \delta < \varepsilon$, tal que se $t \in (0, \delta)$, então $g_j(t) > g_j(0)$ (a função g_j deve crescer, localmente, à direita de 0, pois $g'_j(0) > 0$). Logo, para $t \in (0, \delta)$, $h_j(\gamma(t)) > h_j(\gamma(0)) = h_j(\mathbf{x}) = 0$, uma vez que $j \in \mathcal{A}(\mathbf{x})$. Por outro lado, a curva γ é factível, de modo que para todo $t \in (0, \delta) \subset [0, \varepsilon]$, deve-se ter $h_j(\gamma(t)) \leq 0$, o que é uma contradição. Portanto, $\mathcal{D}(\mathbf{x}) \subseteq \mathcal{H}'_0(\mathbf{x})$. $\square$

Teorema B.16. Considere o problema (B.12). Seja $\mathbf{x}^*$ um ponto factível tal que f e h_j são diferenciáveis em $\mathbf{x}^*$, para $j \in \mathcal{A}(\mathbf{x}^*)$, e h_j é contínua em $\mathbf{x}^*$, para $j \notin \mathcal{A}(\mathbf{x}^*)$. Se $\mathbf{x}^*$ é um mínimo local, então $\mathcal{F}_0(\mathbf{x}^*) \cap \mathcal{H}_0(\mathbf{x}^*) = \emptyset$.

Demonstração. Como $\mathbf{x}^*$ é mínimo local, pelo teorema B.15, $\mathcal{F}_0(\mathbf{x}^*) \cap \mathcal{D}(\mathbf{x}^*) = \emptyset$. Mas, pelo lema B.4, $\mathcal{H}_0(\mathbf{x}^*) \subseteq \mathcal{D}(\mathbf{x}^*)$. Logo, $\mathcal{F}_0(\mathbf{x}^*) \cap \mathcal{H}_0(\mathbf{x}^*) = \emptyset$. $\square$

B.2.2.1 Condições de otimalidade para restrições de desigualdade

A fim de demonstrar o teorema das Condições Necessárias de Fritz-John, utiliza-se o seguinte teorema, cuja demonstração será omitida, tendo em vista que requer um estudo mais aprofundado sobre convexidade.

Teorema B.17. (Teorema de Gordan) Seja $A \in \mathbb{R}^{m \times n}$ uma matriz real. Então exatamente um dos seguintes sistemas tem solução:

- $A\mathbf{x} < \mathbf{0}$, para algum $\mathbf{x} \in \mathbb{R}^n$;
- $A^t \mathbf{p} = \mathbf{0}$ e $\mathbf{p} \geq \mathbf{0}$, para algum $\mathbf{p} \in \mathbb{R}^m - \{\mathbf{0}\}$.

Teorema B.18. (Condições Necessárias de Fritz-John) Considere o problema (B.12). Seja $\mathbf{x}^*$ um ponto factível tal que f e h_j são diferenciáveis em $\mathbf{x}^*$, para $j \in \mathcal{A}(\mathbf{x}^*)$, e h_j é contínua em $\mathbf{x}^*$, para $j \notin \mathcal{A}(\mathbf{x}^*)$. Se $\mathbf{x}^*$ é um mínimo local, então existem escalares ν_0^* e

ν_j^* , para cada $j \in \mathcal{A}(\mathbf{x}^*)$, tais que

$$\nu_0^* \nabla f(\mathbf{x}^*) + \sum_{j \in \mathcal{A}(\mathbf{x}^*)} \nu_j^* \nabla h_j(\mathbf{x}^*) = \mathbf{0} \quad (\text{B.13a})$$

$$\nu_0^*, \nu_j^* \geq 0, \quad \text{para } j \in \mathcal{A}(\mathbf{x}^*) \quad (\text{B.13b})$$

$$(\nu_0^*, \boldsymbol{\nu}_{\mathcal{A}(\mathbf{x}^*)}^*) \neq (0, \mathbf{0}) \quad (\text{B.13c})$$

em que $\boldsymbol{\nu}_{\mathcal{A}(\mathbf{x}^*)}^*$ é o vetor cujas componentes são ν_j^* , com $j \in \mathcal{A}(\mathbf{x}^*)$. Se, adicionalmente, h_j for diferenciável em $\mathbf{x}^*$, para todo $j \notin \mathcal{A}(\mathbf{x}^*)$, as condições supracitadas podem ser escritas da seguinte forma:

$$\nu_0^* \nabla f(\mathbf{x}^*) + \sum_{j=1}^p \nu_j^* \nabla h_j(\mathbf{x}^*) = \mathbf{0} \quad (\text{B.14a})$$

$$\nu_j^* h_j(\mathbf{x}^*) = 0, \quad \text{para } j = 1, \dots, p \quad (\text{B.14b})$$

$$\nu_0^*, \nu_j^* \geq 0, \quad \text{para } j = 1, \dots, p \quad (\text{B.14c})$$

$$(\nu_0^*, \boldsymbol{\nu}^*) \neq (0, \mathbf{0}) \quad (\text{B.14d})$$

em que $\boldsymbol{\nu}^*$ é o vetor cujas componentes são ν_j^* , para $j = 1, \dots, p$.

Demonstração. Como $\mathbf{x}^*$ é ponto de mínimo local, pelo teorema B.16 não existe $\mathbf{d} \in \mathbb{R}^n$ tal que $\nabla f(\mathbf{x}^*)^t \mathbf{d} < 0$ e $\nabla h_j(\mathbf{x}^*)^t \mathbf{d} < 0$ para todo $j \in \mathcal{A}(\mathbf{x}^*)$. Seja A a matriz cujas linhas são $\nabla f(\mathbf{x}^*)^t$ e $\nabla h_j(\mathbf{x}^*)^t$, com $j \in \mathcal{A}(\mathbf{x}^*)$. Então, pelo que foi dito, o sistema $A\mathbf{d} < \mathbf{0}$ não possui solução. Logo, pelo teorema de Gordan (B.17), existe um vetor não-nulo $\mathbf{p} \geq \mathbf{0}$, tal que $A^t \mathbf{p} = \mathbf{0}$. Chamando as coordenadas de $\mathbf{p}$ de ν_0^* e ν_j^* , para $j \in \mathcal{A}(\mathbf{x}^*)$, obtemos as equações (B.13). Quando h_j é diferenciável em $\mathbf{x}^*$, para $j \notin \mathcal{A}(\mathbf{x}^*)$, a forma equivalente em (B.14) pode ser obtida fazendo $\nu_j^* = 0$, para $j \notin \mathcal{A}(\mathbf{x}^*)$. $\square$

A partir das Condições Necessárias de Fritz-John, pode-se enunciar o seguinte teorema, de fundamental importância teórica em otimização não-linear.

Teorema B.19. (Condições Necessárias de Karush-Kuhn-Tucker para restrições de desigualdade) Considere o problema (B.12). Seja $\mathbf{x}^*$ um ponto factível tal que f e h_j são diferenciáveis em $\mathbf{x}^*$, para $j \in \mathcal{A}(\mathbf{x}^*)$, e h_j é contínua em $\mathbf{x}^*$, para $j \notin \mathcal{A}(\mathbf{x}^*)$. Suponha, ainda, que o conjunto $\{\nabla h_j(\mathbf{x}^*)\}_{j \in \mathcal{A}(\mathbf{x}^*)}$ é linearmente independente (isto é, $\mathbf{x}^*$ é um ponto regular). Se $\mathbf{x}^*$ é um mínimo local, então existem escalares ν_j^* , para cada $j \in \mathcal{A}(\mathbf{x}^*)$, tais que

$$\nabla f(\mathbf{x}^*) + \sum_{j \in \mathcal{A}(\mathbf{x}^*)} \nu_j^* \nabla h_j(\mathbf{x}^*) = \mathbf{0} \quad (\text{B.15a})$$

$$\nu_j^* \geq 0, \quad \text{para } j \in \mathcal{A}(\mathbf{x}^*) \quad (\text{B.15b})$$

Se, adicionalmente, h_j for diferenciável em $\mathbf{x}^*$, para todo $j \notin \mathcal{A}(\mathbf{x}^*)$, as condições supracitadas podem ser escritas da seguinte forma:

$$\nabla f(\mathbf{x}^*) + \sum_{j=1}^p \nu_j^* \nabla h_j(\mathbf{x}^*) = \mathbf{0} \quad (\text{B.16a})$$

$$\nu_j^* h_j(\mathbf{x}^*) = 0, \quad \text{para } j = 1, \dots, p \quad (\text{B.16b})$$

$$\nu_j^* \geq 0, \quad \text{para } j = 1, \dots, p \quad (\text{B.16c})$$

Demonstração. Pelas Condições Necessárias de Fritz-John (B.18), existem reais não todos nulos $\bar{\nu}_0^* \geq 0$ e $\bar{\nu}_j^* \geq 0$, com $j \in \mathcal{A}(\mathbf{x}^*)$, tais que

$$\bar{\nu}_0^* \nabla f(\mathbf{x}^*) + \sum_{j \in \mathcal{A}(\mathbf{x}^*)} \bar{\nu}_j^* \nabla h_j(\mathbf{x}^*) = \mathbf{0}. \quad (\text{B.17})$$

Observe que deve-se ter $\bar{\nu}_0^* > 0$ pois, caso contrário, o conjunto $\{\nabla h_j(\mathbf{x}^*)\}_{j \in \mathcal{A}(\mathbf{x}^*)}$ seria linearmente dependente. Dividindo a equação (B.17) por $\bar{\nu}_0^*$ e fazendo $\nu_j^* = \frac{\bar{\nu}_j^*}{\bar{\nu}_0^*}$, para cada $j \in \mathcal{A}(\mathbf{x}^*)$, obtém-se as equações (B.15). Quando h_j é diferenciável em $\mathbf{x}^*$, para $j \notin \mathcal{A}(\mathbf{x}^*)$, a forma equivalente em (B.16) pode ser obtida fazendo $\nu_j^* = 0$, para $j \notin \mathcal{A}(\mathbf{x}^*)$. $\square$

Do ponto de vista geométrico, as condições necessárias de KKT dizem que o oposto do vetor gradiente da função objetivo pode ser escrito como combinação linear não-negativa dos gradientes das restrições ativas em um ponto de mínimo local.

As condições necessárias de KKT também admitem uma versão de segunda ordem, bem como condições suficientes que garantem que um ponto é de mínimo local. Estes resultados serão apresentados na próxima seção, com o problema de otimização no formato geral.

B.2.3 Restrições de igualdade e desigualdade

Os resultados delineados nas duas seções anteriores servem como base para enunciar as condições de otimalidade para o problema na forma geral:

$$\begin{aligned} &\text{minimizar} && f(\mathbf{x}) \\ &\text{sujeito a} && \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ &&& \mathbf{h}(\mathbf{x}) \leq \mathbf{0} \end{aligned} \quad (\text{B.18})$$

em que $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ e $\mathbf{x} \in \Omega \subseteq \mathbb{R}^n$, com Ω um conjunto aberto.

Os resultados enunciados a seguir constituem a versão das condições de otimalidade para

o problema de otimização com restrições de igualdade e desigualdade. As demonstrações serão omitidas, uma vez que podem ser feitas combinando as ideias apresentadas nas seções anteriores.

Teorema B.20. (Condições Necessárias de Karush-Kuhn-Tucker) Considere o problema (B.18). Seja $\mathbf{x}^*$ um ponto factível tal que f , $\mathbf{g}$ e h_j são diferenciáveis em $\mathbf{x}^*$, para $j \in \mathcal{A}(\mathbf{x}^*)$, e h_j é contínua em $\mathbf{x}^*$, para $j \notin \mathcal{A}(\mathbf{x}^*)$. Suponha, ainda, que o conjunto $\{\nabla g_1(\mathbf{x}^*), \dots, \nabla g_m(\mathbf{x}^*)\} \cup \{\nabla h_j(\mathbf{x}^*)\}_{j \in \mathcal{A}(\mathbf{x}^*)}$ é linearmente independente (isto é, $\mathbf{x}^*$ é um ponto regular). Se $\mathbf{x}^*$ é um mínimo local, então existem escalares unicamente determinados λ_i^* , para $i = 1, \dots, m$ e ν_j^* , para cada $j \in \mathcal{A}(\mathbf{x}^*)$, tais que

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{x}^*) + \sum_{j \in \mathcal{A}(\mathbf{x}^*)} \nu_j^* \nabla h_j(\mathbf{x}^*) = \mathbf{0} \quad (\text{B.19a})$$

$$\nu_j^* \geq 0, \quad \text{para } j \in \mathcal{A}(\mathbf{x}^*) \quad (\text{B.19b})$$

Se, adicionalmente, h_j for diferenciável em $\mathbf{x}^*$, para todo $j \notin \mathcal{A}(\mathbf{x}^*)$, as condições supracitadas podem ser escritas da seguinte forma:

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^p \nu_j^* \nabla h_j(\mathbf{x}^*) = \mathbf{0} \quad (\text{B.20a})$$

$$\nu_j^* h_j(\mathbf{x}^*) = 0, \quad \text{para } j = 1, \dots, p \quad (\text{B.20b})$$

$$\nu_j^* \geq 0, \quad \text{para } j = 1, \dots, p \quad (\text{B.20c})$$

Por fim, se as funções f , $\mathbf{g}$ e $\mathbf{h}$ forem de classe C^2 , e $V(\mathbf{x}^*)$ for a matriz cujas linhas são os gradientes do conjunto $\{\nabla g_1(\mathbf{x}^*), \dots, \nabla g_m(\mathbf{x}^*)\} \cup \{\nabla h_j(\mathbf{x}^*)\}_{j \in \mathcal{A}(\mathbf{x}^*)}$, então

$$\mathbf{v}^t \left(\nabla^2 f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^*) + \sum_{j=1}^p \nu_j^* \nabla^2 h_j(\mathbf{x}^*) \right) \mathbf{v} \geq 0, \quad \text{para todo } \mathbf{v} \in \mathcal{N}(V(\mathbf{x}^*))$$

isto é, a matriz $\nabla^2 f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^*) + \sum_{j=1}^p \nu_j^* \nabla^2 h_j(\mathbf{x}^*)$ é semidefinida positiva em $\mathcal{N}(V(\mathbf{x}^*))$.

Assim como no caso das restrições de igualdade, os valores λ_i^* e ν_j^* são denominados *multiplicadores de Lagrange* ou *variáveis duais* do problema.

Teorema B.21. (Condições Suficientes de Karush-Kuhn-Tucker) Considere o problema (B.18), em que as funções f , $\mathbf{g}$ e $\mathbf{h}$ são de classe C^2 . Suponha, ainda, que para um dado ponto $\mathbf{x}^*$, existam vetores $\boldsymbol{\lambda}^* = (\lambda_1^*, \dots, \lambda_m^*)^t$ e $\boldsymbol{\nu}^* = (\nu_1^*, \dots, \nu_p^*)^t$ satisfazendo as equações (B.20). Defina $\mathcal{A}^+(\mathbf{x}^*) = \{j \in \mathcal{A}(\mathbf{x}^*) : \nu_j^* > 0\}$, $\mathcal{A}^0(\mathbf{x}^*) = \{j \in \mathcal{A}(\mathbf{x}^*) : \nu_j^* = 0\}$

e o cone

$$\mathcal{S}(\mathbf{x}^*) = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{d} \neq \mathbf{0}, \nabla g_i(\mathbf{x}^*)^t \mathbf{d} = 0, \text{ para } i = 1, \dots, m, \\ \nabla h_j(\mathbf{x}^*)^t \mathbf{d} = 0, \text{ para } j \in \mathcal{A}^+(\mathbf{x}^*), \nabla h_j(\mathbf{x}^*)^t \mathbf{d} \leq 0, \text{ para } j \in \mathcal{A}^0(\mathbf{x}^*)\}.$$

Se

$$\mathbf{v}^t \left(\nabla^2 f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^*) + \sum_{j=1}^p \nu_j^* \nabla^2 h_j(\mathbf{x}^*) \right) \mathbf{v} > 0, \text{ para todo } \mathbf{v} \in \mathcal{S}(\mathbf{x}^*),$$

então $\mathbf{x}^*$ é um mínimo local estrito do problema (B.18).

B.3 A função Lagrangiana

Agora que já foram apresentadas as condições necessárias e suficientes para problemas de otimização com restrições de igualdade e desigualdade, pode-se questionar se existe uma função que, em certo sentido, recupere as condições de Karush-Kuhn-Tucker (KKT). Observa-se que, nas condições de KKT, dado um ponto $\mathbf{x}^*$ factível (isto é, $\mathbf{g}(\mathbf{x}^*) = \mathbf{0}$ e $\mathbf{h}(\mathbf{x}^*) \leq \mathbf{0}$), se $\mathbf{x}^*$ é mínimo local, deve satisfazer a expressão $\nabla f(\mathbf{x}^*) + J\mathbf{g}(\mathbf{x}^*)^t \boldsymbol{\lambda}^* + J\mathbf{h}(\mathbf{x}^*)^t \boldsymbol{\nu}^* = \mathbf{0}$, para determinados vetores $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ e $\boldsymbol{\nu}^* \in \mathbb{R}^p$.

Assim sendo, considere a seguinte *função Lagrangiana* nas variáveis $\mathbf{x}$, $\boldsymbol{\lambda}$ e $\boldsymbol{\nu}$:

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \nu_j h_j(\mathbf{x}) \quad (\text{B.21})$$

O gradiente dessa função assume a forma

$$\nabla \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \begin{pmatrix} \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \\ \nabla_{\boldsymbol{\lambda}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \\ \nabla_{\boldsymbol{\nu}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \end{pmatrix}$$

em que

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \nabla f(\mathbf{x}) + J\mathbf{g}(\mathbf{x})^t \boldsymbol{\lambda} + J\mathbf{h}(\mathbf{x})^t \boldsymbol{\nu} \quad (\text{B.22a})$$

$$\nabla_{\boldsymbol{\lambda}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \mathbf{g}(\mathbf{x}) \quad (\text{B.22b})$$

$$\nabla_{\boldsymbol{\nu}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \mathbf{h}(\mathbf{x}) \quad (\text{B.22c})$$

B.3.1 Jogos e a noção de dualidade

A partir de agora, observa-se como a função Lagrangiana (B.21) se relaciona com o problema de otimização original (B.18), conforme Griva et al. (2009). Para tanto, imagine que dois indivíduos, A e B, disputam um jogo. O indivíduo A dispõe de um conjunto de estratégias X , enquanto o indivíduo B dispõe de um conjunto de estratégias Y . Neste jogo, o indivíduo A escolhe uma estratégia $\mathbf{x} \in X$ e o indivíduo B uma estratégia $\mathbf{y} \in Y$, as quais são reveladas concomitantemente. Em decorrência das estratégias envolvidas, o indivíduo A paga uma quantia $Q(\mathbf{x}, \mathbf{y})$ para o indivíduo B (que pode ser negativa, o que implicaria no indivíduo B pagar para o indivíduo A).

Considere que cada jogador deseja maximizar os ganhos obtidos. Para tanto, eles buscam otimizar tomando como base o pior cenário possível. Na pior hipótese, dado que o indivíduo A fez uma escolha $\mathbf{x}$, seu pagamento para o indivíduo B será $Q^*(\mathbf{x}) = \max_{\mathbf{y} \in Y} Q(\mathbf{x}, \mathbf{y})$ ⁴. Assim, para minimizar o pagamento no pior caso possível, o indivíduo A escolhe a estratégia que resolve o problema

$$\min_{\mathbf{x} \in X} Q^*(\mathbf{x})$$

o qual é um problema dito min-max, pois pode ser escrito como

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} Q(\mathbf{x}, \mathbf{y}).$$

Diz-se que este problema a ser resolvido pelo indivíduo A é o *problema primal*, e que $Q^*(\mathbf{x})$ é a *função primal*. Por outro lado, o indivíduo B está preocupado que os pagamentos efetuados pelo indivíduo A sejam baixos. No pior cenário, se o indivíduo B escolhe uma estratégia $\mathbf{y}$, o indivíduo A lhe fará um pagamento $Q_*(\mathbf{y}) = \min_{\mathbf{x} \in X} Q(\mathbf{x}, \mathbf{y})$. Assim, para maximizar seus ganhos, o indivíduo B escolhe a estratégia que resolve o problema

$$\max_{\mathbf{y} \in Y} Q_*(\mathbf{y})$$

o qual é dito um problema max-min, já que pode ser expresso por

$$\max_{\mathbf{y} \in Y} \min_{\mathbf{x} \in X} Q(\mathbf{x}, \mathbf{y}).$$

A este problema resolvido pelo indivíduo B chamamos de *problema dual*, e $Q_*(\mathbf{y})$ é a *função dual*. Os problemas apresentados são ditos o dual um do outro. Observe que, dados $\mathbf{x} \in X$ e $\mathbf{y} \in Y$,

$$Q_*(\mathbf{y}) = \min_{\mathbf{x} \in X} Q(\mathbf{x}, \mathbf{y}) \leq Q(\mathbf{x}, \mathbf{y}) \leq \max_{\mathbf{y} \in Y} Q(\mathbf{x}, \mathbf{y}) = Q^*(\mathbf{x}). \quad (\text{B.23})$$

⁴A rigor, deveria-se utilizar o supremo no lugar do máximo, pois pode ser que não exista $\mathbf{y} \in Y$ para o qual o valor máximo é obtido.

Este resultado é conhecido como *dualidade fraca* e significa que os valores da função dual são sempre menores ou iguais ao da função primal. Como consequência da dualidade fraca, tem-se que

$$\max_{\mathbf{y} \in Y} \min_{\mathbf{x} \in X} \mathcal{Q}(\mathbf{x}, \mathbf{y}) \leq \min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} \mathcal{Q}(\mathbf{x}, \mathbf{y}),$$

isto é, o valor ótimo do problema max-min é um limitante inferior para o problema min-max, ao passo que o valor ótimo do problema min-max é um limitante superior para o problema max-min.

Definição B.16. Um ponto $(\mathbf{x}^*, \mathbf{y}^*) \in X \times Y$ satisfaz a *condição de ponto de sela* para $\mathcal{Q}$ se

$$\mathcal{Q}(\mathbf{x}^*, \mathbf{y}) \leq \mathcal{Q}(\mathbf{x}^*, \mathbf{y}^*) \leq \mathcal{Q}(\mathbf{x}, \mathbf{y}^*), \text{ para todo } \mathbf{x} \in X \text{ e } \mathbf{y} \in Y.$$

Isto significa que $\mathbf{y}^*$ maximiza $\mathcal{Q}$ quando $\mathbf{x}$ está fixo em $\mathbf{x}^*$, e que $\mathbf{x}^*$ minimiza $\mathcal{Q}$ quando $\mathbf{y}$ está fixo em $\mathbf{y}^*$.

Agora, pode-se questionar sob quais hipóteses o valor da função primal é igual ao valor da função dual. O próximo lema indica que isso somente ocorre na existência de um ponto de sela.

Lema B.5. (Dualidade Forte) A condição

$$\max_{\mathbf{y} \in Y} \min_{\mathbf{x} \in X} \mathcal{Q}(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} \mathcal{Q}(\mathbf{x}, \mathbf{y})$$

ocorre se, e somente se, existe um par $(\mathbf{x}^*, \mathbf{y}^*)$ que satisfaz a condição de ponto de sela.

Demonstração. Ver (Griva et al., 2009, p. 529). □

B.3.2 Dualidade Lagrangiana

Considere, agora, a função Lagrangiana associada ao problema (B.18):

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f(\mathbf{x}) + \boldsymbol{\lambda}^t \mathbf{g}(\mathbf{x}) + \boldsymbol{\nu}^t \mathbf{h}(\mathbf{x}),$$

com $\mathbf{x} \in \Omega \subseteq \mathbb{R}^n$, $\boldsymbol{\lambda} \in \mathbb{R}^m$ e $\boldsymbol{\nu} \in \mathbb{R}^p$, $\boldsymbol{\nu} \geq \mathbf{0}$. Verifica-se que o problema (B.18) pode ser escrito da forma min-max, fazendo uma analogia da função $\mathcal{L}$ com a função $\mathcal{Q}$. Defina

$$\mathcal{L}^*(\mathbf{x}) = \max_{\substack{\boldsymbol{\lambda} \in \mathbb{R}^m \\ \boldsymbol{\nu} \geq \mathbf{0}}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \max_{\substack{\boldsymbol{\lambda} \in \mathbb{R}^m \\ \boldsymbol{\nu} \geq \mathbf{0}}} [f(\mathbf{x}) + \boldsymbol{\lambda}^t \mathbf{g}(\mathbf{x}) + \boldsymbol{\nu}^t \mathbf{h}(\mathbf{x})].$$

Se $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ e $\mathbf{h}(\mathbf{x}) \leq \mathbf{0}$, então $\boldsymbol{\lambda}^t \mathbf{g}(\mathbf{x}) = 0$ e $\boldsymbol{\nu}^t \mathbf{h}(\mathbf{x}) \leq 0$, de modo que a expressão acima será maximizada quando $\boldsymbol{\nu}^t \mathbf{h}(\mathbf{x}) = 0$. Isto pode ser feito tomando, por exemplo, $\boldsymbol{\nu} = \mathbf{0}$.

Neste caso, $\mathcal{L}^*(\mathbf{x}) = f(\mathbf{x})$. Por outro lado, se ocorrer $g_i(\mathbf{x}) \neq 0$, para algum $i = 1, \dots, m$, ou $h_j(\mathbf{x}) > 0$, para algum $j = 1, \dots, p$, tomando λ_i ou ν_j (dependendo do caso que ocorre) tão grande quanto se quiser, conclui-se que $\mathcal{L}^*(\mathbf{x}) = +\infty$. Assim,

$$\mathcal{L}^*(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & \text{se } \mathbf{g}(\mathbf{x}) = \mathbf{0} \text{ e } \mathbf{h}(\mathbf{x}) \leq \mathbf{0} \\ +\infty, & \text{caso contrário} \end{cases}$$

Desconsiderando as regiões em que $\mathcal{L}^*(\mathbf{x}) = +\infty$, então o problema $\min_{\mathbf{x} \in \Omega} \mathcal{L}^*(\mathbf{x})$ é equivalente ao problema original (B.18). Isto significa que o problema original pode ser escrito como o seguinte problema min-max:

$$\min_{\mathbf{x} \in \Omega} \max_{\substack{\boldsymbol{\lambda} \in \mathbb{R}^m \\ \boldsymbol{\nu} \geq \mathbf{0}}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}).$$

Utilizando as ideias de dualidade apresentadas pode-se, portanto, estabelecer um problema dual relacionado ao problema de otimização a ser resolvido. Para tanto, considera-se a seguinte *função dual*, para $\boldsymbol{\nu} \geq \mathbf{0}$:

$$\mathcal{L}_*(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \min_{\mathbf{x} \in \Omega} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}). \quad (\text{B.24})$$

O *problema dual* associado é expresso por:

$$\max_{\substack{\boldsymbol{\lambda} \in \mathbb{R}^m \\ \boldsymbol{\nu} \geq \mathbf{0}}} \mathcal{L}_*(\boldsymbol{\lambda}, \boldsymbol{\nu})$$

ou, mais explicitamente,

$$\max_{\substack{\boldsymbol{\lambda} \in \mathbb{R}^m \\ \boldsymbol{\nu} \geq \mathbf{0}}} \min_{\mathbf{x} \in \Omega} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}). \quad (\text{B.25})$$

Assim sendo, um ponto $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\nu}})$ será dito *dual factível* quando $\bar{\mathbf{x}} \in \Omega$, $\bar{\boldsymbol{\nu}} \geq \mathbf{0}$ e $\bar{\mathbf{x}}$ minimizar $\mathcal{L}(\mathbf{x}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\nu}})$. Como na seção anterior, podemos estabelecer relações entre os problemas primal e dual. Do que foi exposto, sabemos que se $\mathbf{x}$ é factível para o problema primal, então $\mathcal{L}^*(\mathbf{x}) = f(\mathbf{x})$. Além disso, se $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\nu}})$ for factível para o dual, então $\mathcal{L}_*(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\nu}}) = \mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\nu}})$. Pela dualidade fraca, deve-se ter $\mathcal{L}_*(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\nu}}) \leq \mathcal{L}^*(\mathbf{x})$. Este resultado é a versão do teorema da dualidade fraca para o problema (B.18).

Teorema B.22. (Dualidade Fraca) Seja $\mathbf{x}$ uma solução factível para o problema primal (B.18), e seja $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\nu}})$ uma solução factível para o problema dual a ele associado. Então

$$\mathcal{L}_*(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\nu}}) = f(\bar{\mathbf{x}}) + \bar{\boldsymbol{\lambda}}^t \mathbf{g}(\bar{\mathbf{x}}) + \bar{\boldsymbol{\nu}}^t \mathbf{h}(\bar{\mathbf{x}}) \leq f(\mathbf{x}) = \mathcal{L}^*(\mathbf{x}).$$

Em decorrência da dualidade fraca, tem-se o seguinte corolário.

Corolário B.4.

$$\max_{\substack{\lambda \in \mathbb{R}^m \\ \nu \geq 0}} \mathcal{L}_*(\lambda, \nu) \leq \min_{x \in \mathcal{X}} f(x)$$

em que $\mathcal{X} = \{x \in \Omega : g(x) = 0 \text{ e } h(x) \leq 0\}$.

Pode-se demonstrar que, em problemas convexos, a solução ótima do problema primal apresenta mesmo valor que a solução ótima do problema dual. Este fato, contudo, pode não ocorrer em problemas não-convexos. Neste caso, diz-se que ocorre um *gap de dualidade*. O exemplo a seguir ilustra este fato.

Exemplo B.4. Considere o seguinte problema primal

$$\begin{aligned} &\text{minimizar} && f(x) = -x^4 - x^2 \\ &\text{sujeito a} && x = 2 \\ &&& x \in \Omega = \{x \in \mathbb{R} : -1 \leq x \leq 3\} \end{aligned}$$

Naturalmente, a solução ótima é $x^* = 2$, e o valor ótimo é $f(x^*) = -20$. A função Lagrangiana associada ao problema é

$$\mathcal{L}(x, \lambda) = -x^4 - x^2 + \lambda(x - 2), \quad x \in \Omega, \quad \lambda \in \mathbb{R}.$$

A função dual é expressa por

$$\mathcal{L}_*(\lambda) = \min_{x \in \Omega} \mathcal{L}(x, \lambda) = \min_{-1 \leq x \leq 3} [-x^4 - x^2 + \lambda(x - 2)].$$

Observe que $\frac{\partial^2 \mathcal{L}}{\partial x^2}(x, \lambda) = -12x^2 - 2 < 0$, para todo $x \in \Omega$. Isto é, a função a ser minimizada para obtenção da função dual, para λ fixo, é côncava. Isto significa que ela assume mínimo em algum dos extremos do intervalo $[-1, 3]$. Comparando os valores, conclui-se que:

$$\mathcal{L}_*(\lambda) = \begin{cases} -2 - 3\lambda, & \text{se } \lambda \geq 22 \\ -90 + \lambda, & \text{se } \lambda < 22. \end{cases}$$

Analisando o gráfico da função dual, apresentado na Figura B.3, conclui-se que o máximo da função dual ocorre para $\lambda^* = 22$. Neste caso, o valor ótimo da função dual é $\mathcal{L}_*(\lambda^*) = -68 < -20 = f(x^*)$. O valor do *gap* de dualidade é a diferença $f(x^*) - \mathcal{L}_*(\lambda^*) = -20 - (-68) = 48$.

O teorema da dualidade forte garante que os valores ótimos primal e dual são iguais se, e somente se, existe um ponto de sela. Por outro lado, o exemplo anterior mostra que um ponto desse tipo pode não existir, mesmo quando os problemas primal e dual têm solução,

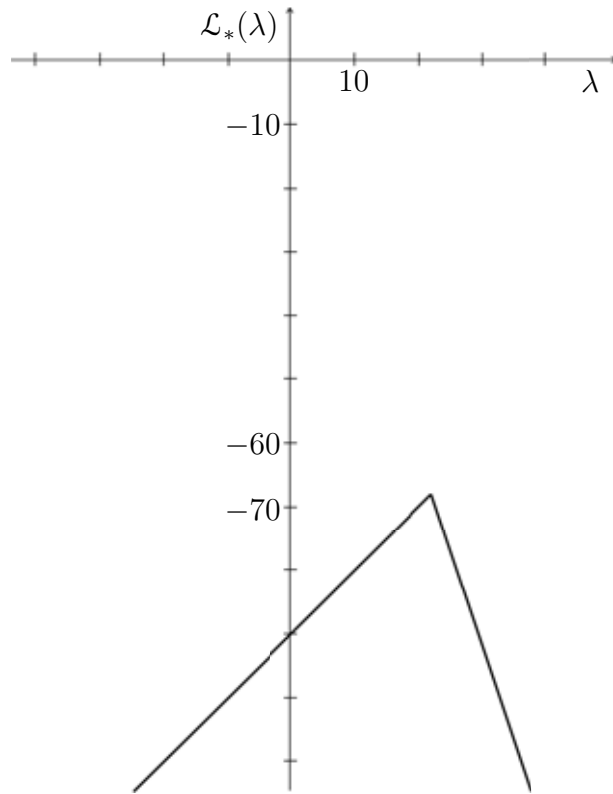


Figura B.3: Gráfico da função $\mathcal{L}_*(\lambda)$ no exemplo B.4.

levando a uma distância entre as soluções nos espaços primal e dual – o *gap* de dualidade. Apesar disso, a função dual apresenta a interessante propriedade de ser sempre côncava, independentemente das características do problema primal. Em contrapartida, a característica de diferenciabilidade nem sempre é verificada, como também mostra o exemplo anterior. A seguir, enuncia-se o Teorema do Ponto de Sela, o qual é uma consequência da Dualidade Forte.

Teorema B.23. (Teorema do Ponto de Sela) Considere o problema (B.18) e $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu})$ a função Lagrangiana associada. Então existem $\mathbf{x}^* \in \mathcal{X}$ e $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, com $\boldsymbol{\nu}^* \geq \mathbf{0}$, satisfazendo a condição de ponto de sela, isto é,

$$\mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}, \boldsymbol{\nu}) \leq \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) \leq \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$$

para todo $\mathbf{x} \in \mathcal{X}$ e $(\boldsymbol{\lambda}, \boldsymbol{\nu})$, com $\boldsymbol{\nu} \geq \mathbf{0}$, se, e somente se, $\mathbf{x}^*$ e $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ são soluções para os problemas primal e dual, respectivamente, com *gap* de dualidade nulo.

De acordo com Griva et al. (2009), em alguns casos é possível estabelecer uma teoria de *dualidade local*, isto é, através da restrição do problema à vizinhança de uma solução ótima local. Sob algumas hipóteses é possível garantir, localmente, a existência de um ponto de sela e, portanto, que os valores ótimos do primal e do dual são iguais.

A fim de desenvolver as ideias de dualidade local, considere que as funções do problema (B.18) são de classe C^2 . Seja $\mathbf{x}^*$ um mínimo local regular, e $\boldsymbol{\lambda}^*$ e $\boldsymbol{\nu}^* \geq \mathbf{0}$ os vetores de multiplicadores de Lagrange tais que

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) = \nabla f(\mathbf{x}^*) + J\mathbf{g}(\mathbf{x}^*)^t \boldsymbol{\lambda}^* + J\mathbf{h}(\mathbf{x}^*)^t \boldsymbol{\nu}^* = \mathbf{0}.$$

Assume-se que a matriz hessiana da função Lagrangiana em relação à $\mathbf{x}$,

$$\nabla_{\mathbf{x}\mathbf{x}}^2 \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) = \nabla^2 f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 g_i(\mathbf{x}^*) + \sum_{j=1}^p \nu_j^* \nabla^2 h_j(\mathbf{x}^*),$$

é definida positiva. Observe que esta hipótese é mais forte que a das condições suficientes de KKT, em que essa mesma matriz apenas precisa ser definida positiva em um cone, relacionado às restrições ativas.

Analisando o sistema de equações expresso por

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \nabla f(\mathbf{x}) + J\mathbf{g}(\mathbf{x})^t \boldsymbol{\lambda} + J\mathbf{h}(\mathbf{x})^t \boldsymbol{\nu} = \mathbf{0},$$

sabe-se que $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ é uma solução conhecida e que a jacobiana em relação a $\mathbf{x}$ no referido ponto, a matriz $\nabla_{\mathbf{x}\mathbf{x}}^2 \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, é definida positiva. Portanto, pelo Teorema da Função Implícita existem vizinhanças de $\mathbf{x}^*$ e $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, bem como uma função ξ de classe C^2 , tal que todo ponto $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ na vizinhança de $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ possui um único ponto correspondente $\xi(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \mathbf{x}$ que resolve o sistema. Além disso, $\nabla_{\mathbf{x}\mathbf{x}}^2 \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu})$ é definida positiva. Seja X tal vizinhança de $\mathbf{x}^*$. Então todo $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ suficientemente próximo de $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ possui um único $\mathbf{x}$ correspondente que é mínimo global de $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu})$ em X . Se restringirmos nossa atenção a essa vizinhança, o problema primal pode ser reescrito como:

$$\begin{aligned} & \text{minimizar} && f(\mathbf{x}) \\ & \text{sujeito a} && \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ & && \mathbf{h}(\mathbf{x}) \leq \mathbf{0} \\ & && \mathbf{x} \in X. \end{aligned}$$

A função dual, neste caso, é dada por

$$\mathcal{L}_*(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \min_{\mathbf{x} \in X} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}).$$

Para $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ suficientemente próximo de $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, tem-se que

$$\mathcal{L}_*(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \mathcal{L}(\xi(\boldsymbol{\lambda}, \boldsymbol{\nu}), \boldsymbol{\lambda}, \boldsymbol{\nu}) = f(\xi(\boldsymbol{\lambda}, \boldsymbol{\nu})) + \boldsymbol{\lambda}^t \mathbf{g}(\xi(\boldsymbol{\lambda}, \boldsymbol{\nu})) + \boldsymbol{\nu}^t \mathbf{h}(\xi(\boldsymbol{\lambda}, \boldsymbol{\nu}))$$

em que $\xi(\boldsymbol{\lambda}, \boldsymbol{\nu})$ é o único ponto na vizinhança X tal que

$$\nabla f(\xi(\boldsymbol{\lambda}, \boldsymbol{\nu})) + J\mathbf{g}(\xi(\boldsymbol{\lambda}, \boldsymbol{\nu}))^t \boldsymbol{\lambda} + J\mathbf{h}(\xi(\boldsymbol{\lambda}, \boldsymbol{\nu}))^t \boldsymbol{\nu} = \mathbf{0}.$$

O problema

$$\underset{\boldsymbol{\nu} \geq \mathbf{0}}{\text{maximizar}} \mathcal{L}_*(\boldsymbol{\lambda}, \boldsymbol{\nu})$$

restrito às vizinhanças citadas anteriormente, é chamado *problema dual local*. Neste caso, pode-se garantir que os valores ótimos dos problemas primal e dual são iguais.

Teorema B.24. (Dualidade Local) Considere o problema (B.18) em que as funções envolvidas são de classe C^2 . Seja $\mathbf{x}^*$ um ponto de mínimo local regular e $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, com $\boldsymbol{\nu}^* \geq \mathbf{0}$, os vetores de multiplicadores de Lagrange associados. Suponha, ainda, que $\nabla_{\mathbf{x}\mathbf{x}}^2 \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ é definida positiva. Então $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, com $\boldsymbol{\nu}^* \geq \mathbf{0}$, é uma solução do problema dual local, $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ é uma solução dual factível, e os valores ótimos primal e dual são iguais.

Apêndice C

Banco de Dados dos Sistemas

Este apêndice apresenta os dados dos sistemas elétricos utilizados nos resultados numéricos para o PDE-PPV e PDE-PV-RT. Na coluna tipo de barra, a barra *slack* é representada pelo número 3, as barras PV são representadas pelo número 2 e as barras PQ pelo número 1. Os dados são apresentados em pu, em uma base de 100 MVA. Os valores nulos nas colunas de *tap* indicam que é linha de transmissão. No caso do sistema de 6 Barras de Wood e Wollenberg, os dados das demandas de potência ativa e reativa foram alterados, para que coincidissem com o artigo de Ravi et al. (2006). No caso dos sistemas IEEE-14 Barras e IEEE-30 Barras, as potências geradas do banco de dados foram alteradas para coincidir com a do ponto inicial do método PDPCBLM, pois o cálculo dos coeficientes-B assume potências geradas não-nulas nas barras de geração (no cálculo dos elementos da matriz Ψ).

C.1 Sistema de 6 Barras de Wood e Wollenberg

Tabela C.1: Dados de Barra – 6 Barras

Barra	Tipo	Dados de Barra								
		V	θ	P_G	Q_G	$\underline{Q}$	$\overline{Q}$	P_L	Q_L	b^{sh}
1	3	1,05	0	0	0	-100	100	0	0	0
2	2	1,05	0	0,5	0	-100	100	0	0	0
3	2	1,07	0	0,6	0	-100	100	0	0	0
4	1	1	0	0	0	0	0	0,65	0,65	0
5	1	1	0	0	0	0	0	0,65	0,65	0
6	1	1	0	0	0	0	0	0,65	0,65	0

Tabela C.2: Dados de Linha – 6 Barras

Dados de Linha					
Linha		r	x	b^{sh}	Tap
1	2	0,1	0,2	0,04	0
1	4	0,05	0,2	0,04	0
1	5	0,08	0,3	0,06	0
2	3	0,05	0,25	0,06	0
2	4	0,05	0,1	0,02	0
2	5	0,1	0,3	0,04	0
2	6	0,07	0,2	0,05	0
3	5	0,12	0,26	0,05	0
3	6	0,02	0,1	0,02	0
4	5	0,2	0,4	0,08	0
5	6	0,1	0,3	0,06	0

C.2 IEEE-14 Barras

Tabela C.3: Dados de Barra – 14 Barras

Dados de Barra										
Barra	Tipo	V	θ	P_G	Q_G	$\underline{Q}$	$\overline{Q}$	P_L	Q_L	b^{sh}
1	3	1,06	0	0,55	-0,169	0	0,1	0	0	0
2	2	1,045	-4,98	0,35	0,424	-0,4	0,5	0,217	0,127	0
3	2	1,01	-12,72	0,475	0,234	0	0,4	0,942	0,19	0
4	1	1,019	-10,33	0	0	0	0	0,478	-0,039	0
5	1	1,02	-8,78	0	0	0	0	0,076	0,016	0
6	2	1,07	-14,22	0,275	0,122	-0,06	0,24	0,112	0,075	0
7	1	1,062	-13,37	0	0	0	0	0	0	0
8	2	1,09	-13,36	0,275	0,174	-0,06	0,24	0,3	0	0
9	1	1,056	-14,94	0	0	0	0	0,295	0,166	0,19
10	1	1,051	-15,1	0	0	0	0	0,09	0,058	0
11	1	1,057	-14,79	0	0	0	0	0,035	0,018	0
12	1	1,055	-15,07	0	0	0	0	0,061	0,016	0
13	1	1,05	-15,16	0	0	0	0	0,135	0,058	0
14	1	1,036	-16,04	0	0	0	0	0,149	0,05	0

Tabela C.4: Dados de Linha – 14 Barras

Dados de Linha					
Linha		r	x	b^{sh}	Tap
1	2	0,01938	0,05917	0,0528	0
1	5	0,05403	0,22304	0,0492	0
2	3	0,04699	0,19797	0,0438	0
2	4	0,05811	0,17632	0,034	0
2	5	0,05695	0,17388	0,0346	0
3	4	0,06701	0,17103	0,0128	0
4	5	0,01335	0,04211	0	0
4	7	0	0,20912	0	0,978
4	9	0	0,55618	0	0,969
5	6	0	0,25202	0	0,932
6	11	0,09498	0,1989	0	0
6	12	0,12291	0,25581	0	0
6	13	0,06615	0,13027	0	0
7	8	0	0,17615	0	0
7	9	0	0,11001	0	0
9	10	0,03181	0,0845	0	0
9	14	0,12711	0,27038	0	0
10	11	0,08205	0,19207	0	0
12	13	0,22092	0,19988	0	0
13	14	0,17093	0,34802	0	0

C.3 IEEE-30 Barras

Tabela C.5: Dados de Barra – 30 Barras

Barra	Tipo	V	θ	Dados de Barra						
				P_G	Q_G	$\underline{Q}$	$\overline{Q}$	P_L	Q_L	b^{sh}
1	3	1,06	0	1,25	-0,161	0	0,1	0	0	0
2	2	1,043	-5,48	0,5	0,5	-0,4	0,5	0,217	0,127	0
3	1	1,021	-7,96	0	0	0	0	0,024	0,012	0
4	1	1,012	-9,62	0	0	0	0	0,076	0,016	0
5	2	1,01	-14,37	0,325	0,37	-0,4	0,4	0,942	0,19	0
6	1	1,01	-11,34	0	0	0	0	0	0	0
7	1	1,002	-13,12	0	0	0	0	0,228	0,109	0
8	2	1,01	-12,1	0,225	0,373	-0,1	0,4	0,3	0,3	0
9	1	1,051	-14,38	0	0	0	0	0	0	0
10	1	1,045	-15,97	0	0	0	0	0,058	0,02	0,19
11	2	1,082	-14,39	0,2	0,162	-0,06	0,24	0	0	0
12	1	1,057	-15,24	0	0	0	0	0,112	0,075	0
13	2	1,071	-15,24	0,26	0,106	-0,06	0,24	0	0	0
14	1	1,042	-16,13	0	0	0	0	0,062	0,016	0
15	1	1,038	-16,22	0	0	0	0	0,082	0,025	0
16	1	1,045	-15,83	0	0	0	0	0,035	0,018	0
17	1	1,04	-16,14	0	0	0	0	0,09	0,058	0
18	1	1,028	-16,82	0	0	0	0	0,032	0,009	0
19	1	1,026	-17	0	0	0	0	0,095	0,034	0
20	1	1,03	-16,8	0	0	0	0	0,022	0,007	0
21	1	1,033	-16,42	0	0	0	0	0,175	0,112	0
22	1	1,033	-16,41	0	0	0	0	0	0	0
23	1	1,027	-16,61	0	0	0	0	0,032	0,016	0
24	1	1,021	-16,78	0	0	0	0	0,087	0,067	0,043
25	1	1,017	-16,35	0	0	0	0	0	0	0
26	1	1	-16,77	0	0	0	0	0,035	0,023	0
27	1	1,023	-15,82	0	0	0	0	0	0	0
28	1	1,007	-11,97	0	0	0	0	0	0	0
29	1	1,003	-17,06	0	0	0	0	0,024	0,009	0
30	1	0,992	-17,94	0	0	0	0	0,106	0,019	0

Tabela C.6: Dados de Linha – 30 Barras

Dados de Linha					
Linha		r	x	b^{sh}	Tap
1	2	0,0192	0,0575	0,0528	0
1	3	0,0452	0,1652	0,0408	0
2	4	0,057	0,1737	0,0368	0
3	4	0,0132	0,0379	0,0084	0
2	5	0,0472	0,1983	0,0418	0
2	6	0,0581	0,1763	0,0374	0
4	6	0,0119	0,0414	0,009	0
5	7	0,046	0,116	0,0204	0
6	7	0,0267	0,082	0,017	0
6	8	0,012	0,042	0,009	0
6	9	0	0,208	0	0,978
6	10	0	0,556	0	0,969
9	11	0	0,208	0	0
9	10	0	0,11	0	0
4	12	0	0,256	0	0,932
12	13	0	0,14	0	0
12	14	0,1231	0,2559	0	0
12	15	0,0662	0,1304	0	0
12	16	0,0945	0,1987	0	0
14	15	0,221	0,1997	0	0
16	17	0,0524	0,1923	0	0
15	18	0,1073	0,2185	0	0
18	19	0,0639	0,1292	0	0
19	20	0,034	0,068	0	0
10	20	0,0936	0,209	0	0
10	17	0,0324	0,0845	0	0
10	21	0,0348	0,0749	0	0
10	22	0,0727	0,1499	0	0
21	22	0,0116	0,0236	0	0
15	23	0,1	0,202	0	0
22	24	0,115	0,179	0	0
23	24	0,132	0,27	0	0
24	25	0,1885	0,3292	0	0
25	26	0,2544	0,38	0	0
25	27	0,1093	0,2087	0	0
28	27	0	0,396	0	0,968
27	29	0,2198	0,4153	0	0
27	30	0,3202	0,6027	0	0
29	30	0,2399	0,4533	0	0
8	28	0,0636	0,2	0,0428	0
6	28	0,0169	0,0599	0,013	0