

UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"  
FACULDADE DE ENGENHARIA DE BAURU  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

ALGORITMO GENÉTICO DIRECIONADO E SUA ANÁLISE  
ESTOCÁSTICA PARA RESOLUÇÃO DO PROBLEMA DE DESPACHO  
ECONÔMICO

LÍVIA TERESA MINAMI BORGES

Bauru  
2024

LÍVIA TERESA MINAMI BORGES

**ALGORITMO GENÉTICO DIRECIONADO E SUA ANÁLISE  
ESTOCÁSTICA PARA RESOLUÇÃO DO PROBLEMA DE DESPACHO  
ECONÔMICO**

Defesa de doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Estadual Paulista "Júlio De Mesquita Filho", UNESP, Campus de Bauru, como requisito para obtenção do título de Doutor em Engenharia Elétrica.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Edilaine Martins Soler.

Coorientador: Prof. Dr. Leonardo Nepomuceno

Bauru  
2024

B732a

Borges, Livia Teresa Minami  
Algoritmo Genético Direcionado e sua Análise Estocástica para  
Resolução do Problema de Despacho Econômico / Livia Teresa  
Minami Borges. -- Bauru, 2024  
154 p.

Tese (doutorado) - Universidade Estadual Paulista (UNESP),  
Faculdade de Engenharia, Bauru  
Orientadora: Edilaine Martins Soler  
Coorientador: Leonardo Nepomuceno

1. otimização. 2. meta-heurística. 3. convergência. 4. taxa de  
convergência. 5. fluxo de carga linearizado. I. Título.

## IMPACTO POTENCIAL DA PESQUISA

O desenvolvimento do Algoritmo Genético Direcionado (AGD) e do Algoritmo Evolutivo Adaptativo (AEA) para resolver o Problema de Despacho Econômico considerando os efeitos de pontos de carregamento de válvula e zonas de operação proibida (PDEPVZ) e também a transmissão (PDEPVZ-RR) possui um impacto significativo em diversas áreas, tanto do ponto de vista econômico quanto social.

Do ponto de vista econômico, a aplicação dos algoritmos propostos permite uma utilização mais eficiente dos recursos energéticos, resultando em uma operação mais econômica das usinas e, conseqüentemente, na redução dos custos de produção de energia elétrica. Além disso, a resolução eficaz do PDEPVZ e do PDEPVZ-RR com os novos algoritmos pode minimizar os custos associados à operação fora das zonas proibidas e aos impactos dos pontos de carregamento de válvula, otimizando a alocação de geração de energia. Concessionárias de energia que implementam essas metodologias podem se tornar mais competitivas no mercado, oferecendo tarifas mais atrativas e melhorando seus índices de eficiência.

Do ponto de vista social, a capacidade de planejar e operar de maneira otimizada reduz a probabilidade de falhas no sistema elétrico, garantindo um fornecimento de energia mais estável e confiável para a sociedade. A melhoria na operação dos sistemas de geração de energia pode evitar paradas repentinas, que causam prejuízos a diversos setores, incluindo a produção de bens e serviços essenciais. Além disso, uma operação mais eficiente e econômica pode levar à redução dos custos de produção e, conseqüentemente, ao alívio das tarifas pagas pelos consumidores, contribuindo para a acessibilidade e sustentabilidade financeira do fornecimento de energia.

## POTENTIAL IMPACT OF THE RESEARCH

The development of the Directed Genetic Algorithm (DGA) and the Adaptive Evolutionary Algorithm (AEA) to solve the Economic Dispatch Problem considering Valve-Point Loading Effects and Prohibited Operating Zones (EDPVPZ), as well as the transmission constraints (NC-EDPVPZ), has a significant impact on various areas, both from an economic and a social perspective.

From an economic standpoint, the application of the proposed algorithms allows for a more efficient use of energy resources, resulting in a more economical operation of power plants and, consequently, reducing the costs of electricity production. Additionally, the effective resolution of the EDPVPZ and NC-EDPVPZ with the new algorithms can minimize costs associated with operating outside prohibited zones and the impacts of valve-point loading effects, optimizing the allocation of power generation. Energy utilities that implement these methodologies may become more competitive in the market by offering more attractive tariffs and improving their efficiency indices.

From a social perspective, the ability to plan and operate in an optimized manner reduces the likelihood of failures in the power system, ensuring a more stable and reliable energy supply for society. Improvements in the operation of energy generation systems can prevent sudden outages, which cause losses across various sectors, including the production of essential goods and services. Moreover, a more efficient and economical operation can lead to lower production costs and, consequently, alleviate the tariffs paid by consumers, contributing to the accessibility and financial sustainability of energy supply.

**ATA DA DEFESA PÚBLICA DA TESE DE DOUTORADO DE LÍVIA TERESA MINAMI BORGES, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, DA FACULDADE DE ENGENHARIA - CÂMPUS DE BAURU.**

Aos 14 dias do mês de outubro do ano de 2024, às 08:30 horas, por meio de Videoconferência, realizou-se a defesa de TESE DE DOUTORADO de LÍVIA TERESA MINAMI BORGES, intitulada **Algoritmo Genético Direcionado e sua Análise Estocástica para Resolução do Problema de Despacho Econômico**. A Comissão Examinadora foi constituída pelos seguintes membros: Profa. Dra. EDILAINE MARTINS SOLER (Orientador(a) - Participação Virtual) do(a) Departamento de Matemática / Faculdade de Ciências de Bauru UNESP, Prof. Dr. ANDRE CHRISTOVAO PIO MARTINS (Participação Virtual) do(a) Departamento de Engenharia Elétrica / Universidade Estadual Paulista (UNESP) - Faculdade de Ciências, Câmpus de Bauru, Prof. Dr. ANTONIO ROBERTO BALBO (Participação Virtual) do(a) Departamento de Matemática / Universidade Estadual Paulista (UNESP) - Faculdade de Ciências, Câmpus de Bauru, Prof. Dr. DIEGO NUNES DA SILVA (Participação Virtual) do(a) Departamento de Matemática / Instituto Federal de Educação, Ciência e Tecnologia de São Paulo - IFSP, Prof. Dr. LEANDRO BATISTA MORGADO (Participação Virtual) do(a) Departamento de Matemática / Universidade Federal de Santa Catarina. Após a exposição pela doutoranda e arguição pelos membros da Comissão Examinadora que participaram do ato, de forma presencial e/ou virtual, a discente recebeu o conceito final: APROVADA \_\_\_\_\_. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelo(a) Presidente(a) da Comissão Examinadora.

Profa. Dra. EDILAINE MARTINS SOLER

*Dedico este trabalho ao meu papito Celso (in memoriam) que tinha um coração maior que ele, à minha mãe Tereza que me deu suporte físico e emocional para enfrentar os desafios, às minhas amadas filhas Larisse e Milene por serem minha fonte de energia e amor, ao meu marido Fabiano por estar sempre ao meu lado e por me fazer sorrir mesmo nos momentos mais difíceis e por fim à minha família e amigos, pois cada um do seu jeitinho contribuiu para que este sonho de voltar a estudar se tornasse realidade.*

# Agradecimento

Agradeço primeiramente a Deus, que me orienta e me dá forças para seguir.

Ao meu marido, companheiro e melhor amigo Fabiano pelo apoio, pelas conversas durante as caminhadas no fim de tarde, pelas palavras de encorajamento enquanto tomávamos milhares de garrafas de café e por estar sempre ao meu lado.

Às minhas meninas Larisse e Milene, por tornarem os dias de estudo mais leves e por sempre me darem muitos beijinhos quando as coisas não saíam como o esperado.

À minha cunhada Fernanda pelas risadas, pelas conversas, por me mostrar que o Python não morde e por estar sempre pronta para esclarecer minhas dúvidas de iniciante.

À minha família pelo carinho e pela torcida. Obrigada mãe pelo suporte e por me socorrer sempre que eu gritava.

À Prof<sup>a</sup>. Dr<sup>a</sup>. Edilaine Martins Soler, pela paciência, dedicação, parceria e por sempre estar sorrindo, isso torna o caminho mais suave e realmente faz toda diferença.

Ao Prof. Dr. Leonardo Nepomuceno, pela competência, entusiasmo e por fazer dessa etapa uma experiência de extremo conhecimento pra mim, não só como estudante, mas como pessoa.

Aos membros da banca, minha mais profunda gratidão pela atenção dedicada e pelos valiosos *insights* fornecidos. Com certeza suas contribuições e questionamentos enriquecerão imensamente minha pesquisa e me ajudarão a aprimorar meu trabalho.

Aos colegas que, mesmo não nos conhecendo pessoalmente, tornaram o ensino remoto mais leve, engraçado e produtivo.

Aos professores e funcionários do programa de pós-graduação em engenharia elétrica

da UNESP Bauru e do Departamento de Matemática.

Ao programa de pós-graduação da FEB em Engenharia Elétrica e a UNESP pela oportunidade de realização do curso de doutorado.

Ao IFSP pelo incentivo e aos professores da Matemática do Campus Birigui, pelo suporte e palavras de apoio.

*“A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que ninguém ainda  
pensou sobre aquilo que todo mundo vê.”*  
*(Arthur Schopenhauer)*

# Resumo

Neste trabalho, visando resolver o Problema de Despacho Econômico considerando os efeitos de pontos de carregamento de válvula e zonas de operação proibida (PDEPVZ), vamos propor o Algoritmo Genético Direcionado (AGD), um método baseado no Algoritmo Genético (AG) que, utilizando as singularidades da função objetivo e sem perder a capacidade de exploração do AG, realiza uma busca inteligente focada principalmente nessas singularidades do problema. Além disso, através de resultados da Teoria de Probabilidade e processos estocásticos provamos que o AGD converge. Para que pudéssemos validar o AGD, o algoritmo foi implementado no Python e testes foram realizados para resolver o PDEPV e o PDEPVZ para sistemas com 3, 5, 6 e 40 geradores. Propomos ainda o Algoritmo Evolutivo Adaptativo (AEA), que incorpora uma mutação adaptativa ao Algoritmo Evolutivo convencional com elitismo, e provamos que o AEA apresenta uma Taxa Média de Convergência (TMC) linear, quando aplicado à funções que são Lipschitz contínuas e que satisfazem determinadas hipóteses ( $(\mathcal{A}_1)$  ou  $(\mathcal{A}_2)$ ). Adaptações foram feitas aos teoremas relacionados à TMC, para que toda a teoria pudesse ser aplicada ao PDEPV. Além disso, para tornar o PDEPVZ ainda mais representativo, propõe-se um modelo linear para a rede de transmissão (PDEPVZ-RR), onde tanto as perdas quanto os limites de fluxo de potência são aplicados a todos os indivíduos da população. Este modelo foi incorporado ao AGD através da criação de um operador que calcula as perdas e faz o balanço de potência. Testes foram realizados para os sistemas IEEE-118 barras e IEEE-300 barras.

**Palavras-chave:** otimização; meta-heurística; singularidades; convergência; processos estocástico; mutação positiva-adaptativa; função Lipschitz contínua; fluxo de carga linearizado.

# Abstract

In this work, aiming to solve the Economic Dispatch Problem considering the effects of valve-point loading and prohibited operating zones (EDPVPZ), we propose the Directed Genetic Algorithm (DGA), a method based on the Genetic Algorithm (GA) that, by utilizing the singularities of the objective function without losing the GA's exploratory capability, performs an intelligent search primarily focused on these problem singularities. Additionally, through results from Probability Theory and stochastic processes, we prove that the DGA converges. To validate the DGA, the algorithm was implemented in Python, and tests were conducted to solve the EDPVP and EDPVPZ for systems with 3, 5, 6, and 40 generators.

We also propose the Adaptive Evolutionary Algorithm (AEA), which incorporates an adaptive mutation into the conventional Evolutionary Algorithm with elitism. We prove that the AEA exhibits a linear Average Convergence Rate (ACR) when applied to functions that are Lipschitz continuous and satisfy certain hypotheses ( $(\mathcal{A}_1)$  or  $(\mathcal{A}_2)$ ). Adaptations were made to the theorems related to ACR to apply the entire theory to the EDPVP. Additionally, to make the EDPVPZ even more representative, a linear model for the transmission network (TR-EDPVPZ) is proposed, where both losses and power flow limits are applied to all individuals in the population. This model was incorporated into the DGA through the creation of an operator that calculates losses and performs the power balance. Tests were conducted for the IEEE-118 bus and IEEE-300 bus systems.

**Keywords:** optimization; meta-heuristic; singularities; convergence; stochastic processes; positive-adaptive mutation; Lipschitz continuous function; linearized power flow.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>24</b>
<b>2</b>	<b>Modelos Matemáticos para o Problema de Despacho Econômico</b>	<b>33</b>
2.1	Formulação Clássica do PDE . . . . .	34
2.2	Formulação do PDEPV . . . . .	34
2.3	Formulação do PDEPVZ . . . . .	37
<b>3</b>	<b>Algoritmo Genético Direcionado - AGD</b>	<b>39</b>
3.1	Descrição do método . . . . .	39
3.1.1	Algoritmo Genético - AG . . . . .	39
3.1.2	Singularidades associadas ao problema . . . . .	42
3.1.3	Pseudocódigo e Operadores . . . . .	46
3.1.3.1	Pseudocódigo do AGD . . . . .	46
3.1.3.2	População Inicial . . . . .	47
3.1.3.3	Indivíduo pivô . . . . .	48
3.1.3.4	Crossover . . . . .	49
3.1.3.5	Mutação . . . . .	49
3.1.3.6	Avaliação . . . . .	49
3.1.3.7	Elitismo . . . . .	50

3.1.3.8	Critério de parada . . . . .	50
<b>4</b>	<b>Convergência</b>	<b>51</b>
4.0.1	Noções preliminares . . . . .	51
4.0.1.1	Funções contínuas em compactos . . . . .	52
4.0.1.2	Teoria de Probabilidade: alguns conceitos e propriedades . . . . .	52
4.0.1.3	Variáveis Aleatórias . . . . .	55
4.0.1.4	Cadeia de Markov . . . . .	57
4.0.2	Prova da convergência . . . . .	59
4.1	Testes numéricos . . . . .	66
4.1.1	AGD na resolução do PDEPV . . . . .	66
4.1.1.1	Sistema com 3 geradores . . . . .	66
4.1.1.2	Sistema com 5 geradores . . . . .	68
4.1.1.3	Sistema com 6 geradores . . . . .	69
4.1.1.4	Sistema com 40 geradores . . . . .	70
4.1.2	AGD na resolução do PDEPVZ . . . . .	72
4.1.2.1	Sistema com 3 geradores . . . . .	73
4.1.2.2	Sistema com 6 geradores . . . . .	74
4.1.2.3	Sistema com 40 geradores . . . . .	77
4.1.3	Influência do indivíduo pivô . . . . .	79
<b>5</b>	<b>Taxa Média de Convergência - TMC</b>	<b>81</b>
5.1	Algoritmo Evolutivo Adaptativo - AEA . . . . .	81
5.1.1	População Inicial . . . . .	83
5.1.2	Função avaliação . . . . .	83
5.1.3	Mutação . . . . .	83

5.1.4	Seleção . . . . .	84
5.1.5	Elitismo . . . . .	84
5.1.6	Critério de parada . . . . .	84
5.2	Noções preliminares . . . . .	85
5.3	TMC linear para otimização de problemas com funções Lipschitz contínuas	89
5.4	Aplicações do AEA . . . . .	96
5.4.1	Algoritmos Implementados . . . . .	96
5.4.2	Função Ackley . . . . .	98
5.4.3	Função Alpine . . . . .	102
5.4.4	Função $f(x, y) = x^4 + y^4$ . . . . .	104
5.4.5	Discussão: $x^*$ desconhecido . . . . .	105
5.4.6	PDEPV . . . . .	108
5.4.6.1	Sistema com 3 geradores . . . . .	111
5.4.6.2	Sistema com 5 geradores . . . . .	113
5.4.6.3	Sistema com 40 geradores . . . . .	115
<b>6</b>	<b>Algoritmo Genético Direcionado Adaptativo - AGDA</b>	<b>119</b>
6.1	Testes numéricos . . . . .	119
6.1.1	Sistema com 3 geradores . . . . .	119
6.1.2	Sistema com 5 geradores . . . . .	120
6.1.3	Sistema com 40 geradores . . . . .	121
<b>7</b>	<b>PDEPVZ com restrições de rede (PDEPVZ-RR)</b>	<b>124</b>
7.1	Modelo . . . . .	125
7.2	Adaptações ao AGD para resolução do PDEPVZ-RR . . . . .	130
7.2.1	Pseudocódigo . . . . .	130

7.2.2	Operador de Balanço de Potência . . . . .	130
7.2.3	Função de Avaliação . . . . .	132
7.3	Testes numéricos . . . . .	133
7.3.1	Sistema IEEE-118 barras . . . . .	134
7.3.2	Sistema IEEE-300 barras . . . . .	136
<b>8</b>	<b>Considerações finais e perspectivas futuras</b>	<b>140</b>
	<b>Trabalhos publicados</b>	<b>142</b>
	<b>Trabalhos submetidos</b>	<b>143</b>
	<b>Referências Bibliográficas</b>	<b>143</b>
<b>A</b>	<b>Apêndice</b>	<b>153</b>

# Lista de Figuras

2.1	Curva de custo de um gerador sem considerar efeitos de ponto de válvula (linha tracejada) e com efeitos de ponto de válvula (linha contínua). . . . .	35
2.2	Superfície de custo de um sistema de dois geradores considerando efeitos de ponto de carregamento de válvula. . . . .	36
2.3	Curva de custo de um gerador $i$ considerando efeitos de ponto de carregamento de válvula e com zonas de operação proibida (área hachurada). . . . .	38
4.1	Cadeia de Markov obtida para a população $POP_k$ . . . . .	60
4.2	Convergência do método AGD no Problema de Despacho Econômico considerando os efeitos de pontos de carregamento de válvula e zonas proibidas (PDEPVZ) com 6 geradores. . . . .	76
4.3	Média e desvio padrão das curvas de convergência para o AGD no PDEPVZ com 6 geradores. . . . .	77
4.4	Comparação do <i>fitness</i> médio do AGD e de todos os métodos usados nas comparações. . . . .	79
5.1	Taxa Média de Convergência do AGA (Teorema 5.3 e Corolário 5.4) e do AG em 2000 gerações, na resolução do Problema 5.4.1. . . . .	101
5.2	Taxa Média de Convergência do EEA (Teorema 5.3 e Corolário 5.4) e do EE em 2000 gerações, na resolução do Problema 5.4.1. . . . .	102
5.3	Uma comparação entre a TMC do AGA (Teorema 5.3), EEA (Teorema 5.3), AG e EE em 2000 gerações na resolução do Problema (5.4.3) com $n = 10$ . . . . .	104
5.4	Comparação entre a TMC do AGA (Teorema 5.5), EEA (Teorema 5.5), AG and EE em 500 gerações na minimização de $f(x, y) = x^4 + y^4$ . . . . .	105

5.5	Comparação entre a TMC de diferentes mutações (Teorema 5.3, $\bar{\sigma}$ e $\hat{\sigma}$ ) aplicadas ao EEA e EE na resolução do Problema 5.4.1. . . . .	107
5.6	Comparação entre a TMC de diferentes mutações (Teorema 5.3, $\bar{\sigma}$ e $\hat{\sigma}$ ) aplicadas ao EEA e EE na resolução do Problema (5.4.3). . . . .	107
5.7	Comparação entre a TMC de diferentes mutações (Teorema 5.3, $\bar{\sigma}$ e $\hat{\sigma}$ ) aplicadas ao EEA e EE na minimização de $f(x, y) = x^4 + y^4$ . . . . .	108
5.8	Comparação entre a TMC do AGA (Corolários 5.6 e 5.7) e AG em 100 gerações na resolução do Problema de Despacho Econômico considerando os efeitos de pontos de carregamento de válvula (PDEPV) com 3 unidades geradoras. . . . .	112
5.9	Comparação entre a TMC do EEA (Corolários 5.6 e 5.7) e EE em 100 gerações na resolução do PDEPV com 3 unidades geradoras. . . . .	113
5.10	Comparação entre a TMC do EEA (Corolários 5.6 e 5.7) e EE em 100 gerações na resolução do PDEPV com 3 unidades geradoras. . . . .	114
5.11	Comparação entre a TMC do EEA (Corolários 5.6 e 5.7) e EE em 100 gerações na resolução do PDEPV com 3 unidades geradoras. . . . .	115
5.12	Comparação entre a TMC do Estratégia Evolutiva Adaptativa (EEA) (Corolários 5.6 e 5.7) e Estratégia Evolutiva (EE) em 200 gerações na resolução do PDEPV com 40 unidades geradoras. . . . .	118

# Lista de Tabelas

3.1	Incidência de soluções nas singularidades do PDEPV segundo SECUI (2015) para sistemas de 10 e 13 geradores. . . . .	46
4.1	Dados da curva de custo para o PDEPV com 3 geradores. . . . .	67
4.2	Singularidades de cada um dos geradores para o PDEPV com 3 geradores. . . . .	67
4.3	Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 3 unidades geradoras. . . . .	67
4.4	Dados da curva de custo para o PDEPV com 5 geradores. . . . .	68
4.5	Singularidades de cada um dos geradores para o PDEPV com 5 geradores. . . . .	68
4.6	Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 5 unidades geradoras. . . . .	69
4.7	Dados da curva de custo para o PDEPV com 6 geradores. . . . .	69
4.8	Singularidades associadas a cada um dos geradores para o PDEPV com 6 geradores. . .	70
4.9	Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 6 unidades geradoras. . . . .	70
4.10	Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 40 unidades geradoras. . . . .	71
4.11	Zonas de operação proibida para o PDEPVZ com 3 geradores. . . . .	73
4.12	Singularidades de cada um dos geradores para o PDEPVZ com 3 geradores. . . . .	73
4.13	Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 3 unidades geradoras. . . . .	74

4.14	Zonas de operação proibida para o PDEPVZ com 6 geradores. . . . .	75
4.15	Singularidades de cada um dos geradores para o PDEPVZ com 6 geradores. . . . .	75
4.16	Comparação entre custo mínimo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 6 unidades geradoras. . . . .	76
4.17	Comparação entre custo mínimo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 40 unidades geradoras. . . . .	77
4.18	Comparação de custos entre AG, AGDsp e AGD. . . . .	79
5.1	Constantes para o PDEPV com 3 geradores. . . . .	111
5.2	Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 3 unidades geradoras. . . . .	112
5.3	Constantes para o PDEPV com 5 geradores. . . . .	113
5.4	Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 5 unidades geradoras. . . . .	114
5.5	Comparação entre custo mínimo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 40 unidades geradoras. . . . .	116
6.1	Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 3 unidades geradoras. . . . .	120
6.2	Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 5 unidades geradoras. . . . .	121
6.3	Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 40 unidades geradoras. . . . .	121
7.1	Comparação de custo mínimo, custo médio, desvio padrão (DP) e tempo de execução: sistema IEEE 118 barras. . . . .	134
7.2	Comparação de custo mínimo, custo médio, desvio padrão (DP) e tempo de execução: sistema IEEE 300 barras. . . . .	136

7.3	Resultados do teste de Wilcoxon comparando o AGD com outros algoritmos em diferentes instâncias, com base no custo mínimo, custo médio e tempo computacional. . . . .	138
7.4	Resultados do Teste de Nemenyi para todas as comparações de métodos. .	139

# Lista de Siglas

**PO** Pesquisa Operacional

**PDE** Problema de Despacho Econômico

**PDEPV** Problema de Despacho Econômico considerando os efeitos de pontos de carregamento de válvula

**PDEPVZ** Problema de Despacho Econômico considerando os efeitos de pontos de carregamento de válvula e zonas proibidas

**PDEPVZ-RR** Problema de Despacho Econômico considerando os efeitos de pontos de carregamento de válvula, zonas proibidas e restrições de rede

**MA-PDEPVZ** Problema de Despacho Econômico múltiplas áreas considerando os efeitos de pontos de carregamento de válvula, zonas proibidas

**MO-PDEPVZ** Problema de Despacho Econômico multi-objetivo considerando os efeitos de pontos de carregamento de válvula, zonas proibidas

**SEP** Sistema Elétrico de Potência

**ZOP** Zonas de Operação Proibida

**AG** Algoritmo Genético

**AGD** Algoritmo Genético Direcionado

**AGA** Algoritmo Genético Adaptativo

**AGDsp** Algoritmo Genético Direcionado sem pivô

**AGDA** Algoritmo Genético Direcionado Adaptativo

**AE** Algoritmo Evolutivo

**AEA** Algoritmo Evolutivo Adaptativo

**EE** Estratégia Evolutiva

**EEA** Estratégia Evolutiva Adaptativa

**ACO** Ant Colony Optimization

**ED** Evolução Diferencial

**EDOC** Evolução Diferencial Otimizada Colaborativa

**PSO** Particle Swarm Optimization

**EEAMC** Estratégia Evolutiva de Adaptação de Matriz de Covariância

**CJADE** Evolução Diferencial Adaptativa Caótica com Arquivo Externo Opcional

**HSES** Estratégia Evolutiva de Amostragem Híbrida

**TMC** Taxa Média de Convergência

**TMCA** Taxa Média de Convergência Alternativa

# Capítulo 1

## Introdução

A Pesquisa Operacional (PO) é uma área da matemática aplicada que tem como finalidade desenvolver modelos matemáticos e algoritmos de solução que auxiliem na tomada de decisão. Ela surgiu na Segunda Guerra Mundial, com a necessidade de se fazer um melhor gerenciamento de tarefas e recursos. Com o fim da guerra, a indústria se interessou pelas técnicas desenvolvidas, visando utilizá-las no planejamento e controle da produção, tendo chegado ao Brasil na década de 1960 (ARENALES ET AL., 2006). Quando um problema de PO consiste em minimizar ou maximizar uma função, chamada de função objetivo, satisfazendo determinadas restrições, estamos diante de um problema de otimização. Dependendo das características das funções envolvidas, o problema pode ser classificado como problema de programação linear ou não-linear. Pode-se ainda classificar um problema de otimização de acordo com a natureza de suas variáveis: se elas forem discretas, tem-se um problema combinatório, se forem contínuas, tem-se um problema contínuo e se houver ambos os tipos de variáveis, tem-se um problema inteiro misto.

Diversos métodos determinísticos têm sido propostos na literatura para resolver os problemas de otimização, tais como Método Gradiente (NESTEROV, 2015), Métodos de Ponto Interior (BERTSEKAS, 1997), Método de Newton (WOOD ET AL., 2013) e Algoritmo de Programação Quadrática (EXLER & SCHITTKOWSKI, 2007), entre outros. Uma das principais dificuldades ocorre para problemas mais complexos com função objetivo multimodal. Nesses casos, os métodos determinísticos podem ficar presos em mínimos locais. Isso ocorre pois a maioria dos métodos determinísticos usa a derivada para determinar a direção de busca e, como se sabe a partir do Cálculo Diferencial, a derivada

é um conceito usado para analisar o comportamento da função localmente, fazendo com que o método não tenha uma visão de todo o espaço de busca e acabe executando uma trajetória em direção ao mínimo local mais próximo.

Um importante problema de otimização envolvendo os sistemas elétricos de potência é o Problema de Despacho Econômico (PDE) que, segundo HAPP (1977), surgiu por volta de 1920, e visa determinar a combinação ótima de saídas de potência ativa para um conjunto de geradores do Sistema Elétrico de Potência (SEP), buscando minimizar seus custos de produção de energia, e satisfazendo determinadas restrições, tais como o atendimento de demanda e limites da capacidade de geração de cada gerador. Na formulação matemática de um PDE, pode-se considerar vários aspectos de modelagem de modo a tornar o problema mais representativo da operação e/ou planejamento real. Porém, à medida que mais aspectos são considerados no modelo do PDE, mais este problema se torna complexo e de difícil resolução. Com a evolução das técnicas matemáticas de solução de problemas de otimização, a formulação matemática do PDE tem evoluído de modo a incorporar novas características físicas do processo de produção de energia, tais como: i) a representação dos efeitos provocados pela abertura e fechamento das válvulas de admissão de vapor, chamados efeitos de pontos de carregamento de válvula (BALAMURUGAN ET AL., 2014); ii) a representação das zonas de operação proibida de determinadas unidades de geração (POURAKBARI-KASMAEI ET AL., 2019); iii) a representação de unidades geradoras que têm a capacidade de operar utilizando múltiplos combustíveis (ARUL ET AL., 2013); iv) a representação de alguns aspectos relacionados aos fluxos de potência ativa e reativa no sistema de transmissão (POURAKBARI-KASMAEI ET AL., 2019), dentre outros.

O Problema de Despacho Econômico considerando os efeitos de pontos de carregamento de válvula (PDEPV) é um problema de programação não linear, não convexo, multimodal e não diferenciável. O aspecto multimodal faz com que métodos tradicionais tenham dificuldade para a determinação de um mínimo global e a não diferenciabilidade da função objetivo impossibilita o uso direto do gradiente para a definição de direções de busca.

As Zonas de Operação Proibida (ZOP) consistem em faixas de geração nas quais o gerador não pode produzir energia, devido a possíveis problemas mecânicos de operação

nestas regiões. A função objetivo do PDEPVZ se torna mais complexa e gera ainda mais pontos não diferenciáveis para o problema (além dos pontos de válvula já considerados no PDEPV).

A representação do sistema de transmissão pode ser incluída no PDEPVZ e é realizada por meio de restrições que garantem o atendimento da demanda em cada barra (balanço de potência) e dos limites operacionais relacionados às linhas de transmissão e transformadores. Em geral, as restrições de balanço de potência em cada barra podem ser formuladas considerando apenas o balanço de potência ativa ou o balanço de potências ativa e reativa. Quando apenas o balanço de potência ativa é considerado, essas equações costumam ser linearizadas. Devido às restrições adicionais relacionadas à rede de transmissão, a dimensão do problema aumenta consideravelmente em comparação com o PDEPVZ, pois as variáveis, parâmetros e restrições devem ser consideradas para cada barra do sistema.

Diversas estratégias têm sido utilizadas para aprimorar a suavidade da função objetivo do PDEPVZ e mitigar suas descontinuidades, de modo a permitir sua resolução por abordagens de otimização determinísticas. Em PINHEIRO ET AL. (2022), os autores utilizam a função de Interpolação Polinomial Segmentada para lidar com as descontinuidades causadas pelas ZOP e tratam a função objetivo para torná-la linear, possibilitando a resolução do problema por meio de métodos baseados em gradientes. Em GONÇALVES ET AL. (2019), os autores apresentam uma abordagem determinística para resolver o Problema de Despacho Econômico Ambiental usando uma técnica de suavização para lidar com a não diferenciabilidade da função objetivo.

O termo meta-heurística foi utilizado pela primeira vez em 1986, em um artigo de Fred Glover que tratava sobre o algoritmo de Busca Tabu (BROWNLEE, 2011). Segundo MELIÁN ET AL. (2003), este termo vem do prefixo grego “meta”, que significa além de e da palavra grega “*euriskw*” que significa inventar, descobrir, encontrar. Juntas, trazem um significado de descoberta de alto nível. As meta-heurísticas são estratégias que, através da experiência anterior, conduzem o processo de busca para obter a solução de um problema, buscando explorar o espaço de soluções de forma mais eficaz. Segundo SÖRENSEN & GLOVER (2013), o termo meta-heurística pode ser definido da seguinte forma:

“Uma meta-heurística é uma estrutura algorítmica independente de pro-

blemas de alto nível que fornece um conjunto de diretrizes ou estratégias para desenvolver algoritmos de otimização heurística. O termo também é usado para se referir a uma implementação específica do problema de um algoritmo de otimização heurística de acordo com as diretrizes expressas em tal estrutura” (SÖRENSEN & GLOVER, 2013).

As características de não convexidade e não diferenciabilidade do PDEPVZ têm dificultado a aplicação direta de métodos determinísticos para a solução deste problema. Assim, abordagens de otimização não exatas, envolvendo a utilização de meta-heurísticas, têm sido propostas para a solucioná-lo. Além de lidar diretamente com essas características, as meta-heurísticas permitem uma maior exploração do espaço de busca e tendem a dificultar a convergência precoce para ótimos locais, a qual ocorre frequentemente em métodos determinísticos tradicionais. Dentre as principais meta-heurísticas utilizadas na solução de problemas de otimização complexos, pode-se destacar: os Algoritmos Genéticos (GASPAR-CUNHA ET AL., 2012), a Otimização por Enxame de Partículas (ENGELBRECHT, 2007; HOSSEINNEZHAD & BABAEI, 2013), os Algoritmos de Otimização por Gafanhotos (MANDAL & ROY, 2021), a Busca Tabu (BROWNLEE, 2011; SA-NGIAMVIBOOL ET AL., 2011), o Recozimento Simulado (GASPAR-CUNHA ET AL., 2012; VISHWAKARMA ET AL., 2012), Algoritmo do morcego (PANG ET AL., 2023), a Otimização por Colônia de Formigas (ACO-Ant Colony Optimization)(POTHIYA ET AL., 2010; SECUI, 2015), dentre muitas outras. Entre as clássicas e recentes referências sobre meta-heurísticas aplicadas na resolução do PDE, pode-se mencionar (SINGH & DHILLON, 2019; SUN ET AL., 2020; CHAVEZ ET AL., 2019; GHASEMI ET AL., 2016; SRIVASTAVA & DAS, 2020; NETO ET AL., 2017; LIU ET AL., 2016; MOHAMMADIAN ET AL., 2018; GHOLAMGHASEMI ET AL., 2019; XIN-GANG ET AL., 2020).

Apesar de existirem muitos métodos evolutivos propostos na literatura, não encontrou-se nenhum método especificamente desenvolvido para resolver o PDE, cuja prova de convergência tenha sido apresentada.

Neste trabalho propõe-se o Algoritmo Genético Direcionado (AGD) para resolver o PDEPVZ. O AGD baseia-se no Algoritmo Genético (AG) convencional, que é uma meta-heurística da classe dos algoritmos evolutivos e cuja metáfora envolve a evolução natural proposta por Charles Darwin. O método proposto busca reduzir o custo compu-

tacional através da diminuição do espaço de busca e evitar a convergência precoce, sem perder a capacidade de exploração, uma vez que realiza a busca de forma direcionada às singularidades da função objetivo. As singularidades são os pontos onde a função deixa de ser diferenciável e a motivação está no fato de que, segundo ZHAN ET AL. (2014a) e ZHAN ET AL. (2014b), os mínimos locais de um PDEPV estão nas singularidades (exceto possivelmente em um dos geradores, chamado de gerador *slack*), na grande maioria dos casos. Para problemas de pequeno porte, não haveria necessidade de nenhum método, pois todas as singularidades do problema poderiam ser investigadas e comparadas. Porém, para problemas de grande porte, tal análise teria um custo computacional muito alto e, neste caso, o AGD seria bastante útil.

Outra contribuição deste trabalho está no fato de que o AGD é interpretado em um contexto de sistemas dinâmicos estocásticos, o que permitiu provar a convergência do método através de ferramentas da Teoria de Probabilidade. Mais especificamente, demonstra-se que, com probabilidade 1, a trajetória percorrida pelo melhor indivíduo na população se aproxima de um ótimo global à medida que o número de iterações aumenta. A prova também assegura que o algoritmo é capaz de alcançar exatamente o ótimo global quando este está localizado em um ponto de singularidade da função objetivo.

Foram realizados testes numéricos com o AGD na resolução do PDEPV e do PDEPVZ para sistemas com 3, 5, 6 e 40 geradores. O método mostrou-se eficaz e obteve resultados de qualidade quando comparado com outros métodos da literatura, melhor até que o próprio AG convencional, comprovando que os operadores criados trouxeram melhorias ao método.

Algoritmos Evolutivos, como Algoritmos Genéticos (AG) (WANG ET AL., 2023; SUKKERD & WUTTIPORNPUN, 2016; LI & GAO, 2016; FAN ET AL., 2022), Estratégia Evolutiva (EE) (AUGER, 2005; BEYER, 2013) e Evolução Diferencial (ED) (ZHANG ET AL., 2024; MOLINA-PÉREZ ET AL., 2024; YANG ET AL., 2024), têm sido amplamente empregados na resolução de problemas não convexos, não diferenciáveis ou multimodais, como pode ser visto em SONG & LIN (2021), WANG & PENG (2020), SEMENOV & TERKEL (2003) e HE & YU (2001a). No entanto, mesmo com garantias formais de que a população  $X_t$  converge para o conjunto de soluções globais ótimas  $\mathcal{S}^*$ , uma questão fundamental surge: quão rapidamente a aptidão  $f(X_t)$  da população  $X_t$  converge para a

aptidão ótima  $f^*$ ?

Nos artigos DING & KANG (2001a) e HE & KANG (1999b), é apresentada uma análise de convergência em distribuição para algoritmos evolutivos com elitismo. Utilizando a teoria ergódica em cadeias de Markov e técnicas na álgebra de Banach, DING & KANG (2001a) determinam taxas de convergência de ordem exponencial. Por outro lado, HE & KANG (1999b) inicialmente analisam algoritmos evolutivos com operadores genéticos invariantes no tempo, estabelecendo um limite para a taxa de convergência no espaço geral. Para algoritmos adaptativos, eles determinam um limite para a taxa de convergência no espaço de estado finito. De forma similar, HE & YU (2001b) estudam a convergência em distribuição de algoritmos evolutivos, fornecendo uma análise teórica das condições necessárias e suficientes. Os autores também estabelecem limites superiores e inferiores para as taxas de convergência dos algoritmos evolutivos. AKIMOTO ET AL. (2020) investigam estratégias evolutivas aplicadas a funções quadráticas convexas, explorando o conceito de melhoria em um único passo. TARLOWSKI (2023) analisa a taxa de convergência assintótica através de erros de aproximação esperados durante um processo estocástico em tempo discreto, determinando um limite superior para a taxa de convergência e discutindo a relação entre a taxa de convergência no espaço objetivo e no espaço de busca.

Além disso, RUDOLPH (1997a) apresenta uma condição de convergência para casos onde não há uma sequência monotonicamente decrescente de pontos viáveis em direção ao mínimo global e utiliza a razão de erro de uma geração para fornecer limites precisos para a taxa de convergência de funções, como quadráticas, com matriz hessiana definida positiva. Recentemente, na busca por uma compreensão teórica da eficiência do algoritmo (1+1)-ES, os autores em MORINAGA ET AL. (2023) realizaram uma análise da taxa de convergência do algoritmo para uma classe mais ampla de funções convexas, especificamente aquelas que são localmente L-fortes convexas com gradiente U-Lipschitz contínuo. Outros estudos investigaram a taxa de convergência de algoritmos evolutivos, como RUDOLPH (1997b); HE & KANG (1999a); AUGER & HANSEN (2016); DING & KANG (2001b); JÄGERSKÜPPER (2006, 2007).

Considere  $f_t := \mathbb{E}[f(X_t)]$  o valor esperado de  $f(X_t)$  e  $e_t := |f_t - f^*|$ , onde  $X_t$  é uma população na iteração  $t$ . Em HE & LIN (2015) e CHEN & HE (2021), os autores observam

que a razão de erro  $e_t/e_{t-1}$  usada em RUDOLPH (1997a) contém um ruído significativo devido à aleatoriedade, tornando-a impraticável para experimentos numéricos. Portanto, eles propõem uma Taxa Média de Convergência (TMC), dada por  $R_t = 1 - (e_t/e_0)^{1/t}$ , que mede a rapidez com que o erro de aproximação de um algoritmo evolutivo converge para zero por geração. Ela incorpora a média geométrica das taxas de erro, fornecendo uma estimativa mais estável em comparação com  $e_t/e_{t-1}$ . Recentemente, TARLOWSKI (2023) utilizou a definição da TMC para mostrar que alguns algoritmos não convergem rapidamente, quando utilizados na resolução de problemas de otimização contínua não trivial. Outros trabalhos que utilizaram TMC para medir a velocidade de convergência de algoritmos evolutivos incluem DONG ET AL. (2018); DHIVYAPRABHA ET AL. (2018); JANIGA ET AL. (2019); LI ET AL. (2019).

Mais especificamente, em CHEN & HE (2021), os autores apresentam uma análise teórica da TMC em otimização contínua para métodos evolutivos e provam que a TMC é linear quando a mutação é positiva-adaptativa, mas sublinear quando a mutação é invariável ou zero-adaptativa. Para alcançar uma TMC linear, a mutação deve ser positiva-adaptativa. No entanto, até onde é do conhecimento do autor, nenhum trabalho fornece um resultado que garanta uma modificação na mutação para torná-la positiva-adaptativa e ainda apresente explicitamente um limitante inferior positivo para a TMC, exceto para o caso em que a função objetivo  $f$  é simples e a mutação ocorre apenas em uma das entradas (veja o exemplo (59) dado em (CHEN & HE, 2021, p. 216)).

Neste trabalho, com base em CHEN & HE (2021), consideramos um problema de otimização em que a função objetivo é Lipschitz contínua. Para esses problemas, propomos duas alternativas para o operador de mutação e provamos que, em ambos os casos, obtemos operadores que são positivos-adaptativos. Além disso, obtemos estimativas para o limitante inferior da TMC em termos da constante de Lipschitz e da dimensão do espaço (Teorema 5.3, Corolário 5.4 e Teorema 5.5). Os operadores de mutação propostos são integrados a um Algoritmo Evolutivo (AE) com elitismo, resultando no Algoritmo Evolutivo Adaptativo (AEA) proposto. O AEA adapta dinamicamente as taxas de mutação ao longo do processo de busca, visando alcançar uma TMC linear.

Para ilustrar os resultados teóricos apresentados, considera-se problemas de otimização tendo como função objetivo a função de Ackley, função Alpine e  $f(x, y) = x^4 + y^4$ ,

sujeitas a restrições canalizadas. Estas são funções bem conhecidas e amplamente utilizadas para testar algoritmos. Nestas aplicações, verifica-se que as funções satisfazem as condições necessárias e especifica-se as estimativas fornecidas no Teorema 5.3, Corolário 5.4 e Teorema 5.5. Alguns gráficos relacionados à taxa  $R_t$  são analisados, onde pode-se observar os resultados teóricos obtidos. Além disso, adaptações foram feitas às hipóteses do Teorema 5.3 e Corolário 5.4, para que pudessem ser aplicados ao PDEPV (Corolários 5.6 e 5.7) e testes foram realizados na resolução do PDEPV utilizando o AEA.

Fez-se ainda uma junção entre o AGD e o AEA, incorporando a mutação adaptativa ao AGD, originando o Algoritmo Genético Direcionado Adaptativo (AGDA). Testes foram realizados para o PDEPV com 3, 5 e 40 geradores. Foi possível notar que o AGDA apresenta resultados ainda melhores que os outros métodos propostos na tese.

Quando incorporamos a transmissão da rede ao PDEPVZ, temos um modelo ainda mais representativo e próximo da realidade, denominado Problema de Despacho Econômico considerando os efeitos de pontos de carregamento de válvula, zonas proibidas e restrições de rede (PDEPVZ-RR). A representação do sistema de transmissão é geralmente feita utilizando a abordagem da matriz de coeficientes  $\mathbf{B}$  (para mais detalhes ver GRAINGER & STEVENSON JR (1994), Seção 13.3). No entanto, essa abordagem ou negligencia os limites de fluxo de potência nas linhas de transmissão ou representa tais limites apenas para as linhas de interconexão, como visto nos modelos PDEPVZ-MA. Neste trabalho, um novo modelo de sistema de transmissão é desenvolvido para o PDEPVZ-RR, capaz de representar tanto as perdas quanto os limites de fluxo de potência na rede. Este modelo é implementado no AGD por meio de um novo operador, capaz de calcular as perdas e realizar o balanço do fluxo de potência. Implementações são apresentadas na resolução dos sistemas IEEE-118 barras e IEEE-300 barras.

A partir das discussões apresentadas, as principais contribuições deste trabalho são:

- Propor o AGD, um método de busca direcionada para a resolução do PDEPV e PDEPVZ;
- Provar a convergência do AGD através da interpretação da trajetória do melhor indivíduo num contexto de sistema dinâmico estocástico (Proposição 4.22 e Teorema 4.24);

- Propor o AEA, um método que incorpora a um algoritmo evolutivo um operador de mutação positivo-adaptativo;
- Demonstrar que, para uma função objetivo que seja Lipschitz contínua e satisfaça determinadas hipóteses, o AEA apresenta TMC linear (Teorema 5.3, Corolário 5.4 e Teorema 5.5);
- Fornecer explicitamente uma estimativa para o limitante inferior da TMC do AEA, em termos da constante de Lipschitz e da dimensão do problema;
- Propor o AGDA, método que incorpora a mutação adaptativa do AEA ao AGD para a resolução do PDEPV (Corolários 5.6 e 5.7);
- Desenvolver um novo modelo de sistema de transmissão para o PDEPVZ-RR, capaz de representar tanto as perdas quanto os limites de fluxo de potência na rede e incorporá-lo no AGD através de um novo operador genético chamado “operador de balanço de potência”.

Este trabalho está organizado da seguinte forma: o Capítulo 2 apresenta a formulação do PDE com efeitos de carregamento de ponto de válvula e zonas de operação proibida. No Capítulo 3, tem-se a descrição do AGD, método proposto neste trabalho e que usa as singularidades da função objetivo para realizar sua busca de forma direcionada, prova-se a sua convergência e apresenta-se os testes realizados para validar o AGD na resolução do PDEPV e do PDEPVZ. Faz-se ainda comparações com outros métodos existentes na literatura. No Capítulo 5 propõe-se o AEA, um algoritmo que incorpora uma mutação adaptativa ao AE com elitismo e apresenta-se a TMC, uma forma de medir a velocidade com que o erro médio de um algoritmo tende a zero (HE & LIN, 2015; CHEN & HE, 2021). Ainda neste capítulo, conduz-se uma análise teórica relacionada à TMC em problemas de otimização com função objetivo que seja Lipschitz contínua, prova-se que com as mutações propostas obtém-se uma TMC linear e determina-se estimativas para o limitante inferior da TMC. Além disso, aplica-se os resultados obtidos para minimizar as funções: Ackley, Alpine,  $f(x, y) = x^4 + y^4$  e na resolução do PDEPV. Já no Capítulo 6 apresenta-se o AGDA, um método que incorpora a mutação adaptativa ao AGD. No Capítulo 7 apresenta-se o modelo PDEPVZ-RR e os testes numéricos realizados. Por fim, no Capítulo 8, conclusões e perspectivas futuras são apresentadas.

## Capítulo 2

# Modelos Matemáticos para o Problema de Despacho Econômico

O Problema de Despacho Econômico (PDE) tem sido utilizado como uma ferramenta computacional fundamental para a operação e planejamento de sistemas de potência. O PDE tem como objetivo encontrar o despacho ótimo de geração de energia que atenda uma demanda de energia pré-estabelecida de modo a minimizar os custos de produção, e satisfazendo restrições associadas ao sistema de geração de energia, tais como os limites operacionais de cada unidade geradora.

Várias formulações matemáticas têm sido adotadas para o PDE de modo a representar, de forma cada vez mais detalhada, o processo de geração de energia e os limites operacionais do sistema de geração. Assim, alguns aspectos de modelagem foram sendo progressivamente considerados no PDE tais como: a representação de zonas de operação proibida para algumas unidades geradoras, a representação dos pontos de carregamento de válvula, a representação de unidades que operam com múltiplos combustíveis e a representação da rede de transmissão. A representação destes aspectos de modelagem aproxima o PDE da operação/planejamento de um sistema de geração mais realista, porém pode tornar o PDE um problema de grande dimensão e de difícil solução, devido a natureza dos problemas de otimização obtidos.

Neste capítulo, apresentam-se algumas formulações que têm sido propostas na literatura para o PDE. Na Seção 2.1, descreve-se o modelo para o PDE clássico, proposto

na década de 1920 e reformulado como um problema de otimização em HAPP (1977). Na Seção 2.2, é descrito o modelo para o PDEPV, em que é introduzida a representação dos pontos de carregamento de válvula para as unidades termelétricas e na Seção 2.3 é descrito o modelo para o PDEPVZ, em que é introduzida a representação das zonas de operação proibida.

## 2.1 Formulação Clássica do PDE

Em sua formulação clássica, o PDE consiste em determinar o despacho ótimo de potência ativa  $P_i$  de cada unidade geradora  $i$ , buscando minimizar o custo total de produção, de tal forma que as potências geradas estejam dentro dos limites operacionais de cada unidade geradora dados por  $P_i^{min}$  e  $P_i^{max}$ , respectivamente, a fim de suprir uma determinada demanda  $D$ . O PDE clássico é modelado como um problema de otimização com função objetivo convexa e quadrática, proposta pela primeira vez por STEINBERG ET AL. (1943). Assim, sendo tem-se que o PDE clássico é descrito pelo modelo (2.1.1):

$$\begin{aligned} \min_{\mathbf{P}} \quad & \sum_{i=1}^{n_G} [ a_i P_i^2 + b_i P_i + c_i ] \\ \text{s.a. :} \quad & \sum_{i=1}^{n_G} P_i = D \\ & P_i^{min} \leq P_i \leq P_i^{max}, \quad i = 1, \dots, n_G, \end{aligned} \quad (2.1.1)$$

em que  $\mathbf{P} = (P_1, \dots, P_{n_G})$  é o vetor de potências ativas geradas,  $n_G$  é a quantidade de geradores do sistema e  $a_i$ ,  $b_i$  e  $c_i$  são os coeficientes da função de custo do gerador  $i$ .

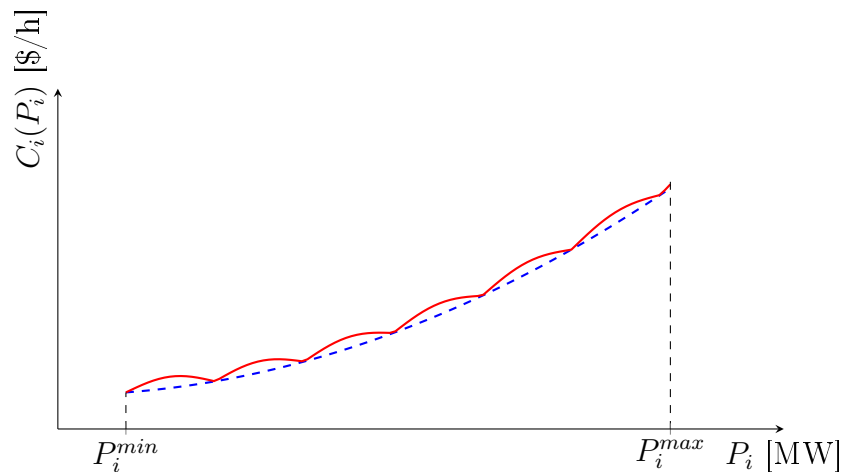
## 2.2 Formulação do PDEPV

Nos geradores termelétricos, existem válvulas de admissão de calor cuja abertura parcial provoca perdas de energia durante a passagem de vapor, aumentando os custos de produção. Além disso, quanto mais estrangulado o vapor, isto é, quanto menor a abertura da válvula, menor o rendimento da mesma. Para melhorar o rendimento global, o

vapor é controlado por múltiplas válvulas parciais. Portanto, no momento de abertura de cada uma dessas válvulas parciais há uma alteração na curva de rendimento global e, por conseguinte, na curva de custos. Este fenômeno é chamado de efeito de ponto de carregamento de válvula, e sua inserção no PDE resulta no problema de despacho denominado de PDEPV. O ponto de carregamento de válvula é definido em HAPP (1977) como o ponto imediatamente anterior à abertura da próxima válvula. Na função objetivo de custo de combustível do PDE utilizada para as unidades termelétricas, os efeitos provocados pelos pontos de carregamento de válvula podem ser expressos por meio da adição de termos modulares. Dessa forma, a inserção destes termos modulares para a representação dos pontos de carregamento de válvula torna a função objetivo do problema não convexa e não-diferenciável nestes pontos de carregamento de válvula. O problema passa, então, a ser multimodal, dificultando a determinação de um mínimo global.

Graficamente, os efeitos de ponto de carregamento de válvula para um gerador podem ser vistos na Figura 2.1, em que a linha tracejada indica a curva de custo do gerador  $i$  sem considerar os efeitos de ponto de carregamento de válvula e a linha contínua indica a função custo do mesmo gerador, considerando os efeitos de ponto de carregamento de válvula. Os pontos de não diferenciabilidade ficam evidentes no gráfico da função.

Figura 2.1: Curva de custo de um gerador sem considerar efeitos de ponto de válvula (linha tracejada) e com efeitos de ponto de válvula (linha contínua).



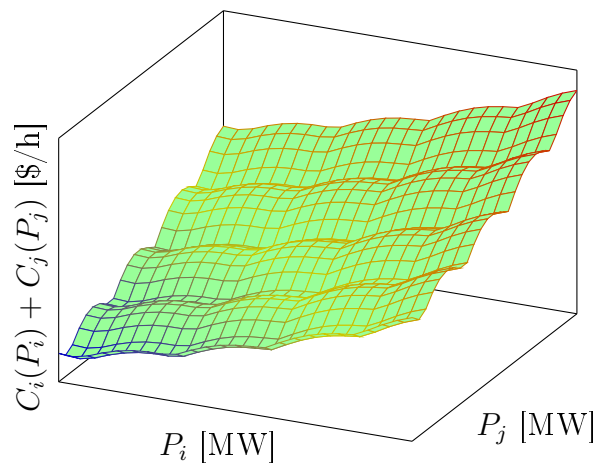
Fonte: Elaborada pelo autor.

Com a introdução dos pontos de carregamento de válvula no PDEPV, os métodos determinísticos clássicos ficam impossibilitados de serem diretamente utilizados para a sua solução, pois a não diferenciabilidade da função objetivo impede a determinação do vetor

gradiente, utilizado para determinar a direção de busca, e de matrizes hessianas, quando necessário, em alguns métodos. Mesmo que a não diferenciabilidade possa ser tratada para serem aplicados métodos determinísticos, por meio de funções aproximantes, ou por meio de reformulação através de um modelo equivalente, o PDEPV ainda continua sendo um problema multimodal, de modo que os métodos determinísticos tendem a ficar presos nos mínimos locais. Uma abordagem que tem sido amplamente utilizada para contornar este problema consiste em utilizar métodos meta-heurísticos para a solução do PDEPV. Mesmo não tendo garantias matemáticas de otimalidade de suas soluções, as meta-heurísticas podem ser capazes de escapar de mínimos locais e encontrar boas aproximações para o mínimo global. Por outro lado, uma desvantagem das meta-heurísticas é que, para problemas de grande porte, o espaço de busca pode se tornar muito amplo, dificultando a determinação do mínimo global.

Conforme se aumenta a quantidade de unidades geradoras, tem-se um aumento substancial da quantidade de mínimos locais. De fato, para um sistema de geração com duas unidades geradoras  $i$  e  $j$ , o gráfico da função custo pode ser visto na Figura 2.2, a qual apresenta uma quantidade bem maior de mínimos locais, dificultando a determinação de um mínimo global e aumentando as chances de métodos determinísticos ficarem presos. Para sistemas com muitas unidades geradoras, o problema resultante pode atingir números excessivamente altos de pontos de mínimo e máximo locais.

Figura 2.2: Superfície de custo de um sistema de dois geradores considerando efeitos de ponto de carregamento de válvula.



Fonte: Elaborada pelo autor.

A partir da discussão acima, o modelo de PDEPV possui as mesmas restrições do problema de PDE clássico, porém sua função objetivo possui um termo que corresponde

ao módulo de uma senoide, que representa os pontos de carregamento de válvula, conforme mostrado em (2.2.1):

$$\begin{aligned}
 & \min_{\mathbf{P}} \sum_{i=1}^{n_G} [ a_i \cdot P_i^2 + b_i \cdot P_i + c_i + | e_i \cdot \text{sen}(f_i \cdot (P_i^{min} - P_i)) | ] \\
 & \text{s.a.} \\
 & \sum_{i=1}^{n_G} P_i = D \\
 & P_i^{min} \leq P_i \leq P_i^{max}, \quad i = 1, \dots, n_G,
 \end{aligned} \tag{2.2.1}$$

em que as constantes  $a_i$ ,  $b_i$ ,  $c_i$ ,  $e_i$  e  $f_i$  são os coeficientes da curva de custo de combustível do gerador  $i$ .

## 2.3 Formulação do PDEPVZ

De acordo com DE OLIVEIRA ET AL. (2008), as ZOP de uma unidade termelétrica podem existir devido a faltas nas máquinas, oscilações das válvulas a vapor no eixo da máquina ou nos serviços auxiliares como caldeiras, bombas de alimentação, entre outras causas. A melhor maneira de tratar esse problema é evitando a operação destas unidades nas regiões proibidas. Assim, se forem considerados os efeitos de ponto de carregamento de válvula e as zonas de operação proibida, a função de custo de produção do problema passa a apresentar regiões de descontinuidade, como mostra a Figura 2.3, em que as zonas de operação proibida são as regiões hachuradas.

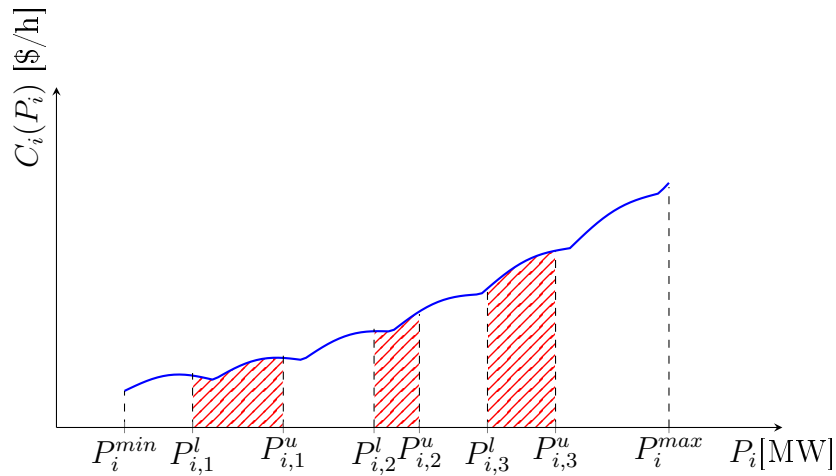
Dada uma unidade geradora  $i$ , denotando o limite mínimo e máximo da  $p$ -ésima zona proibida por  $P_{i,p}^l$  e  $P_{i,p}^u$ , respectivamente, as zonas de operação permitida do gerador  $i$  podem ser descritas conforme (2.3.1):

$$\begin{aligned}
 & P_i^{min} \leq P_i \leq P_{i,1}^l \\
 & P_{i,p-1}^u \leq P_i \leq P_{i,p}^l, \quad p = 2, \dots, z_i \\
 & P_{i,z_i}^u \leq P_i \leq P_i^{max}
 \end{aligned} \tag{2.3.1}$$

em que  $z_i$  é a quantidade de zonas de operação proibida deste gerador.

Denotando-se  $P_i^{min} = P_{i,0}^u$  e  $P_i^{max} = P_{i,z_i+1}^l$ , tem-se que o PDEPVZ pode ser mode-

Figura 2.3: Curva de custo de um gerador  $i$  considerando efeitos de ponto de carregamento de válvula e com zonas de operação proibida (área hachurada).



Fonte: Elaborada pelo autor.

lado conforme descrito em (2.3.2):

$$\begin{aligned}
 & \min_{\mathbf{P}} \sum_{i=1}^{n_G} \left[ a_i P_i^2 + b_i P_i + c_i + |e_i \sin(f_i(P_i^{min} - P_i))| \right] \\
 & \text{s.a:} \\
 & \sum_{i=1}^{n_G} P_i = D \\
 & P_i \in \bigcup_{p=0}^{z_i} [P_{i,p}^u, P_{i,p+1}^l], \quad i = 1, \dots, n_G.
 \end{aligned} \tag{2.3.2}$$

Além dos modelos apresentados neste capítulo para o PDE, podemos ainda considerar as características do sistema de transmissão, resultando em um PDEPVZ com restrições de rede (PDEPVZ-RR), que será investigado no Capítulo 7, bem como as adaptações propostas ao AGD para o tratamento das restrições adicionais associadas à rede de transmissão. Para a resolução de problemas de despacho econômico, no próximo capítulo, além do AG convencional e da definição de singularidades (conceito fundamental para a construção do AGD), são apresentados os detalhes do AGD e cada um de seus operadores.

## Capítulo 3

# Algoritmo Genético Direcionado - AGD

Este capítulo tem como objetivo descrever o AGD, método proposto neste trabalho para a solução do PDEPVZ (também pode ser aplicado ao PDEPV). Inicialmente, descreve-se AG em sua versão clássica, incluindo a “metáfora” que inspira o método e seus operadores fundamentais. Em seguida, as singularidades relacionadas ao PDEPVZ são descritas, bem como a sua importância na resolução do problema. Então apresenta-se o método AGD, método inspirado no AG clássico e que utiliza as propriedades relacionadas às singularidades da função objetivo do PDEPVZ no seu processo de busca. Além disso, a convergência do AGD é provada e testes numéricos são realizados.

### 3.1 Descrição do método

#### 3.1.1 Algoritmo Genético - AG

Os métodos evolutivos são algoritmos que se baseiam na Teoria da Evolução de Charles Darwin e utilizam o cruzamento (*crossover*) e a seleção natural (sobrevivência do mais apto) para determinar uma solução para um problema. Estes métodos têm sido amplamente utilizados na solução de problemas de engenharia, *machine learning* e inteligência artificial. Alguns exemplos de métodos evolutivos envolvem: algoritmos genéticos, programação evolutiva e estratégias evolutivas, evolução diferencial, dentre outros. Segundo GASPAR-CUNHA ET AL. (2012), os métodos evolutivos surgiram em 1960 com o trabalho do americano John Henry Holland, da Universidade de Michigan.

Neste trabalho, o método evolutivo abordado é o Algoritmo Genético (AG). Nesta seção, descreve-se a metáfora utilizada como inspiração para o desenvolvimento do método e detalha-se os principais operadores utilizados na construção do AG.

O AG tem sua inspiração na Teoria da Evolução criada por Charles Darwin no século XIX, em que um indivíduo mais adaptado ao meio tende a sobreviver em determinada condição e se reproduzir. Ao fazê-lo, suas características mais “vantajosas” passam para outras gerações (descendentes) através da reprodução. Em contrapartida, os indivíduos menos adaptados, tendem a se extinguir. De acordo com Darwin, existe uma luta constante pela sobrevivência e nessa luta somente os mais aptos sobrevivem. Este processo é chamado seleção natural e durante sua realização, podem ocorrer as chamadas mutações, evitando que os filhos sejam cópias de seus pais.

A seleção natural pode ser de três tipos: direcional (um determinado fenótipo tem sua frequência aumentada), estabilizadora (os fenótipos intermediários são selecionados e os extremos são eliminados) e disruptiva (os extremos são favorecidos se comparados com os intermediários pois os extremos são mantidos na população). Os genes fornecem todas as características que serão passadas aos descendentes, como tamanho, peso, cor, dentre outros.

A ideia do AG é gerar uma população inicial onde cada indivíduo é uma solução potencial para o problema a ser resolvido, e cada um de seus genes corresponde ao valor que uma das variáveis do problema deve assumir. Em seguida, ocorre um processo de seleção em que são definidos os pais que irão se reproduzir e gerar iterativamente indivíduos filhos a partir da população anterior, buscando sempre melhorar os indivíduos e transmitir as melhores características de cada um deles. A mutação é incorporada ao método através de uma pequena perturbação em um ou mais genes dos indivíduos filhos.

No AG, a cada iteração obtém-se uma geração de indivíduos em que cada um deles representa uma solução potencial para o problema que se deseja resolver. De forma

bastante geral, as principais etapas envolvidas no algoritmo são enumeradas a seguir:

- (i) Estabeleça uma representação cromossomial;
- (ii) Inicialize a população de indivíduos;
- (iii) Avalie cada indivíduo criado;
- (iv) Selecione os pais;
- (v) Faça o *crossover*;
- (vi) Faça a mutação;
- (vii) Avalie os filhos;
- (viii) Verifique o critério de parada.

Na etapa (i), cada indivíduo é representado por uma sequência numérica, chamada de cromossomo. Cada cromossomo representa os valores a serem assumidos por cada variável do problema de otimização que está sendo resolvido. O cromossomo é geralmente descrito por meio de uma codificação que pode ser realizada utilizando-se números reais ou binários.

Na etapa (ii), determinam-se os valores iniciais que devem ser assumidos pela população inicial. Em geral, estes valores são estabelecidos de forma aleatória.

Na etapa (iii), cada indivíduo criado deve ser avaliado por meio de uma função avaliação que mede o grau de adaptabilidade (*fitness*) do indivíduo. Esta função deve levar em consideração o valor da função objetivo e o grau de factibilidade do indivíduo, no que diz respeito ao atendimento das restrições do problema de otimização que se pretende resolver.

Na etapa (iv), a operação de seleção escolhe os pais que deverão participar no processo de reprodução. A seleção pode ser feita aleatoriamente, com pesos iguais para todos os indivíduos, ou associando cada indivíduo a uma probabilidade de ser selecionado, de acordo com o valor obtido na função de avaliação anterior. Dessa última forma, os indivíduos que tiverem melhor avaliação terão maiores chances de serem escolhidos no processo seletivo. Outra forma de o indivíduo ser selecionado envolve o operador denominado de elitismo. Por meio deste operador, os melhores indivíduos sempre são escolhidos para compor a próxima geração, sem a necessidade de cruzamento com outro indivíduo.

Na etapa (v), os novos indivíduos (filhos) são gerados a partir do cruzamento dos indivíduos selecionados (pais) mais aptos, de modo a buscar uma evolução continuada dos indivíduos em cada geração. De acordo com ENGELBRECHT (2007), este cruzamento (*crossover*) pode ser assexual (um pai gera toda a prole), sexual (dois pais são utilizados para produzir um ou dois filhos) ou multi-recombinação (mais de dois pais são utilizados para gerar um ou mais descendentes). No cruzamento sexual, uma possibilidade é escolher aleatoriamente um ponto de corte para o cromossomo dos indivíduos selecionados, realizar o corte e recombinar as partes, obtendo dois novos indivíduos (filhos).

Na etapa (vi), os indivíduos gerados (filhos) podem estar sujeitos a sofrer mutação. O operador de mutação consiste em uma perturbação aleatória nos genes dos indivíduos. O grau de incidência da mutação em cada gene do indivíduo é determinado pelo chamado coeficiente de mutação, que estabelece uma taxa (porcentagem) que define se a mutação deve ou não ocorrer em cada gene. Assim, o coeficiente de mutação precisa ser estabelecido como um número pequeno para que o algoritmo não perca as boas características genéticas já adquiridas dos indivíduos. Por outro lado, quando utilizada com coeficientes de mutação altos, a mutação permite o surgimento de indivíduos que podem ser geneticamente diversos de seus pais, promovendo maior variabilidade ao algoritmo.

Na etapa (vii), a população é atualizada com os melhores indivíduos entre a população de pais e os indivíduos filhos, tomando o cuidado de manter o tamanho da população fixo em cada geração.

Por fim, na etapa (viii), verifica-se o critério de parada do algoritmo, que pode envolver dois critérios básicos: a quantidade máxima de gerações é atingida ou ocorre uma estagnação (não melhora) dos valores de *fitness* no melhor indivíduo de cada geração. Após esta etapa, caso o critério de parada não seja atingido, o algoritmo retorna à etapa (iv) para um novo processo de seleção, e assim, sucessivamente.

### 3.1.2 Singularidades associadas ao problema

Na literatura, alguns trabalhos que propõem métodos determinísticos para resolver o PDEPV e o PDEPVZ utilizam técnicas para suavizar a função objetivo nos pontos não diferenciáveis ou reescrever a função modular de forma equivalente (PINHEIRO ET AL.,

2022). Nos métodos determinísticos, a reformulação da parcela não diferenciável da função objetivo é absolutamente necessária, pois estes métodos dependem intrinsecamente da obtenção das derivadas da função (bem como das funções associadas às restrições) para compor suas direções de busca, as quais são baseadas no gradiente e/ou hessiana. Por outro lado, essas questões de diferenciabilidade não são importantes para os algoritmos baseados em meta-heurísticas, uma vez que nestes algoritmos não se faz distinção entre pontos diferenciáveis e não-diferenciáveis. No entanto, no AGD aqui proposto, dá-se ênfase exatamente nos pontos onde a função objetivo não é diferenciável e portanto algumas definições se fazem necessárias.

Nesta seção define-se pontos específicos de uma função, denominados singularidades, e evidencia-se a sua importância na resolução do PDEPV. Mostra-se que grande parte das soluções numéricas que têm sido obtidas para o PDEPV apresenta valores de potência ativa ótima das unidades geradoras posicionados exatamente nos pontos de singularidade da função objetivo. Assim, o AGD realiza a busca da solução ótima focando principalmente nestes pontos de singularidade da função objetivo, cuja definição será dada a seguir.

**Definição 3.1.** Considere a função  $C_i : D_i \subset \mathbb{R} \rightarrow \mathbb{R}$  definida por

$$C_i(x_i) = a_i x_i^2 + b_i x_i + c_i + |e_i \text{sen}(f_i(P_i^{\min} - x_i))|,$$

em que  $D_i$  é um conjunto compacto. Um ponto  $x_i \in D_i$  é chamado de singularidade de  $C_i$  se for um ponto de carregamento de válvula do gerador  $i$ , um limitante de sua capacidade de geração ou um limitante de alguma zona de operação proibida deste gerador. Para a função

$$F(x_1, \dots, x_n) = \sum_{j=1}^n C_j(x_j), \quad n \geq 2,$$

dizemos que  $x = (x_1, \dots, x_n) \in D = D_1 \times \dots \times D_n$  é uma singularidade de  $F$  se existe uma lista de  $n - 1$  coordenadas de  $x$  tal que para todo  $x_j$  nesta lista, temos que  $x_j$  é uma singularidade de  $C_j$ .

De acordo com a definição, uma singularidade associada ao PDEPVZ (ou ao PDEPV) possui na  $i$ -ésima entrada, uma singularidade de  $C_i$ , para todas as suas coordenadas, exceto, possivelmente, em uma delas. Portanto, a fim de determinar as singularidades de  $F$ , vamos determinar as singularidades factíveis associadas a cada gerador e combinar os

valores obtidos em  $n - 1$  coordenadas. Os pontos onde a função  $C_i$  não é diferenciável são aqueles onde o termo modular se anula (derivada muda de sinal), ou seja,

$$\begin{aligned} \text{sen}(f_i \cdot (P_i^{min} - P_i)) &= 0 \\ \Rightarrow f_i \cdot (P_i - P_i^{min}) &= k_i \cdot \pi, \quad k_i \in \mathbb{Z} \\ \Rightarrow P_i &= \frac{k_i \cdot \pi}{f_i} + P_i^{min}. \end{aligned} \tag{3.1.1}$$

Para que as potências geradas respeitem os limites da capacidade de geração do gerador  $i$ , dadas por

$$P_i^{min} \leq P_i \leq P_i^{max},$$

é necessário que

$$0 \leq k_i \leq N_i,$$

onde

$$N_i = \left\lfloor f_i \cdot \frac{P_i^{max} - P_i^{min}}{\pi} \right\rfloor.$$

Acima,  $\lfloor y \rfloor$  denota a função piso, ou seja, ela transforma o número real  $y$  no maior inteiro menor ou igual a  $y$ .

Além destes pontos, devemos incluir  $P_i^{max}$  ( $P_i^{min}$  já foi considerado para  $k_i = 0$ ) e todos os limites (inferior e superior) de todas as zonas de operação proibida. E ainda excluir as singularidades que pertencerem às zonas de operação proibida. Com este processo, obtemos todas as singularidades associadas a cada gerador do sistema e, conseqüentemente, as singularidades associadas ao PDEPVZ.

O conceito de singularidade é fundamental para o AGD pois, utilizando técnicas baseadas no AG, realiza sua busca de forma direcionada às singularidades associadas ao problema. Este direcionamento nas buscas é utilizado para que possamos diminuir o espaço de busca, aumentando assim a velocidade de convergência, e se faz através das singularidades pois, segundo ZHAN ET AL. (2014a), quando resolvemos o PDEPV, na grande maioria das vezes, o mínimo global se encontra em uma singularidade do problema.

Em ZHAN ET AL. (2014a), os autores analisam primeiramente o PDEPV com duas variáveis, considerando constantes as outras variáveis do problema. Esta análise é feita através do sinal das derivadas laterais de primeira e segunda ordem da função objetivo e,

fazendo uma pequena perturbação (para mais e para menos) nas coordenadas do ponto, verificando se o ponto inicial é um mínimo local ou não. Mais precisamente, dado um ponto factível  $(x, y)$ , para um valor pequeno qualquer  $\epsilon$ , consideram-se os pontos  $(x + \epsilon, y - K \cdot \epsilon)$  e  $(x - \epsilon, y + K \cdot \epsilon)$ , onde  $K$  é um valor considerado para que o atendimento de demanda continue sendo respeitado. Em seguida, compara-se o valor da função objetivo nos três pontos, buscando concluir se  $(x, y)$  é um mínimo local. Após realizarem este procedimento com pontos que estão em cada área da região factível (singularidades em ambas as coordenadas, em apenas uma delas ou em nenhuma delas), foi possível criar uma tabela com todas as situações possíveis e caracterizar os mínimos locais do problema.

Ainda de acordo com ZHAN ET AL. (2014a), para o PDEPV com  $n$  geradores os autores observaram que se  $(x_1, \dots, x_n)$  é um mínimo local então, se tomarmos duas variáveis quaisquer, teremos um mínimo local no problema de duas variáveis, considerando as outras variáveis constantes. Isso é válido pois a função objetivo do PDEPV (também PDEPVZ) não possui termos mistos (termos envolvendo mais de uma variável). Como os mínimos dos problemas de duas variáveis foram analisados e caracterizados, os autores puderam dividir os mínimos locais de um problema de  $n$  variáveis em dois tipos:

- Tipo 1** - Existe no máximo uma unidade geradora cuja saída de potência ativa gerada não se localiza em uma singularidade;
- Tipo 2** - Existem duas ou mais unidades geradoras cujas saídas de potência ativa gerada não se localizam em singularidades.

Note que o Tipo 1 significa que o mínimo local está em um ponto singular e o Tipo 2 ocorre quando o mínimo local não está em um ponto singular da região factível. Além disso, os autores analisam a proporção entre os mínimos do Tipo 1 e Tipo 2, e detectam que o Tipo 2 ocorre em apenas 2.75% das vezes e que, portanto, este tipo de mínimo local é muito pouco provável de ocorrer. Assim, os autores concluem que na grande maioria dos casos, os mínimos locais estão em pontos de singularidade da região factível (i.e. geradores com saída de potência ativa em pontos de singularidade, exceto possivelmente em um deles). Os autores calculam ainda o erro causado ao modificarmos um ponto do Tipo 2 para o Tipo 1 e concluem que esta mudança provoca uma perda muito pequena na precisão. Em geral, o gerador que se permite gerar fora das suas singularidades é o responsável pelo ajuste de potência necessário para que ocorra o atendimento da demanda.

### 3.1.3 Pseudocódigo e Operadores

Nesta seção, são fornecidos os detalhes do AGD, método proposto neste trabalho e que, visando melhorar o desempenho do AG na resolução do PDEPVZ e aumentar a velocidade de convergência, propõe modificações ao método descrito na Seção 3.1.1, de modo a direcionar o processo de busca, priorizando mínimos locais do Tipo 1.

A motivação para a criação de novos operadores no AG são as informações apresentadas na Seção 3.1.2 sobre as singularidades da função objetivo do problema e sua estreita relação com as soluções do PDEPV (e também do PDEPVZ). Além disso, como visto na Seção 3.1.2, como as singularidades são os pontos em que a função não é diferenciável, no PDEPVZ estes pontos correspondem àqueles em que ocorre o carregamento de válvula, os extremos da capacidade de geração e também os extremos das zonas de operação proibida.

Outra motivação para o AGD está no trabalho descrito em SECUI (2015), onde os autores utilizam a busca direcionada a singularidades para resolver um PDEPV dinâmico por meio de um método que pertence ao Ant Colony Optimization (ACO), um grupo de algoritmos baseados no comportamento das formigas. Nos testes realizados, após determinarem as soluções, os autores também observaram a alta incidência de ótimos posicionados em pontos de carregamento de válvula e em limites da capacidade de geração, constatando que aproximadamente 90% das soluções encontradas estavam nas singularidades associadas ao problema. Os dados percentuais obtidos em SECUI (2015), com relação à incidência de mínimos nas singularidades, estão resumidos na Tabela 3.1.

Tabela 3.1: Incidência de soluções nas singularidades do PDEPV segundo SECUI (2015) para sistemas de 10 e 13 geradores.

	10 geradores (%)	13 geradores (%)
Pontos de válvula	48,34	72,44
$P^{min}$ ou $P^{max}$	38,33	18,27
Total	86,67	90,71

Fonte: SECUI (2015).

#### 3.1.3.1 Pseudocódigo do AGD

O AGD recebe este nome pois realiza sua busca de forma direcionada às singularidades associadas ao problema, uma vez que esta característica é levada em consideração na criação dos indivíduos e na presença do indivíduo pivô, que será descrito em detalhes

a seguir. O pseudocódigo do AGD é mostrado no Algoritmo 1, sendo que cada operador será detalhado nas próximas seções.

---

**Algoritmo 1** Pseudocódigo do AGD

---

- 1: Inicialize os parâmetros:  $t = 0$ ,  $n_I$  (número de indivíduos) e  $p_M$  (probabilidade de ocorrência da mutação);
  - 2: Determine todas as singularidades factíveis de cada gerador;
  - 3: Gere a população inicial  $POP$  de acordo com a Seção 3.1.3.2;
  - 4: Gere o indivíduo pivô  $I_p$  de acordo com a Seção 3.1.3.3 e acrescentar à  $POP$ ;
  - 5: Ajuste todos os indivíduos de  $POP$ , para que todos se tornem factíveis de acordo com Seção 3.1.3.2;
  - 6: Avalie cada indivíduo de  $POP$  de acordo com Seção 3.1.3.6;
  - 7: **while** condição de parada não é satisfeita **do**
  - 8:     **for**  $i = 1, \dots, \frac{n_I}{2}$  **do**
  - 9:         Selecione aleatoriamente dois indivíduos de  $POP$ ;
  - 10:         Aplique o crossover de acordo com a Seção 3.1.3.4;
  - 11:         Aplique a mutação com probabilidade  $p_M$  de acordo com a Seção 3.1.3.5;
  - 12:         Ajuste os indivíduos filhos, para que se tornem factíveis de acordo com Seção 3.1.3.2;
  - 13:         Compare os dois filhos obtidos com os pais e mantenha os dois melhores;
  - 14:     **end for**
  - 15:     Atualize a população;
  - 16:     Faça a mutação do indivíduo pivô  $I_p$  como descrito na Seção 3.1.3.3, compare com o pivô anterior e acrescente o melhor deles em  $POP$ ;
  - 17:      $t = t + 1$
  - 18: **end while**
  - 19: Retorne o melhor indivíduo de  $POP$  como solução.
- 

### 3.1.3.2 População Inicial

Para gerar a população inicial de tamanho  $n_I \geq 2$ , determina-se primeiramente todas as singularidades de cada gerador  $i$ ,  $i = 1, \dots, n_G$ . Feito isso, cada indivíduo será da forma  $I^j = (P_1^j, \dots, P_{n_G}^j)$ ,  $j = 1, \dots, n_I$ , e cada gene  $P_i^j$  é escolhido aleatoriamente entre as singularidades do gerador  $i$ .

Note que ao realizar a escolha de seus genes de forma aleatória, as restrições de limite de capacidade de geração são intrinsecamente satisfeitas, uma vez que as singularidades criadas já obedecem tais limites. Por outro lado, os indivíduos gerados não levam em consideração a restrição de atendimento da demanda. Para que tenhamos um indivíduo  $I^j$  factível com respeito também ao atendimento de demanda  $D$ , alguns ajustes precisam ser realizados:

- Calcula-se  $\Delta^j$ :

$$\Delta^j = D - \sum_{i=1}^{n_G} P_i^j.$$

- Se  $\Delta^j > 0$ , significa que a demanda não foi atendida em sua integralidade e é necessário aumentar um pouco a potência total despachada. Verifica-se então se existe algum gerador capaz de gerar o que já estava previsto mais essa diferença dada por  $\Delta^j$ , sem violar sua capacidade de geração. Se existir tal gerador, este será o gerador *slack*. Caso contrário, é necessário distribuir a diferença, de forma sucessiva, entre os demais geradores, começando por aquele que possui maior capacidade livre, até que a demanda seja atendida;
- Se  $\Delta^j < 0$ , significa que a potência total despachada está acima da demanda e, portanto, é necessário reduzir as gerações para o atendimento da demanda. De forma semelhante ao caso anterior, verifica-se, inicialmente, se existe algum gerador do qual pode-se retirar todo o excesso de geração sem violar nenhuma de suas restrições. De forma análoga ao caso anterior, se existir tal gerador, este será o gerador *slack*. Caso contrário, é necessário retirar a diferença, de forma sucessiva, dos demais geradores, até que a demanda seja satisfeita.

A partir do procedimento acima, obtém-se uma população inicial composta somente por indivíduos factíveis. Este ajuste também será realizado sempre antes de um indivíduo ser avaliado.

### 3.1.3.3 Indivíduo pivô

Na Seção 3.1.2 mostrou-se que no PDEPV os mínimos ocorrem, em sua grande maioria, nas singularidades da função objetivo. Entretanto, a aplicação de alguns operadores genéticos, tais como a mutação (descrita em 3.1.3.5) ou o ajuste (3.1.3.6), podem fazer com que os indivíduos saiam das singularidades. Neste caso, como as singularidades são pontos (medida nula), se todos os indivíduos saírem das singularidades, existe a probabilidade de nunca mais voltarem, comprometendo o desempenho do método. Para evitar que isso ocorra, propõe-se neste trabalho um novo operador genético chamado de indivíduo pivô  $I_p$ , que realizará sua busca apenas nas singularidades (ele caminhará apenas nas singularidades, buscando a melhor solução). Em cada geração, o indivíduo pivô tem a

possibilidade de “pular” de uma singularidade para outra, se esta for uma solução melhor que a anterior.

### 3.1.3.4 Crossover

O *crossover* é realizado para que possamos gerar novos indivíduos. Para realizá-lo, escolhe-se aleatoriamente dois indivíduos da população e realiza-se o *crossover* da seguinte forma: sorteia-se um ponto de corte e realiza-se a troca das entradas a partir deste ponto. Então, se os indivíduos escolhidos para sofrerem o *crossover* forem  $I^i = (P_1^i, \dots, P_{n_G}^i)$ , e  $I^j = (P_1^j, \dots, P_{n_G}^j)$ , e se o ponto de corte sorteado for  $k$ , então os filhos serão dados por:

$$filho_1 = (P_1^i, \dots, P_k^i, P_{k+1}^j, \dots, P_{n_G}^j) \quad \text{e} \quad filho_2 = (P_1^j, \dots, P_k^j, P_{k+1}^i, \dots, P_{n_G}^i).$$

### 3.1.3.5 Mutação

Com uma taxa de mutação  $p_M$ , cada indivíduo filho obtido pode sofrer uma mutação feita através da soma de uma perturbação a cada gene do indivíduo. Cada perturbação é obtida aleatoriamente da distribuição normal de média 0 e desvio padrão 1, denotada por  $\mathcal{N}(0, 1)$ . Assim, se o indivíduo  $I^j = (P_1^j, \dots, P_{n_G}^j)$  sofrer a mutação, este indivíduo passa a ser dado por:

$$I^j + \xi = (P_1^j + \xi_1, \dots, P_{n_G}^j + \xi_{n_G}),$$

em que  $\xi_i \in \mathcal{N}(0, 1)$ ,  $i = 1, \dots, n_G$ .

### 3.1.3.6 Avaliação

A função  $F^A(I^j)$  tem como objetivo fazer a avaliação de cada indivíduo  $I^j$ , de modo que seja possível estabelecer parâmetros de adaptabilidade (i.e. *fitness*) que permitam selecionar os indivíduos que devem sobreviver e identificar aqueles que serão extintos.

Antes de avaliar os indivíduos, é necessário executar o procedimento de ajuste descrito na Seção 3.1.3.2, de modo que o atendimento de demanda seja viabilizado, e para que as restrições impostas pela capacidade de geração de cada gerador não sejam violadas, de modo que somente indivíduos factíveis sejam avaliados, assim como realizado na criação da população inicial. Portanto, após a realização deste procedimento, não existe a neces-

sidade de penalização dos indivíduos pelo não atendimento das restrições do PDEPVZ. Nesse caso, dado que todos os indivíduos são factíveis, a função de avaliação deve levar em conta somente a função objetivo original do problema, conforme descrito a seguir:

$$F^A(I^j) = \frac{1}{C(I^j)} = \left( \sum_{i=1}^{n_G} C_i(P_i^j) \right)^{-1}$$

em que  $C(I^j)$  é o custo de combustíveis associado ao indivíduo  $I^j = (P_1^j, \dots, P_{n_G}^j)$ . Portanto, a partir da função de avaliação descrita acima, o indivíduo com menor custo terá a melhor avaliação.

### 3.1.3.7 Elitismo

O elitismo é um operador que impede a piora na função objetivo da população. No AGD, o elitismo ocorre através da sobrevivência do melhor indivíduo para a próxima geração. Além disso, após a avaliação dos filhos gerados, eles são comparados com seus pais e apenas os dois indivíduos com melhor avaliação vão sobreviver.

### 3.1.3.8 Critério de parada

No AGD, assim como na grande maioria das meta-heurísticas, utiliza-se dois critérios de parada: a quantidade de gerações (i.e. o número de iterações do método)  $n_g$  e a não melhora nas soluções das últimas iterações (estagnação), o que ocorrer primeiro.

# Capítulo 4

## Convergência

Nesta seção demonstra-se a convergência do AGD, método proposto neste trabalho para resolver o PDEPVZ, o qual se baseia no AG e cujos detalhes são descritos na Seção 3.1. Para facilitar a compreensão da prova de convergência, primeiramente são fornecidos alguns conceitos preliminares relacionados à Análise e Teoria de Probabilidade. Entretanto, estes temas não serão descritos de forma aprofundada, uma vez que fogem aos objetivos deste trabalho. Para maiores detalhes envolvendo estes temas, sugere-se as referências BARRY (2008), ELON (2009) e ALLEN (2010). Em seguida, demonstra-se que se o PDEPVZ tiver pelo menos um ótimo global e este estiver localizado em uma singularidade, o método será capaz de encontrá-lo e se o PDEPVZ tiver pelo menos um ótimo global e este não estiver localizado em uma singularidade, o AGD conseguirá chegar tão próximo quanto se queira deste ótimo global.

### 4.0.1 Noções preliminares

Nesta seção são fornecidos alguns conceitos necessários para o entendimento da prova da convergência do AGD. Para isso, criou-se as seguintes seções: funções contínuas em compactos, teoria de probabilidade, variáveis aleatórias e cadeia de Markov.

#### 4.0.1.1 Funções contínuas em compactos

**Definição 4.1.** Considere a função  $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ . Dizemos que  $f$  é contínua em  $a \in X$  se para todo  $\epsilon > 0$  existe um  $\delta > 0$  tal que

$$x \in X \text{ e } \|x - a\| \leq \delta \Rightarrow |f(x) - f(a)| \leq \epsilon.$$

Dizemos que  $f$  é contínua se for contínua em todo seu domínio.

**Definição 4.2.** Dizemos que  $K \subset \mathbb{R}^n$  é um conjunto compacto se ele for limitado e fechado.

Uma propriedade que segue da definição é que se  $K_1, \dots, K_p$  são compactos em  $\mathbb{R}^n$ , então  $K_1 \cup \dots \cup K_p$  é compacto. Esta propriedade será útil, uma vez que no problema considerado, a existência das zonas de operação proibida fazem da região factível uma união de compactos.

**Proposição 4.3.** (Weierstrass) Toda função real contínua  $f : K \rightarrow \mathbb{R}$ , definida em um conjunto compacto  $K \subset \mathbb{R}^n$ , atinge seu máximo global e seu mínimo global em  $K$ , isto é, existem pontos  $x_0, x_1 \in K$  tais que  $f(x_0) \leq f(x) \leq f(x_1)$  para todo  $x \in K$ .

#### 4.0.1.2 Teoria de Probabilidade: alguns conceitos e propriedades

Uma medida num conjunto  $\Omega$  é uma função que a cada subconjunto de  $\Omega$  associa um número real não negativo, de tal forma que a propriedade aditiva seja preservada (a medida da união de dois conjuntos disjuntos é a soma das suas medidas). Uma medida importante é a medida de Lebesgue, que atribui comprimento, área ou volume aos conjuntos, dependendo do espaço  $\Omega$ . Outra medida também bastante importante é a medida de probabilidade, que associa cada evento à sua probabilidade de ocorrer (uma medida que assume valores em  $[0, 1]$ ). Existem muitas medidas interessantes e bastante úteis, mas neste trabalho utilizaremos as duas citadas anteriormente.

Como pode ser visto em detalhes em CABRAL (2010), nem sempre é possível atribuir (de forma consistente) uma medida a todos os subconjuntos de um conjunto qualquer  $\Omega$ . Para isso, se faz necessário criar uma coleção especial  $\mathcal{F}$  de subconjuntos de  $\Omega$  onde a medida está definida, chamada  $\sigma$ -álgebra de subconjuntos de  $\Omega$ .

**Definição 4.4.** Seja  $\Omega$  um conjunto e  $\mathcal{F}$  uma família de subconjuntos de  $\Omega$ . Dizemos que  $\mathcal{F}$  é uma  $\sigma$ -álgebra em  $\Omega$  se satisfaz as seguintes propriedades:

- (i)  $\emptyset \in \mathcal{F}$ ;
- (ii)  $A \in \mathcal{F} \implies A^C \in \mathcal{F}$ , ( $A^C = \Omega \setminus A$ , complemento de  $A$  em  $\Omega$ );
- (iii)  $A_1, A_2, \dots, A_n, \dots \in \mathcal{F} \implies A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ .

Os subconjuntos de  $\Omega$  que pertencem a uma  $\sigma$ -álgebra  $\mathcal{F}$  são ditos  $\mathcal{F}$ -mensuráveis em Teoria da Medida e, no contexto de Teoria de Probabilidade, são chamados *eventos*. O par  $(\Omega, \mathcal{F})$  é chamado de espaço mensurável.

Uma  $\sigma$ -álgebra bem conhecida na reta real, denominada  $\sigma$ -álgebra de Borel e denotada por  $\mathcal{B}$  (neste caso o espaço mensurável é  $(\mathbb{R}, \mathcal{B})$ ), é quando consideramos a menor  $\sigma$ -álgebra que contém todos os conjuntos abertos da reta real e, portanto, os fechados (por causa do item (ii) da definição anterior). Os elementos desta  $\sigma$ -álgebra são chamados *borelianos*. Em termos intuitivos, um boreliano é um conjunto que pode ser obtido de um número enumerável de abertos, aplicando-se as operações união, interseção e complementar um número enumerável de vezes. Esta definição pode ser estendida naturalmente para  $\mathbb{R}^n$  considerando seus abertos, obtendo assim o espaço mensurável  $(\mathbb{R}^n, \mathcal{B})$ .

A seguir, é apresentada uma construção axiomática de probabilidade que se deve a Kolmogorov, e que permite um olhar “mais matemático” se comparado à definição “frequentista” ou “estatística” de probabilidade.

**Definição 4.5.** Dada uma  $\sigma$ -álgebra  $\mathcal{F}$  e uma função  $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ , dizemos que  $\mathbb{P}$  é uma *medida de probabilidade* no espaço de medida  $(\Omega, \mathcal{F})$  se satisfaz:

- (i)  $\mathbb{P}(\emptyset) = 0$ ,  $\mathbb{P}(\Omega) = 1$ ;
- (ii) se  $A_1, A_2, \dots \in \mathcal{F}$  são dois a dois disjuntos, então  $\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ .

(Os eventos são dois a dois disjuntos se são mutuamente exclusivos, isto é,  $A_i \cap A_j = \emptyset$  se  $i \neq j$ ). Um *espaço de probabilidade* é composto pela tripla  $(\Omega, \mathcal{F}, \mathbb{P})$ .

Quando uma função  $\mu : \mathcal{F} \rightarrow [0, \infty]$  satisfaz apenas a condição aditiva (ii) acima, diz-se simplesmente que  $\mu$  é uma medida.

**Definição 4.6.** Seja  $(\Omega, \mathcal{F}, \mathbb{P})$  um espaço de probabilidade. Se  $B \in \mathcal{F}$  e  $\mathbb{P}(B) > 0$ , a *probabilidade condicional* de  $A$  dado  $B$  é definida por:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad A \in \mathcal{F}.$$

A seguir, será enunciado o Lema de Borel-Cantelli que fornece uma ferramenta para verificar se uma sequência de eventos  $(A_n)$  ocorre infinitas vezes. Este lema diz que se a soma infinita das probabilidades dos eventos é finita, então o conjunto de todos os resultados que são “repetidos” infinitamente (muitas vezes) devem ocorrer com probabilidade zero. Dizer que a soma das probabilidades é finita significa que, a menos de um número finito, as probabilidades são estritamente menores que 1. Mais que isso, conforme  $n$  cresce, a probabilidade desses eventos precisa se tornar tão pequena quanto se queira. Sua demonstração pode ser encontrada em RUFFINO (2012).

**Lema 4.7. Borel-Cantelli** Seja  $(A_n)$  uma sequência de eventos em um espaço de probabilidade e  $B_n = A_n \cup A_{n+1} \cup \dots$ . Se

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty,$$

então

$$\mathbb{P}(B_1 \cap B_2 \cap \dots) = 0.$$

**Observação 4.8.** O conjunto  $\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ , chamado de limite superior da sequência  $(A_k)$ , é o evento “ocorrência de um número infinito dos  $A_n$ ”.

Para definir a medida de Lebesgue, define-se a medida exterior de Lebesgue, criada a partir do conceito de comprimento de intervalos. Feito isso, utilizando do método de Carathéodory ((CABRAL, 2010), p.10), obtém-se a  $\sigma$ -álgebra de Lebesgue e a medida de Lebesgue, através da extensão da medida exterior de Lebesgue sobre a  $\sigma$ -álgebra gerada. Por sua vez, esta  $\sigma$ -álgebra obtida é grande o suficiente para conter os borelianos, ou seja, os borelianos são Lebesgue mensuráveis ((CABRAL, 2010), Teorema 1.29).

**Definição 4.9.** (Medida exterior de Lebesgue) Seja  $I = (a, b) \subset \mathbb{R}$  com  $\|I\| = b - a$  e  $\|\emptyset\| = 0$ . A medida exterior de Lebesgue de  $A \subset \mathbb{R}$  é dada por:

$$\tilde{\mu}(A) = \inf \left\{ \sum_{j=0}^{\infty} \|I_j\| : (I_j)_{j \in \mathbb{N}} \text{ é uma sequência de intervalos abertos, tal que } A \subset \bigcup_{j \in \mathbb{N}} I_j \right\}.$$

A  $\sigma$ -álgebra de subconjuntos de  $\mathbb{R}$  gerada por  $\tilde{\mu}$  é dada por

$$\tilde{\mathcal{F}} = \{A \subset \mathbb{R} \mid \tilde{\mu}(E) = \tilde{\mu}(E \cap A) + \tilde{\mu}(E \setminus A) \text{ para todo } E \subset \mathbb{R}\},$$

e seus elementos são chamados de conjuntos mensuráveis a Lebesgue.

**Definição 4.10.** (Medida de Lebesgue) A medida de Lebesgue  $\mu$  é dada por  $\mu : \tilde{\mathcal{F}} \rightarrow [0, \infty)$ , onde  $\mu(A) = \tilde{\mu}(A)$ .

**Definição 4.11.** (Medida (de Lebesgue) nula) Dizemos que  $A \subset \mathbb{R}$  tem medida (de Lebesgue) nula se para todo  $\epsilon > 0$ , existe uma sequência  $(I_n)_{n \in \mathbb{N}}$  de intervalos abertos e limitados, tal que

$$A \subset \bigcup_{n=1}^{\infty} I_n \quad \text{e} \quad \sum_{n=1}^{\infty} \|I_n\| \leq \epsilon,$$

sendo  $\|I\| = b - a$ , se  $I = (a, b)$ .

Um conjunto  $A$  tem medida de Lebesgue positiva se não tiver medida de Lebesgue nula. Intuitivamente, pode-se pensar que a região possui área ou volume positivo, dependendo da dimensão em que se encontra. Todos os conceitos podem ser estendidos de forma natural para  $\mathbb{R}^n$ .

A proposição a seguir é utilizada por vários autores como uma forma de caracterizar um conjunto Lebesgue-mensurável, como sendo aquele que pode ser “espremido” entre um conjunto fechado contido nele e um conjunto aberto que o contém.

**Proposição 4.12.** Se o conjunto  $V \subset \mathbb{R}^n$  é Lebesgue-mensurável então para todo  $\epsilon > 0$  existe um aberto  $A$  e um fechado  $F$ , tal que  $F \subset V \subset A$  e  $\mu(A \setminus F) < \epsilon$ , onde  $\mu$  é a medida de Lebesgue.

#### 4.0.1.3 Variáveis Aleatórias

Uma variável aleatória é uma função  $X$  que associa os elementos de  $\Omega$  a números reais, tal que a imagem inversa de qualquer boreliano  $U \subset \mathbb{R}$  é um evento da  $\sigma$ -álgebra  $\mathcal{F}$ . Isso é exigido para que se possa associar uma probabilidade da variável aleatória  $X$  assumir seus valores num determinado boreliano  $U \subset \mathbb{R}$ . Formalmente, tem-se a seguinte definição de variável aleatória.

**Definição 4.13.** Seja  $(\Omega, \mathcal{F}, \mathbb{P})$  um espaço de probabilidade e  $(\mathbb{R}, \mathcal{B})$  um espaço mensurável. A função  $X : \Omega \rightarrow \mathbb{R}$  é uma *variável aleatória* (ou mensurável) se, para todo boreliano  $U \in \mathcal{B}$ , temos que:

$$X^{-1}(U) = \{\omega \in \Omega; X(\omega) \in U\} \in \mathcal{F}.$$

Na literatura é bastante comum ocorrer a omissão do termo  $\omega$  e denotar simplesmente por

$$\{X \in U\} \in \mathcal{F},$$

portanto faz sentido escrever

$$\mathbb{P}(\{X \in U\}).$$

**Definição 4.14.** As variáveis aleatórias  $X_1, \dots, X_n$ , definidas no mesmo espaço de probabilidade  $(\Omega, \mathcal{F}, \mathbb{P})$ , são independentes se

$$\mathbb{P}(X_1 \in U_1, X_2 \in U_2, \dots, X_n \in U_n) = \prod_{i=1}^n \mathbb{P}(X_i \in U_i),$$

onde cada  $U_i$ ,  $i = 1, \dots, n$  é um boreliano da reta real.

A seguir são dadas algumas definições sobre distribuição de uma variável aleatória.

**Definição 4.15.** A variável aleatória  $X$  é absolutamente contínua se existe uma função  $f(t) \geq 0$ , tal que

$$F_X(x) := \mathbb{P}(X \leq x) = \int_{-\infty}^x f(t) dt, \quad \forall x \in \mathbb{R}.$$

Neste caso, dizemos que  $f$  é *função densidade de probabilidade* de  $X$  ou simplesmente densidade de  $X$ . A função  $F_X$  é chamada de função de distribuição de  $X$ .

**Definição 4.16.** A variável aleatória  $X$  possui distribuição normal com média  $\bar{x}$  e desvio padrão  $\sigma$  se  $X$  tem densidade

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\bar{x})^2/2\sigma^2}, \quad x \in \mathbb{R}.$$

Notação:  $X \sim N(\bar{x}, \sigma)$ .

**Definição 4.17.** Seja  $(X_1, \dots, X_n)$  um vetor aleatório e  $F$  sua função de distribuição. Se existe uma função  $f \geq 0$ , tal que

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_n} \dots \int_{-\infty}^{x_1} f(t_1, \dots, t_n) dt_1 \dots dt_n, \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n,$$

então  $f$  é chamada densidade do vetor aleatório  $(X_1, \dots, X_n)$  ou densidade conjunta das variáveis aleatórias  $X_1, \dots, X_n$ .

A seguir serão dados alguns resultados importantes para a prova da convergência do AGD.

**Proposição 4.18.** (BARRY (2008), p. 63) Se  $X_1, \dots, X_n$  são variáveis aleatórias independentes e possuem densidades  $f_{X_1}, \dots, f_{X_n}$ , então a função

$$f(x_1, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i), \quad x_i \in \mathbb{R}$$

é densidade conjunta das variáveis aleatórias  $X_1, \dots, X_n$ , isto é,  $f = f_{X_1, \dots, X_n}$ .

**Proposição 4.19.** (BARRY (2008), p. 52) Suponha que a variável aleatória  $X$  possui função densidade  $f_X$ . Seja  $Y$  uma variável aleatória dada por  $Y = bX + c$ , onde  $b > 0$  e  $c \in \mathbb{R}$ . Então a função densidade de  $Y$  será:

$$f_Y(y) = \frac{1}{b} \cdot f_X\left(\frac{y-c}{b}\right), \quad y \in \mathbb{R}.$$

#### 4.0.1.4 Cadeia de Markov

Um *processo estocástico discreto* é uma sequência de variáveis aleatórias  $X_n$ , ou seja, para cada “tempo”  $n \in \mathbb{N}$ ,  $X_n$  é uma variável aleatória em  $(\Omega, \mathcal{F})$ . Se fixarmos  $\omega \in \Omega$  e deixarmos variar  $n$ , a função  $n \rightarrow X_n(\omega)$  é chamada de trajetória (ou realização) de  $X_n$ . Os valores assumidos pela variável aleatória são chamados *estados*. Para maiores detalhes sobre espaço de probabilidade para uma sequência de variáveis aleatórias ver por exemplo OLIVEIRA ET AL. (2017).

Uma cadeia de Markov (ou processo de Markov) é um processo estocástico discreto  $(X_n)_{n \in \mathbb{N}}$  cuja probabilidade de transição para determinados estados depende apenas do estado presente em que processo se encontra, ou seja, seu futuro não depende do passado. Mais precisamente, considere um espaço de estados com um número finito (ou enumerável) de elementos  $E = \{e_1, \dots, e_n\}$ . Então  $(X_n)_{n \in \mathbb{N}}$  é uma *cadeia de Markov* se a probabilidade condicional satisfizer

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n), \quad (4.0.1)$$

para todo  $n \geq 1$  e para toda sequência  $x_0, x_1, \dots, x_{n+1}$  de elementos do espaço de estados  $E$ . A probabilidade condicional

$$p_{ij} := \mathbb{P}(X_{n+1} = e_j | X_n = e_i)$$

é chamada *probabilidade de transição* do estado  $e_i$  para  $e_j$ , e representa a probabilidade do processo estar em  $e_j$  no tempo  $n + 1$  dado que estava no estado  $e_i$  no tempo  $n$ .

Os estados de uma cadeia de Markov, de acordo com a dinâmica gerada pelas probabilidades de transição, podem ser classificados como recorrentes, transientes, absorventes, dentre outros. Um estado  $e_i$  é dito absorvente se, uma vez que o processo chegar nele permanece para sempre, ou seja,  $p_{ii} = 1$ . Dizemos que um estado é recorrente se a probabilidade do processo retornar a ele é 1, do contrário, é dito transiente: existe probabilidade positiva do processo não retornar. Formalmente temos:

**Definição 4.20.** Dizemos que um estado  $e_j$  é recorrente se

$$\mathbb{P}(X_{n+m} = e_j \text{ para algum } m \geq 1 \mid X_n = e_j) = 1,$$

caso contrário é dito transiente.

Assim como temos várias formas de convergência para uma sequência de funções, o mesmo ocorre para uma sequência de variáveis aleatórias: como se trata de uma sequência de funções num ambiente probabilístico, surgem naturalmente vários tipos de convergência, como por exemplo, convergência quase sempre, convergência em probabilidade e a convergência em distribuição. Para o AGD foi possível provar a convergência quase sempre (ver Definição 4.21), que é mais “forte” que as outras duas. De fato, a convergência quase sempre implica em convergência em probabilidade (prova em BARRY (2008) p.195), e a convergência em probabilidade implica em convergência em distribuição (prova em BARRY (2008) p.248).

**Definição 4.21.** Seja  $(X_n)_{n \in \mathbb{N}}$  uma sequência de variáveis aleatórias. Dizemos que  $(X_n)_{n \in \mathbb{N}}$  converge quase sempre (quase certamente, fortemente ou com probabilidade 1) para a variável aleatória  $X$ , se

$$\mathbb{P}(\lim_{n \rightarrow \infty} X_n = X) = 1.$$

## 4.0.2 Prova da convergência

De maneira geral, o PDEPVZ pode ser modelado da seguinte forma:

$$\left\{ \begin{array}{l} \min_x F(x) \\ s.a. : \\ h(x) = 0 \\ x \in \mathfrak{X} \end{array} \right. \quad (4.0.2)$$

onde o conjunto factível  $\mathfrak{X} \subset \mathbb{R}^n$  é compacto (união de compactos), a função objetivo  $F$  é contínua no compacto  $\mathfrak{X}$  e a função  $h$  é contínua.

Para esta seção, a fim de mostrarmos a convergência do método AGD, consideraremos algumas hipóteses:

- ( $\mathcal{H}_1$ ) A população inicial é não vazia;
- ( $\mathcal{H}_2$ ) Para todo  $x$  factível, seja  $B(x, \delta)$  a bola aberta de centro  $x$  e raio  $\delta > 0$ . A medida de Lebesgue do conjunto  $B(x, \delta) \cap \mathfrak{X}$  é positiva, para todo  $\delta$ . Isso significa que se tomarmos qualquer vizinhança de um ponto factível, a interseção dela com a região factível possui volume (ou área) positivo;
- ( $\mathcal{H}_3$ ) Existe pelo menos um ótimo global que não é singularidade, denotado por  $x_{ns}^*$ , e pelo menos um ótimo global que é singularidade, denotado por  $x_s^*$ . Note que ambos possuem o mesmo custo, uma vez que são ótimos globais.

A região factível associada ao problema PDEPVZ satisfaz a condição ( $\mathcal{H}_2$ ). Em geral, esta é uma condição usada em métodos evolutivos em geral, como pode ser visto, por exemplo, em WANG ET AL. (2005), e a utilizaremos na prova do item (*iv*) da Proposição 4.22 abaixo. Por outro lado, a hipótese ( $\mathcal{H}_3$ ) nos permite apresentar as dinâmicas estocásticas mais gerais a serem realizadas pelo AGD. Casos específicos são discutidos na Observação 4.26.

Considere

$$\mathcal{S} = \{x \in \mathfrak{X} : x \text{ é singularidade}\}$$

e

$$Q_{ns} = \{x \in \mathfrak{R} \setminus \mathcal{S} : |F(x) - F(x_{ns}^*)| < \epsilon\},$$

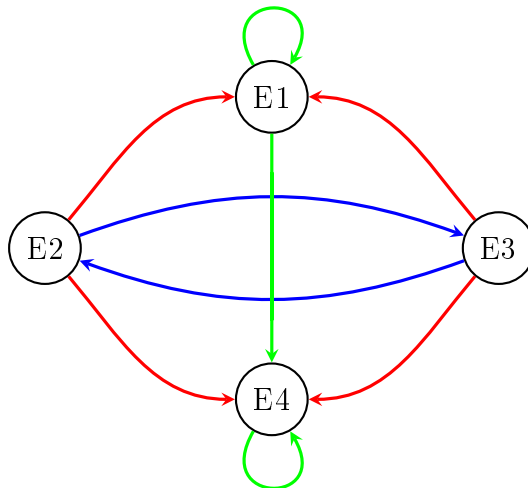
para um dado  $\epsilon > 0$ .

No que se segue, construiremos um processo estocástico discreto (mostrado na Figura 4.1) realizado pelo melhor indivíduo  $x_k^*$  da população  $POP_k$ , onde  $k = 1, 2, \dots$ . Para esse fim, assumimos o seguinte critério de desempate para a seleção do melhor indivíduo: se houver dois indivíduos, um em uma singularidade e o outro não, com o mesmo valor da função objetivo, o algoritmo deve escolher o indivíduo na singularidade.

Além disso, definimos os seguintes estados:

1. A população  $POP_k$  está no estado  $E_1$  se  $x_k^* \in Q_{ns}$ ;
2. A população  $POP_k$  está no estado  $E_2$  se  $x_k^* \notin \mathcal{S}$  e  $x_k^* \notin Q_{ns}$ ;
3. A população  $POP_k$  está no estado  $E_3$  se  $x_k^* \in \mathcal{S}$  e  $F(x_k^*) \neq F(x_s^*)$ ;
4. A população  $POP_k$  está no estado  $E_4$  se  $x_k^* \in \mathcal{S}$  e  $F(x_k^*) = F(x_s^*)$ .

Figura 4.1: Cadeia de Markov obtida para a população  $POP_k$ .



Fonte: Elaborada pelo autor.

Seja  $p_{ij}$  a probabilidade de transição da população  $POP_k$  do estado  $E_i$  para o estado  $E_j$ . Como a probabilidade  $p_{ij}$  não depende dos estados já visitados pela população anteriormente, temos uma Cadeia de Markov.

O resultado a seguir nos dará ferramentas para que possamos aplicar o Lema 4.7 e assim mostrar que teremos a convergência quase sempre da sequência gerada pelas imagens do melhor indivíduo de cada geração para o ótimo global do problema.

**Proposição 4.22.** (Probabilidades de transição) Nas condições assumidas anteriormente temos:

- (i)  $p_{12} = p_{13} = 0$ ;
- (ii)  $p_{44} = 1$ ;
- (iii)  $p_{23} > 0$ ,  $p_{24} > 0$  e  $p_{34} > 0$ ;
- (iv)  $p_{32} > 0$ ;
- (v)  $p_{21} > 0$  e  $p_{31} > 0$ .

*Demonstração.* De acordo com a Seção 3.1.3.7, como o melhor indivíduo de cada geração sempre sobrevive, temos que os itens (i) e (ii) são verdadeiros.

Para que  $p_{24} > 0$  e  $p_{34} > 0$  ocorram, basta notar que o indivíduo pivô  $I_P$  pertence a  $POP_k$  e portanto, existe a probabilidade de  $I_P$  sofrer mutação e se tornar exatamente  $x_s^*$ , fazendo com que a população esteja no estado  $E_4$ .

Como  $I_P$  pode sofrer mutação e se tornar o melhor indivíduo da população na próxima iteração, temos que  $p_{23} > 0$ . Temos ainda que  $p_{32} > 0$  pois, através da mutação, se o melhor indivíduo está em uma singularidade e que não seja ótimo global pode sair dela na próxima geração.

Por fim, provaremos o item (v) generalizando algumas técnicas propostas em WANG ET AL. (2005), que provou a convergência de um algoritmo evolutivo na resolução de um problema de programação não linear bi-nível.

Para que possamos mostrar que  $p_{21} > 0$ , temos que observar que a população  $POP_k$  pode sair do estado  $E_2$  e ir para o estado  $E_1$  através de uma mutação. Por hipótese, a

função  $F$  possui mínimo global em  $x_{ns}^*$  e este ponto não é uma singularidade. Como  $F$  é uma função contínua, temos que:

$$\forall \epsilon > 0, \exists \delta > 0, \text{ tal que } \|x - x_{ns}^*\| < \delta \Rightarrow |F(x) - F(x_{ns}^*)| < \epsilon/2.$$

Seja  $B(x_{ns}^*, \delta) = \{x \in \mathbb{R}^n; \|x - x_{ns}^*\| < \delta\}$  e  $V = (B(x_{ns}^*, \delta) \cap \mathfrak{A}) \setminus \mathcal{S}$ . Vale observar que a medida de Lebesgue de  $V$  continua sendo positiva, pois o conjunto das singularidades  $\mathcal{S}$  tem medida de Lebesgue nula. Além disso, temos que  $V \subset Q_{ns}$  pois, para todo  $x \in V$ , temos que  $x$  é factível, não é singularidade e  $\|x - x_{ns}^*\| < \delta$ . Pela continuidade, temos que  $|F(x) - F(x_{ns}^*)| < \epsilon/2$  e, portanto,  $x \in Q_{ns}$ .

Seja  $POP_k$  uma população no estado  $E_2$ ,  $x$  um indivíduo de  $POP_k$  e  $\psi$  o indivíduo obtido de  $x$  após a mutação descrita na Seção 3.1.3.5, ou seja,  $\psi = x + \xi$ , onde  $\xi = (\xi_1, \dots, \xi_n)$  com  $\xi_i$ ,  $i = 1, \dots, n$ , obtido aleatoriamente da distribuição normal de média 0 e desvio padrão  $\sigma_i$ , denotada por  $N(0, \sigma_i)$ . Como  $\xi_1, \dots, \xi_n$  são variáveis aleatórias independentes, se  $f_{\xi_1}, \dots, f_{\xi_n}$  são as respectivas funções densidade, pela Proposição 4.18 segue que a função

$$f_{\xi}(y_1, \dots, y_n) = \prod_{i=1}^n f_{\xi_i}(y_i), \quad (y_1, \dots, y_n) \in \mathbb{R}^n$$

é sua densidade conjunta (isto é,  $f_{\xi} = f_{\xi_1, \dots, \xi_n}$ ).

Como  $\xi_i \in N(0, \sigma_i)$ , segue que

$$f_{\xi_i}(y_i) = \frac{1}{\sqrt{2\pi} \sigma_i} \cdot e^{-(y_i-0)^2/2\sigma_i^2}, \quad y_i \in \mathbb{R}.$$

Sendo  $x = (x_1, \dots, x_n)$  determinístico, também podemos ver a aplicação  $\psi = x + \xi$  como uma variável aleatória, onde  $\psi_1 = x_1 + \xi_1, \dots, \psi_n = x_n + \xi_n$  são variáveis aleatórias independentes com funções densidade dadas por  $f_{x_1+\xi_1}, \dots, f_{x_n+\xi_n}$  e, portanto,

$$f_{\psi}(y) = f_{x+\xi}(y_1, \dots, y_n) = \prod_{i=1}^n f_{x_i+\xi_i}(y_i), \quad (y_1, \dots, y_n) \in \mathbb{R}^n.$$

Por outro lado, pela Proposição 4.19, segue que

$$f_{\psi_i}(y_i) = f_{x_i+\xi_i}(y_i) = f_{\xi_i}(y_i - x_i), \quad y_i \in \mathbb{R}.$$

Assim,

$$f_{\psi}(y) = f_{x+\xi}(y) = \prod_{i=1}^n f_{\xi_i}(y_i - x_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi} \sigma_i} \cdot e^{-(y_i-x_i)^2/2\sigma_i^2}.$$

Sem perda de generalidade, uma vez que  $\delta$  é suficientemente pequeno, podemos assumir que  $V \subset I_1 \times \dots \times I_n$ , onde cada  $I_i$  é um intervalo fechado e limitado da reta real. Pela independência de  $\psi_1, \dots, \psi_n$ , temos:

$$\begin{aligned}
 \mathbb{P}(\psi \in V) &\leq \mathbb{P}(\psi_1 \in I_1, \dots, \psi_n \in I_n) \\
 &= \prod_{i=1}^n \mathbb{P}(\psi_i \in I_i) \\
 &= \prod_{i=1}^n \int_{I_i} f_{\psi_i}(y_i) dy_i \\
 &= \prod_{i=1}^n \int_{I_i} f_{\xi_i}(y_i - x_i) dy_i.
 \end{aligned} \tag{4.0.3}$$

Notemos então que:

(a)  $0 < \mathbb{P}(\psi \in V) < 1$ ;

De fato, como

$$\mathbb{P}(\psi \in V) = \int_V f_\psi(y) dy,$$

pela Proposição 4.12, existe compacto  $K \subset V$  com medida de Lebesgue positiva e portanto

$$\mathbb{P}(\psi \in V) \geq \int_K f_\psi(y) dy \geq \int_K \min_K f_\psi(y) dy = \min f_\psi(y) \cdot \text{vol}(K) > 0.$$

De (4.0.3) temos

$$\mathbb{P}(\psi \in V) \leq \prod_{i=1}^n \int_{I_i} f_{\xi_i}(y_i - x_i) dy_i < \prod_{i=1}^n \int_{-\infty}^{\infty} f_{\xi_i}(y_i - x_i) dy_i = 1$$

(b) Considere a aplicação  $h : \mathfrak{R} \rightarrow [0, 1]$ , tal que  $h(x) = \mathbb{P}((x + \xi) \in V)$ . Então  $h$  é contínua em  $\mathfrak{R}$ . De fato, de (4.0.3), temos:

$$h(x_1, \dots, x_n) = \int_V \prod_{i=1}^n f_{\xi_i}(y_i - x_i) dy_i.$$

Como  $f_{\xi_i}(y_i - x_i) = \frac{1}{\sqrt{2\pi} \sigma_i} \cdot e^{-(y_i - x_i)^2 / 2\sigma_i^2}$  varia continuamente na variável  $x_i$  e a integral e o produto, não alteram a continuidade de uma função, temos que  $h$  é contínua no compacto  $\mathfrak{R}$  e, portanto, existe  $\tilde{x} \in \mathfrak{R}$ , tal que

$$h(\tilde{x}) = \min\{h(x); x \in \mathfrak{R}\}$$

onde pelo item (a)

$$0 < h(\tilde{x}) < 1.$$

Como  $V \subset Q_{ns}$ , temos:

$$0 < h(\tilde{x}) \leq h(x) \leq p_{21}.$$

A prova de que  $p_{31} > 0$  é obtida de forma análoga.  $\square$

**Observação 4.23.** O fato de termos  $p_{21} > 0$ ,  $p_{23} > 0$  e  $p_{24} > 0$  significa que  $p_{22} \neq 1$ . Então a probabilidade da população ficar no estado  $E_2$  para sempre é zero, pois para determinar esta probabilidade, basta multiplicar  $p_{22}$  por ele mesmo infinitamente. O mesmo ocorre com  $p_{33}$  então também não teremos uma população “presa” no estado  $E_3$ . Mostraremos a seguir que realmente esses casos não podem ocorrer e que também não poderemos ficar alternando entre os estados  $E_2$  e  $E_3$  para sempre.

**Teorema 4.24. (Convergência do AGD)** Considere a sequência  $(POP_k)$ , onde  $POP_k$  é a população obtida na  $k$ -ésima iteração do AGD. Seja  $x_k^*$  o melhor indivíduo da população  $POP_k$  e  $x^*$  o ótimo global do problema (singularidade ou não-singular). Então  $F(x_k^*)$  converge para  $F(x^*)$  quase sempre, ou seja,

$$\mathbb{P}\left(\lim_{k \rightarrow \infty} F(x_k^*) = F(x^*)\right) = 1.$$

*Demonstração.* A demonstração deste teorema será feita com a utilização do Lema de Borel-Cantelli. Para isso, considere  $p_k$  como sendo a probabilidade do evento  $A_k = \{\text{ficarmos “presos” nos estados } E_2 \text{ e } E_3\}$ , ou seja,

$$p_k = \mathbb{P}\{A_k\} := \mathbb{P}\{F(x_k^*) \neq F(x_s^*) \text{ e } |F(x_k^*) - F(x_{ns}^*)| \geq \epsilon\}.$$

Então segue que

$$p_k = \begin{cases} 0, & \text{se } POP_j \in E_1 \text{ para algum } 1 \leq j \leq k \\ 0, & \text{se } POP_j \in E_4 \text{ para algum } 1 \leq j \leq k \\ (p_{22})^{k_1} \cdot (p_{23})^{k_2} \cdot (p_{32})^{k_3} \cdot (p_{33})^{k_4}, & \text{caso contrário} \end{cases}$$

onde  $k_i$ ,  $i = 1, 2, 3, 4$ , representa a quantidade de vezes que a população  $POP_j$ ,  $1 \leq j \leq k$  permaneceu em cada um dos estados (logo  $k_1 + k_2 + k_3 + k_4 = k$ ).

Considerando  $\rho = \max\{p_{22}, p_{33}, p_{23}, p_{32}\}$ , pelo Teorema 4.22, como

$$p_{21} + p_{22} + p_{23} + p_{24} = 1$$

e

$$p_{31} + p_{32} + p_{33} + p_{34} = 1$$

temos que

$$p_{22} \neq 1, p_{23} \neq 1, p_{32} \neq 1, p_{33} \neq 1.$$

Portanto,

$$\sum_{k=1}^{\infty} p_k \leq \sum_{k=1}^{\infty} \rho^k = \frac{\rho}{1-\rho} < \infty,$$

uma vez que é uma série geométrica de razão  $\rho < 1$ .

Por Borel-Cantelli, temos que:

$$\mathbb{P} \left\{ \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \right\} = 0,$$

ou seja, a probabilidade do evento  $A_k$  ocorrer infinitas vezes é igual a 0. Portanto, a probabilidade de  $A_k$  ocorrer um número finito de vezes é igual a 1, que equivale a dizer que  $F(x_k^*)$  converge para  $F(x^*)$  quase sempre (isto é, para quase todo  $\omega$  temos  $|F(x_k^*)(\omega) - F(x^*)| < \epsilon$ , para  $k$  suficientemente grande e que depende de  $\omega$ ).  $\square$

**Corolário 4.25.** A dinâmica estocástica gerada pelo melhor indivíduo de cada população  $POP_k$ , juntamente com as probabilidades de transição  $p_{ij}$ , formam uma cadeia de Markov onde o estado  $E_4$  é absorvente (e portanto recorrente) e os estados  $E_1$ ,  $E_2$  e  $E_3$  são transientes.

*Demonstração.* Como  $p_{44} = 1$ , temos que o estado  $E_4$  é absorvente. Como mostramos a convergência do AGD, temos que em algum momento a população estará no estado  $E_1$  ou  $E_4$ . A partir deste momento, não sairá mais destes estados e portanto, não retornará ao estado  $E_2$  e nem ao estado  $E_3$ , fazendo desses, estados transientes. Com relação ao estado  $E_1$ , temos que a única forma de sair dele é indo para  $E_4$  e se isso ocorrer, como este é um estado absorvente, a população não retornará mais ao estado  $E_1$ .  $\square$

**Observação 4.26.** Se não existir  $x_s^*$ , o Teorema 4.24 continua válido, bastando apenas excluir o estado  $E_4$  e as probabilidades relacionadas a ele. Analogamente, se não existir

$x_{ns}^*$ , devemos excluir o estado  $E_1$  e ainda teremos a convergência. Mais que isso, se não distinguirmos as singularidades das não-singularidades, ficaremos apenas com os estados  $E_1$  e  $E_2$ . Neste caso, como  $0 < p_{21} < 1$  (Teorema 4.22), segue que  $p_{22} - p_{21}$  também pertence ao intervalo  $(0, 1)$ . E assim, continua válido o Teorema 4.24 para  $\rho = p_{22}$ .

Apresenta-se a seguir os testes numéricos realizados com o AGD na resolução do PDEPV e PDEPVZ.

## 4.1 Testes numéricos

Esta seção apresenta os resultados obtidos nos testes de desempenho do AGD na resolução do PDEPV e PDEPVZ. Para isso, utilizou-se um computador com processador Intel Core, Processador i5-8250u (1,8 GHz), 8 GB de RAM e sistema operacional Windows 10 e o algoritmo foi implementado em Python.

### 4.1.1 AGD na resolução do PDEPV

Nesta seção são apresentados os resultados obtidos na utilização do AGD para a resolução do PDEPV para sistemas com 3, 5, 6 e 40 geradores. Comparações são feitas com Algoritmo Genético convencional (AG), Evolução Diferencial de Otimização Colaborativa (EDOC), Estratégia de Evolução de Adaptação da Matriz de Covariância (EEAMC) e Otimização por Enxame de Partículas (PSO). As bibliotecas CMA e PySwarms foram utilizadas para a implementação de EEAMC e PSO, respectivamente. A configuração de parâmetros, como tamanho da população e número máximo de iterações, foi feita de acordo com os valores estipulados para o AGD.

#### 4.1.1.1 Sistema com 3 geradores

Para esta instância, uma demanda de 201.5 MW é considerada. Dados relacionados ao custo e capacidade de geração de cada unidade geradora foram extraídos de RAVI ET AL. (2006) e são apresentados na Tabela 4.1. Foram conduzidas 50 execuções, cada uma delas consistindo de 100 iterações. A população inicial é composta de 20 indivíduos

e ao desvio padrão da mutação foi atribuído o valor 1. Na Tabela 4.2 estão mostradas todas as singularidades do problema.

Tabela 4.1: Dados da curva de custo para o PDEPV com 3 geradores.

Unidade	a ( $\$/MW^2$ )	b ( $\$/MW$ )	c ( $\$$ )	e ( $\$$ )	f	$P_i^{min}$ (MW)	$P_i^{max}$ (MW)
1	0.00533	11.669	213.1	150	0.0630	50	200
2	0.00889	10.333	200	75	0.0315	37.5	150
3	0.00741	10.833	240	110	0.0420	45	180

Fonte: RAVI ET AL. (2006).

Tabela 4.2: Singularidades de cada um dos geradores para o PDEPV com 3 geradores.

Unidade 1 (MW)	Unidade 2 (MW)	Unidade 3 (MW)
50.0	37.5	45.0
99.9	137.2	119.8
149.7	150.0	180.0
199.6		
200.0		

Fonte: Elaborada pelo autor.

A Tabela 4.3 mostra a melhor solução obtida em cada método, seu custo, o custo médio das execuções, desvio padrão (DP) e o tempo gasto na execução que forneceu a melhor solução. O AGD, juntamente com o EEAMC e PSO, encontrou a melhor solução para o problema, com menor custo médio, desvio padrão e tempo computacional. Vale notar que a solução determinada pelo AGD se encontra em uma singularidade do problema, confirmando o que foi descrito na Seção 3.1.2 e justificando a ênfase dada às singularidades na busca pela solução do problema.

Tabela 4.3: Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 3 unidades geradoras.

	AG	EDOC	EEAMC	PSO	AGD
$P_1(MW)$	50.318	50.109	50.0	50.0	50.0*
$P_2(MW)$	38.567	37.678	37.5	37.5	37.5*
$P_3(MW)$	112.615	113.715	114.0	114.0	114.0
$Custo(\$/h)$	3017.64	3010.11	<b>3007.66</b>	<b>3007.66</b>	<b>3007.66</b>
Custo médio ( $\$/h$ )	3070.49	3032.78	3020.12	3120.97	<b>3007.66</b>
$DP$	27.76	10.07	9.48	69.97	$4.55 \cdot 10^{-13}$
tempo (s)	0.884	0.151	0.192	0.067	0.059

Fonte: Elaborada pelo autor. Os valores marcados com asterisco (\*) representam singularidades.

#### 4.1.1.2 Sistema com 5 geradores

Para esta instância do problema foi considerada uma demanda de 292.9 MW. A população criada pelo AGD possui 30 indivíduos e foram feitas 100 iterações em cada uma, das 50 execuções realizadas. Na Tabela 4.4 são apresentados os dados da curva de custo e os limites da capacidade de geração de cada gerador (extraídos de RAVI ET AL. (2006)) e na Tabela 4.5 estão todas as singularidades do problema.

Tabela 4.4: Dados da curva de custo para o PDEPV com 5 geradores.

Unidade	a ( $\$/MW^2$ )	b ( $\$/MW$ )	c ( $\$$ )	e ( $\$$ )	f	$P_i^{min}$ (MW)	$P_i^{max}$ (MW)
1	0.00375	2	0	80	0.073	10	100
2	0.0175	1.75	0	30	0.032	20	50
3	0.0625	1	0	50	0.051	15	80
4	0.00834	3.25	0	25	0.026	10	76
5	0.025	3	0	25	0.026	10	76

Fonte: RAVI ET AL. (2006).

Tabela 4.5: Singularidades de cada um dos geradores para o PDEPV com 5 geradores.

Unidade	1 (MW)	2 (MW)	3 (MW)	4 (MW)	5 (MW)
	10.0	20.0	15.0	10.0	10.0
	53.0	50.0	76.6	76.0	76.0
	96.1		80.0		
	100.0				

Fonte: Elaborada pelo autor.

A Tabela 4.6 mostra o despacho e o custo da melhor solução obtida pelo AGD e pelos métodos utilizados na comparação. Além disso, apresenta o custo médio, desvio padrão e tempo computacional de todos os métodos. Como se pode observar, o AGD foi capaz de obter uma solução com custo menor do que a solução apresentada pelos outros métodos utilizados na comparação. Além disso, o AGD apresentou o menor custo médio, desvio padrão e menor tempo computacional.

Tabela 4.6: Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 5 unidades geradoras.

	AG	EDOC	EEAMC	PSO	AGD
$P_1(MW)$	95.154	95.520	96.073	95.559	96.089*
$P_2(MW)$	49.111	37.816	49.998	43.245	50.0*
$P_3(MW)$	19.174	42.569	19.092	24.771	17.907
$P_4(MW)$	65.157	73.539	75.987	75.350	76.0*
$P_5(MW)$	63.305	43.455	51.749	53.975	49.004
$Custo(\$ / h)$	1019.01	1046.72	998.99	1018.62	<b>976.25</b>
Custo médio ( $\$/h$ )	1087.29	1076.09	1002.39	1153.06	<b>990.87</b>
$DP$	36.64	29.27	22.97	77.91	19.01
tempo (s)	0.819	0.269	0.346	0.132	0.100

Fonte: Elaborada pelo autor. Os valores marcados com asterisco (\*) representam singularidades.

#### 4.1.1.3 Sistema com 6 geradores

Nesta instância foi considerada uma demanda de 290.9 MW, a população criada possui 30 indivíduos, foram feitas 50 execuções, cada uma delas com 100 iterações. Na Tabela 4.7 são apresentados os dados da curva de custo e os limites de capacidade de geração de cada gerador (extraídos de RAVI ET AL. (2006)), enquanto que na Tabela 4.8 estão todas as singularidades associadas a cada unidade geradora.

Tabela 4.7: Dados da curva de custo para o PDEPV com 6 geradores.

Unidade	a ( $\$/MW^2$ )	b ( $\$/MW$ )	c ( $\$$ )	e ( $\$$ )	f	$P_i^{min}$ (MW)	$P_i^{max}$ (MW)
1	0.00375	2	0	120	0.073	50	200
2	0.0175	1.75	0	50	0.032	20	80
3	0.0625	1	0	30	0.051	15	50
4	0.00834	3.25	0	25	0.026	10	35
5	0.025	3	0	25	0.026	10	30
6	0.025	3	0	30	0.048	12	40

Fonte: RAVI ET AL. (2006).

A Tabela 4.9 mostra o despacho e o custo da melhor solução, custo médio, desvio padrão e tempo computacional obtidos pelo AGD e pelos métodos utilizados na comparação. Como se pode observar, a melhor solução obtida pelo AGD só não é melhor que a apresentada pelo EEAMC, porém a diferença entre os valores obtidos pelos dois métodos

Tabela 4.8: Singularidades associadas a cada um dos geradores para o PDEPV com 6 geradores.

Unidade	1 (MW)	2 (MW)	3 (MW)	4 (MW)	5 (MW)	6 (MW)
	50.0	20.0	15.0	10.0	10.0	12.0
	93.4	80.0	50.0	35.0	30.0	40.0
	136.1					
	179.1					
	200.0					

Fonte: Elaborada pelo autor.

é muito pequena (apenas 15 centavos) e o EEAMC teve um custo médio e desvio padrão altos. Além disso, o AGD mais uma vez apresentou o melhor custo médio, menor desvio padrão e um dos menores tempos computacionais, perdendo apenas para o PSO.

Tabela 4.9: Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 6 unidades geradoras.

	AG	EDOC	EEAMC	PSO	AGD
$P_1(MW)$	179.347	179.245	179.110	179.357	179.125*
$P_2(MW)$	45.651	29.247	37.537	20.515	38.779
$P_3(MW)$	18.165	23.403	15.429	15.716	15.0*
$P_4(MW)$	22.053	26.075	31.358	32.680	30.917
$P_5(MW)$	12.532	19.886	15.452	29.037	15.079
$P_6(MW)$	13.152	13.049	12.014	13.594	12.0*
$Custo(\$ / h)$	850.711	852.40	<b>845.30</b>	854.67	845.45
Custo médio ( $\$/h$ )	907.057	941.32	893.61	962.90	<b>849.01</b>
$DP$	26.218	36.88	16.33	59.39	1.93
tempo (s)	0.954	0.218	0.317	0.142	0.151

Fonte: Elaborada pelo autor. Os valores marcados com asterisco (\*) representam singularidades.

#### 4.1.1.4 Sistema com 40 geradores

Por fim, na instância do PDEPV com 40 unidades geradoras, uma demanda de 10500 MW é considerada, foram realizadas 50 execuções, cada uma delas composta por 200 iterações. Cada geração é composta por 100 indivíduos e ao desvio padrão da mutação foi atribuído o valor 1. Os dados relacionados ao custo e capacidade de geração de cada unidade geradora foram extraídos de HOTA & SAHU (2015).

A Tabela 4.10 mostra o despacho e o custo da melhor solução, custo médio, desvio padrão e tempo obtidos pelo AGD e pelos métodos utilizados na comparação. Como se pode observar, a melhor solução obtida pelo AGD é melhor que a apresentada por todos os outros. Além disso, o AGD apresentou o melhor custo médio e menor desvio padrão.

Tabela 4.10: Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 40 unidades geradoras.

	AG	EDOC	EEAMC	PSO	AGD
$P_1(MW)$	87.824	81.478	92.916	88.938	107.953
$P_2(MW)$	111.510	138.732	76.808	68.551	112.033
$P_3(MW)$	81.295	103.287	97.989	118.564	116.267
$P_4(MW)$	146.425	153.686	131.826	128.961	181.434
$P_5(MW)$	51.516	14.761	89.506	86.477	95.699
$P_6(MW)$	120.232	125.726	132.290	135.394	139.674
$P_7(MW)$	255.585	308.615	261.894	135.734	299.669*
$P_8(MW)$	148.032	273.885	290.973	273.469	284.546*
$P_9(MW)$	266.477	201.375	296.920	213.477	287.934
$P_{10}(MW)$	163.586	222.056	206.121	225.561	201.967
$P_{11}(MW)$	319.351	401.569	311.976	179.687	95.383
$P_{12}(MW)$	343.402	377.574	209.223	202.022	167.687
$P_{13}(MW)$	363.683	273.582	369.061	425.518	126.343
$P_{14}(MW)$	432.738	336.676	392.237	472.839	392.205
$P_{15}(MW)$	492.753	367.766	391.292	482.653	393.087
$P_{16}(MW)$	414.380	486.302	340.540	405.214	304.287*
$P_{17}(MW)$	487.977	489.819	400.045	487.469	490.808
$P_{18}(MW)$	288.302	444.645	481.207	402.222	489.101*
$P_{19}(MW)$	458.554	405.989	422.571	478.863	513.818
$P_{20}(MW)$	483.271	448.216	423.163	543.882	516.514
$P_{21}(MW)$	453.548	386.140	521.818	543.778	528.629
$P_{22}(MW)$	474.830	515.195	513.791	526.334	523.636*
$P_{23}(MW)$	405.880	474.739	532.547	462.235	526.518
$P_{24}(MW)$	270.547	433.182	531.922	542.341	535.597

$P_{25}(MW)$	333.520	529.296	532.843	429.738	525.219
$P_{26}(MW)$	338.775	488.696	522.199	435.745	533.308
$P_{27}(MW)$	90.582	107.419	17.535	47.617	10.0*
$P_{28}(MW)$	70.693	71.329	11.782	45.742	11.164
$P_{29}(MW)$	131.450	124.823	14.020	54.828	10.517
$P_{30}(MW)$	85.919	91.617	90.353	91.258	87.525*
$P_{31}(MW)$	97.989	138.423	178.027	176.631	189.529
$P_{32}(MW)$	130.940	150.563	169.805	180.576	190.0*
$P_{33}(MW)$	121.656	20.956	172.911	182.998	162.378
$P_{34}(MW)$	144.887	187.284	158.521	132.150	163.130
$P_{35}(MW)$	129.516	206.643	174.109	165.164	164.735*
$P_{36}(MW)$	159.795	98.768	168.988	184.921	199.010
$P_{37}(MW)$	90.031	79.085	97.727	106.350	110.0*
$P_{38}(MW)$	102.178	60.585	78.780	51.891	109.699
$P_{39}(MW)$	38.507	94.632	88.318	108.924	91.071
$P_{40}(MW)$	1311.592	613.581	505.412	475.278	511.920*
$Custo(\$ / h)$	133044.24	146029.25	125623.73	131356.44	<b>122188.72</b>
Custo médio ( $\$/h$ )	137631.61	150092.94	126743.78	140602.34	<b>122802.62</b>
$DP$	1787.54	10237.58	640.40	4933.23	418.58
tempo (s)	8.938	1.862	6.250	0.878	6.137

#### 4.1.2 AGD na resolução do PDEPVZ

Nesta seção são apresentados os resultados obtidos na utilização do AGD para a resolução do PDEPVZ para sistemas com 3, 6 e 40 geradores. Os dados dos sistemas foram extraídos de RAVI ET AL. (2006) e HOTA & SAHU (2015) (mesmos utilizados na seção anterior), enquanto que os dados referentes às zonas de operação proibida foram obtidos de ARUL ET AL. (2013).

Para fins de comparação, os mesmos métodos da seção anterior foram utilizados. Além disso, o Algoritmo Genético Direcionado sem pivô (AGDsp) foi implementado para avaliar a importância desse operador. Todos os parâmetros do AGD foram ajustados pelo

*Optuna*, uma framework que automatiza o processo de encontrar as combinações ótimas de parâmetros por meio de otimização Bayesiana, guiada pelos melhores resultados obtidos até o momento (para mais detalhes, veja BERGSTRA ET AL. (2011)).

#### 4.1.2.1 Sistema com 3 geradores

Para este caso, uma demanda de 205 MW é considerada. Dados relacionados ao custo e capacidade de geração de cada unidade geradora foram extraídos de RAVI ET AL. (2006), e as respectivas zonas de operação proibida foram obtidas em ARUL ET AL. (2013). Um total de 50 execuções foram conduzidas, cada uma delas consistindo de 161 iterações. A população inicial é composta de 43 indivíduos, a probabilidade de mutação foi de 0.1 e ao desvio padrão da mutação foi atribuído o valor 1.636.

A Tabela 4.11 mostra as zonas de operação proibida de cada unidade geradora e a Tabela 4.12 traz as singularidades do problema. Note que os limitantes das zonas de operação proibida foram incluídos entre as singularidades e as singularidades que pertenciam a alguma zona proibida foram desconsideradas.

Tabela 4.11: Zonas de operação proibida para o PDEPVZ com 3 geradores.

Unidade	Zonas de operação proibida
1	(165, 177)
2	(50, 60) (92, 102)
3	(60, 72)

Fonte: ARUL ET AL. (2013).

Tabela 4.12: Singularidades de cada um dos geradores para o PDEPVZ com 3 geradores.

Unidade	1 (MW)	2 (MW)	3 (MW)
	50.0	37.5	45.0
	99.9	50.0	60.0
	149.7	60.0	72.0
	165.0	92.0	119.8
	177.0	102.0	180.0
	199.6	137.2	
	200.0	150.0	

Fonte: Elaborada pelo autor.

A Tabela 4.13 apresenta os resultados obtidos pelos métodos AGD, AG, EDOC, EE-AMC e PSO. Ela mostra a melhor solução encontrada, o custo em  $\$/h$ , os valores médios para os custos, os desvios padrão e o tempo de execução para esta solução. Observa-se

que o AGD e o PSO obtiveram um custo melhor do que os outros métodos, no entanto, o AGD mostrou-se mais eficaz, uma vez que o desvio padrão é extremamente pequeno. Entre os métodos analisados, o EEAMC apresentou o pior desempenho, manifestando o maior valor da função de custo. Embora o AG tenha demonstrado o terceiro melhor valor de custo, foi observado um desvio padrão significativo, indicando que soluções de alta qualidade não são consistentemente alcançadas com este método. Em relação aos tempos computacionais, todos os métodos tiveram um desempenho parecido. Vale ressaltar que a solução encontrada pelo AGD é uma singularidade do problema.

Tabela 4.13: Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 3 unidades geradoras.

	AG	EDOC	EEAMC	PSO	AGD
$P_1(MW)$	50.13	99.87	99.90	50.00	50.00*
$P_2(MW)$	100.88	49.99	49.98	37.50	37.50*
$P_3(MW)$	53.81	55.13	55.12	117.50	117.50
$Custo(\$ / h)$	3098.02	3104.45	3104.71	<b>3035.65</b>	<b>3035.65</b>
Custo médio ( $\$/h$ )	3228.36	3104.46	3107.99	3200.02	3036.01
$DP$	55.86	0.0058	3.39	36.28	2.47
tempo (s)	0.030	0.006	0.098	0.457	0.076

Fonte: Elaborada pelo autor. Os valores marcados com asterisco (\*) representam singularidades.

#### 4.1.2.2 Sistema com 6 geradores

Para este caso, uma demanda de 283.4 MW é considerada. Como no caso anterior, os dados relacionados ao custo e capacidade de geração para cada gerador foram extraídos de RAVI ET AL. (2006), e os dados referentes às zonas de operação proibida foram obtidos de ARUL ET AL. (2013). Um total de 50 execuções foram feitas, cada uma delas com 200 iterações. A população inicial é composta de 100 indivíduos, probabilidade de mutação de 0.1 e o desvio padrão utilizado é igual a 0.329. A Tabela 4.14 mostra as zonas de operação proibida de cada unidade geradora e a Tabela 4.15 traz todas as singularidades do problema.

Dos resultados apresentados na Tabela 4.16, observamos que o AGD obteve uma solução melhor do que os outros métodos, com um desvio padrão muito pequeno. Vale

Tabela 4.14: Zonas de operação proibida para o PDEPVZ com 6 geradores.

Unidade	Zonas de operação proibida
1 (MW)	(55, 66) (80, 120)
2 (MW)	(21, 24) (50, 60)
3 (MW)	(30, 36)
4 (MW)	(25, 30)
5 (MW)	(25, 28)
6 (MW)	(24, 30)

Fonte: ARUL ET AL. (2013).

Tabela 4.15: Singularidades de cada um dos geradores para o PDEPVZ com 6 geradores.

Unidade	1 (MW)	2 (MW)	3 (MW)	4 (MW)	5 (MW)	6 (MW)
	50.0	20.0	15.0	10.0	10.0	12.0
	55.0	21.0	30.0	25.0	25.0	24.0
	66.0	24.0	36.0	30.0	28.0	30.0
	80.0	50.0	50.0	35.0	30.0	40.0
	120.0	60.0				
	136.1	80.0				
	179.1					
	200.0					

Fonte: Elaborada pelo autor.

ressaltar que isso ocorreu mesmo que a solução encontrada não seja uma singularidade do problema. Entre os métodos analisados, o EDOC apresentou o pior desempenho, com o maior valor da função de custo. Neste cenário, o AG demonstrou o segundo melhor valor de custo, embora com um desvio padrão relativamente alto. Em termos de tempo computacional, todos os métodos obtiveram um desempenho semelhante, com exceção do AGD que apresentou um tempo computacional bem pequeno, se comparado com os demais.

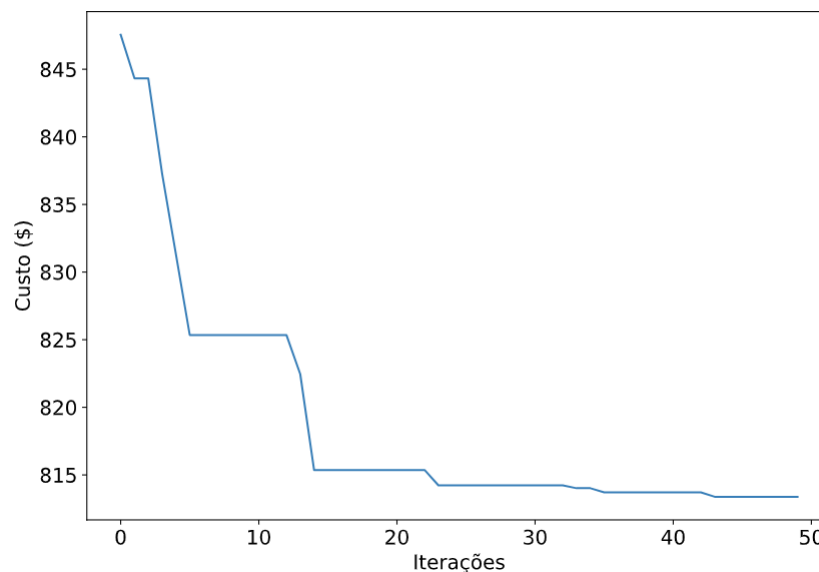
Tabela 4.16: Comparação entre custo mínimo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 6 unidades geradoras.

	AG	EDOC	EEAMC	PSO	AGD
$P_1(MW)$	135.55	79.99	53.40	190.83	179.12*
$P_2(MW)$	67.16	79.98	80.00	20.12	32.66
$P_3(MW)$	24.15	49.70	50.00	21.84	15.25
$P_4(MW)$	17.61	24.95	34.99	12.40	31.19
$P_5(MW)$	19.99	24.79	25.00	22.27	12.97
$P_6(MW)$	18.94	23.97	39.99	15.93	12.20
$Custo(\$/h)$	881.91	1111.64	1109.48	903.23	<b>812.59</b>
Custo médio ( $\$/h$ )	955.23	1111.79	1109.98	1088.92	816.56
$DP$	31.14	2.46	7.11	53.41	1.83
tempo (s)	5.723	9.082	7.319	2.561	0.136

Fonte: Elaborada pelo autor. Os valores marcados com asterisco (\*) representam singularidades.

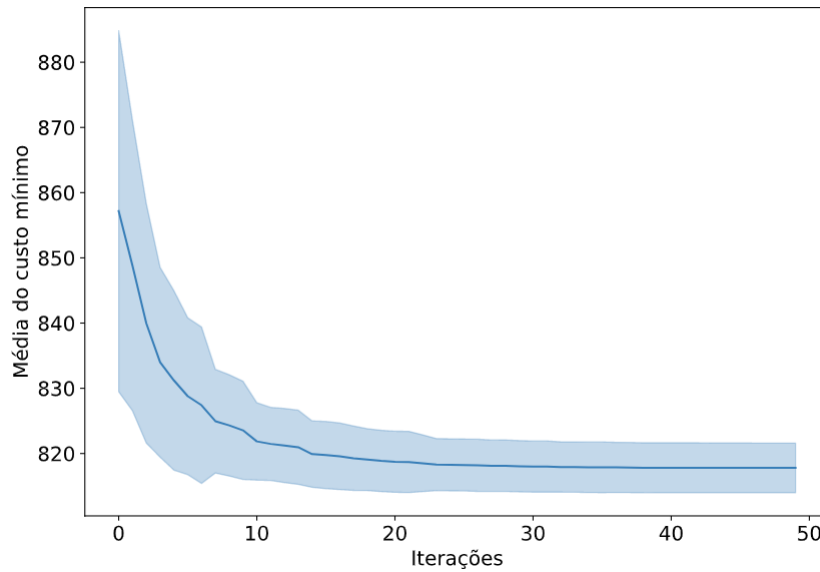
A Figura 4.2 apresenta a curva de convergência do método AGD na realização que obteve o menor custo. Nota-se que a convergência ocorre antes da iteração de número 30. A Figura 4.3 mostra a média das curvas de convergência para o AGD, juntamente com o desvio padrão em cada iteração.

Figura 4.2: Convergência do método AGD no PDEPVZ com 6 geradores.



Fonte: Elaborada pelo autor.

Figura 4.3: Média e desvio padrão das curvas de convergência para o AGD no PDEPVZ com 6 geradores.



Fonte: Elaborada pelo autor.

#### 4.1.2.3 Sistema com 40 geradores

Para esta instância, além dos métodos já utilizados nas comparações, foram implementados: o Evolução Diferencial Adaptativa Caótica com Arquivo Externo Opcional (CJADE) (GAO ET AL., 2019), com o parâmetro de porcentagem superior  $p = 0.05$  e o fator relacionado ao mecanismo de adaptação  $d = 0.1$ , e o Estratégia de Evolução com Amostragem Híbrida (HSES), com taxa de consideração da memória harmônica  $hmcr = 0.9$ , taxa de ajuste de tom  $par = 0.3$  e largura de banda  $bw = 0.01$ . Considerou-se uma demanda de 10500 MW. Os dados relativos ao custo, à capacidade de geração para cada unidade geradora e às zonas de operação proibidas foram obtidos de HOTA & SAHU (2015). Neste caso, foram realizadas 200 execuções, cada uma consistindo em 2500 iterações. A população inicial é composta por 300 indivíduos, com probabilidade de mutação de 0,1, e o desvio padrão utilizado foi 0.537.

A partir da Tabela 4.17, observamos que, entre os métodos analisados, mais uma vez o AGD alcançou resultados superiores com o menor desvio padrão e o AG apresentou o pior desempenho.

Tabela 4.17: Comparação entre custo mínimo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 40 unidades geradoras.

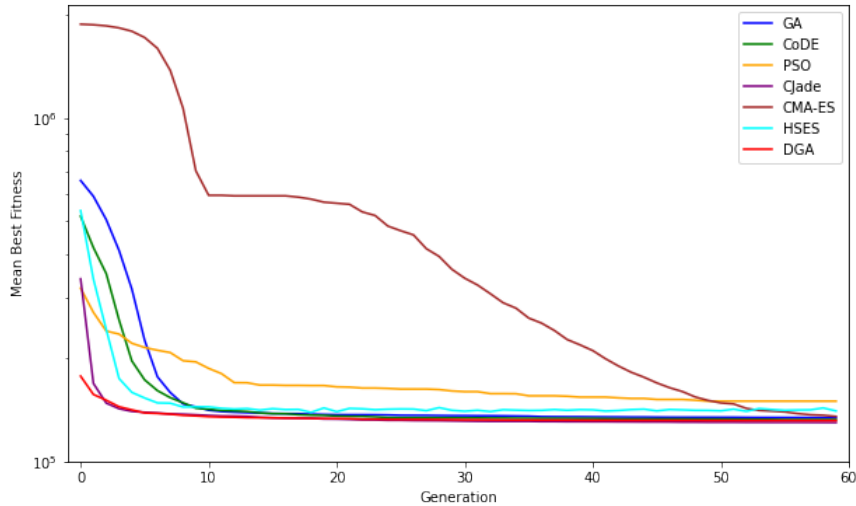
	CJADE	HSES	AG	EDOC	EEAMC	PSO	AGD
--	-------	------	----	------	-------	-----	-----

$P_1(MW)$	112.28	101.90	94.71	73.45	110.82	89.65	112.50
$P_2(MW)$	111.22	85.34	65.59	113.62	110.87	113.66	110.80*
$P_3(MW)$	113.25	103.82	87.70	109.15	120.00	120.00	102.93
$P_4(MW)$	188.12	169.93	173.86	146.10	179.73	129.74	179.73*
$P_5(MW)$	95.91	52.58	85.37	73.46	88.06	96.88	97.00*
$P_6(MW)$	139.99	126.04	115.38	130.45	140.00	105.21	140.00*
$P_7(MW)$	123.92	255.11	300.00	286.39	300.00	259.69	259.60*
$P_8(MW)$	292.38	277.84	277.17	267.00	284.60	300.00	284.60*
$P_9(MW)$	299.66	247.25	238.75	296.60	284.60	284.52	284.60*
$P_{10}(MW)$	230.10	257.97	238.98	129.13	130.00	130.00	130.00*
$P_{11}(MW)$	168.80	286.48	372.98	100.76	94.00	94.00	171.05
$P_{12}(MW)$	168.76	289.37	247.22	94.14	94.00	94.00	94.45*
$P_{13}(MW)$	125.51	360.06	407.12	140.49	125.00	125.00	214.76*
$P_{14}(MW)$	492.11	399.49	488.15	199.71	200.00	394.28	394.28*
$P_{15}(MW)$	304.54	383.30	410.52	502.87	484.04	394.28	395.41
$P_{16}(MW)$	394.28	461.62	497.97	456.71	484.04	484.04	489.28
$P_{17}(MW)$	494.33	402.01	274.56	498.00	489.28	488.60	489.28*
$P_{18}(MW)$	497.07	396.74	440.51	486.38	489.28	489.01	511.28
$P_{19}(MW)$	532.60	480.07	362.81	545.31	511.28	540.22	511.28*
$P_{20}(MW)$	540.21	525.64	416.26	545.86	511.28	550.00	523.28
$P_{21}(MW)$	546.72	478.64	473.59	540.61	523.28	543.75	523.28*
$P_{22}(MW)$	523.25	528.75	453.62	493.65	523.28	550.00	523.28*
$P_{23}(MW)$	523.32	511.83	371.63	538.91	523.28	550.00	523.28*
$P_{24}(MW)$	523.33	500.79	481.37	475.34	523.28	550.00	523.28*
$P_{25}(MW)$	523.27	488.47	466.94	511.68	523.28	537.75	523.28*
$P_{26}(MW)$	523.32	442.19	519.54	541.23	523.28	550.00	523.28*
$P_{27}(MW)$	10.07	79.94	63.46	62.63	10.00	10.00	523.28
$P_{28}(MW)$	10.01	71.29	62.78	80.79	10.00	10.00	10.00*
$P_{29}(MW)$	10.02	89.89	31.19	53.15	10.00	10.00	10.00*
$P_{30}(MW)$	93.79	81.65	78.57	81.92	88.17	93.75	87.80*
$P_{31}(MW)$	189.98	147.94	174.46	144.47	190.00	159.84	190.00*
$P_{32}(MW)$	159.81	98.32	164.07	190.25	190.00	188.94	164.80
$P_{33}(MW)$	189.99	139.56	127.97	189.28	190.00	189.91	164.80
$P_{34}(MW)$	176.23	180.38	198.56	170.60	200.00	90.00	164.80*
$P_{35}(MW)$	169.84	137.23	180.62	187.26	200.00	170.82	110.00
$P_{36}(MW)$	93.31	165.05	143.09	187.45	200.00	200.00	110.00
$P_{37}(MW)$	96.04	49.19	68.29	111.92	110.00	101.81	110.00*
$P_{38}(MW)$	91.72	84.78	63.92	111.45	110.00	89.16	110.00*
$P_{39}(MW)$	109.63	67.37	99.03	100.12	110.00	110.00	110.00*
$P_{40}(MW)$	511.29	494.18	499.06	532.87	511.28	511.28	511.28*
Custo (\$/h)	124056.90	138238.47	133029.95	131374.71	122026.93	123750.53	<b>121547.91</b>
Custo médio (\$/h)	125903.37	141614.39	145736.97	139107.42	122152.37	125315.40	122011.53
DP	813.06	1869.22	5937.43	4088.88	1347.84	686.43	197.76
tempo (s)	2520.678	2377.639	260.134	135.327	263.736	23.395	30.795

A Figura 4.4 ilustra o comportamento médio de convergência de todos os métodos ao

longo de 60 gerações e 20 execuções. O gráfico destaca a eficiência do AGD em progredir consistentemente em direção à solução ótima, demonstrando um desempenho competitivo em comparação com outros algoritmos.

Figura 4.4: Comparação do *fitness* médio do AGD e de todos os métodos usados nas comparações.



Fonte: Elaborada pelo autor.

### 4.1.3 Influência do indivíduo pivô

A Tabela 4.18 fornece uma comparação entre os custos mínimos obtidos pelos métodos AG, AGDsp e AGD na resolução do PDEPVZ em sistemas com 3, 6 e 40 geradores. Esses resultados mostram que o AGD apresenta melhores resultados do que ambos, em todas as instâncias consideradas, justificando o direcionamento da busca e comprovando a importância do indivíduo pivô.

Tabela 4.18: Comparação de custos entre AG, AGDsp e AGD.

	3 geradores	6 geradores	40 geradores
AG	3098.02	881.91	133029.95
AGDsp	<b>3035.65</b>	812.91	122498.71
AGD	<b>3035.65</b>	<b>812.59</b>	<b>121547.91</b>

Fonte: Elaborada pelo autor.

Após mostrar a convergência do AGD e realizar vários testes, surgem questões bastante relevantes: qual o tempo necessário para que esse processo alcance uma solução suficientemente próxima do ótimo global? Existe uma forma de ajustar o desvio padrão

associado à mutação, de tal forma que o método tenha uma taxa de convergência linear? Diante dessas incertezas, decidimos avançar em nossa pesquisa, explorando a possibilidade de analisar a taxa de convergência, embora este método seja intrinsecamente aleatório. A meta nos próximos capítulos é encontrar uma estratégia que permita responder a essas questões de forma mais precisa e eficaz, não só para o AGD, mas para algoritmos evolutivos que possuam determinadas características.

## Capítulo 5

# Taxa Média de Convergência - TMC

Em problemas de otimização discreta, uma das formas de medir a velocidade de convergência de um algoritmo é através do número de iterações ou tempo gasto para obter uma solução suficientemente próxima do ótimo global. Para métodos determinísticos, existe uma taxa de convergência, capaz de medir quão rapidamente as soluções encontradas se aproximam da solução ótima. Já para métodos aleatórios, como os algoritmos evolutivos, não é possível utilizar tal ferramenta devido ao caráter aleatório do método.

A fim de avaliar o desempenho de métodos evolutivos, este capítulo apresenta o conceito de TMC apresentado em HE & LIN (2015) e CHEN & HE (2021). Inicialmente apresenta-se o Algoritmo Evolutivo Adaptativo (AEA), outro método proposto neste trabalho, e o embasamento teórico, com as principais definições e resultados, necessários para o bom entendimento dos resultados apresentados. Em seguida, mostra-se que o AEA possui uma convergência linear e determina-se um limitante inferior para sua TMC. Por fim, aplica-se o AEA na minimização da função Ackley, Alpine,  $f(x, y) = x^4 + y^4$  e no PDEPV.

### 5.1 Algoritmo Evolutivo Adaptativo - AEA

Os algoritmos evolutivos (AE), como por exemplo o Algoritmo Genético (AG) e o Estratégia Evolutiva (EE), são métodos que se baseiam na Teoria da Evolução e cujo conceito envolve a geração de uma população inicial, na qual cada indivíduo é uma potencial solução para o problema em questão, e cada um de seus genes corresponde ao valor

que uma das variáveis do problema deve assumir. Posteriormente, ocorre um processo de seleção no qual os pais são escolhidos para se reproduzir e gerar descendentes iterativamente a partir da população anterior, sempre com o objetivo de melhorar os indivíduos e transmitir as melhores características. A mutação, com probabilidade  $p_M$ , é incorporada ao método por meio de uma leve perturbação em um ou mais genes do indivíduo a ser mutado.

Buscando melhorar o desempenho dos AEs, nesta seção, propomos o AEA, um método baseado no AE que incorpora o elitismo e uma mutação adaptativa. O AEA ajusta dinamicamente o desvio padrão da mutação ao longo do processo de busca, visando alcançar uma TMC linear.

O pseudocódigo do AEA é apresentado no Algoritmo 2, com cada operador detalhado nesta seção.

---

**Algoritmo 2** Algoritmo Evolutivo Adaptativo (AEA)

---

- 1: **Input:** Probabilidade de mutação  $p_M$ , número de indivíduos  $n_I$ , número de gerações  $n_g$
  - 2:  $k \leftarrow 0$
  - 3: Gere a população inicial aleatoriamente com  $n_I$  indivíduos
  - 4: **while**  $k < n_G$  **do**
  - 5:   Avalie indivíduos da população
  - 6:   Selecione pais
  - 7:   Gere uma nova população
  - 8:   Aplique mutação adaptativa com probabilidade  $p_M$
  - 9:   Aplique ajuste individual e elitismo
  - 10:    $k \leftarrow k + 1$
  - 11: **end while**
  - 12: Retorne o melhor indivíduo da população como solução
- 

A seguir, fornecemos uma descrição detalhada de cada operador.

### 5.1.1 População Inicial

A população inicial de tamanho  $n_I \geq 1$  é aleatoriamente gerada dentro da região factível.

### 5.1.2 Função avaliação

Cada indivíduo deve ser avaliado usando uma função de aptidão (*fitness*) que mede o grau de adaptabilidade do indivíduo. Esta função deve levar em consideração tanto o valor da função objetivo quanto o grau de viabilidade do indivíduo em relação à aplicação das restrições do problema de otimização. No AEA, quando um indivíduo infactível é encontrado, ele deve ser ajustado da mesma forma que o AGD, tornando-se factível. Portanto, a aptidão é determinada apenas com base na função objetivo.

### 5.1.3 Mutação

Cada indivíduo tem a probabilidade de sofrer mutação, feita através de uma perturbação aleatória nos genes do indivíduo. A probabilidade de mutação  $p_M$  em cada indivíduo é determinada pelo coeficiente de mutação, uma taxa que decide se a mutação deve ocorrer. Portanto, é crucial definir o coeficiente de mutação para um valor pequeno para preservar as características genéticas favoráveis adquiridas até aquele ponto no algoritmo. No entanto, a mutação facilita o surgimento de indivíduos com diversidade genética a partir de seus pais, promovendo uma maior variabilidade no algoritmo. Diante disso, o ideal é que a diversidade deve estar presente no início do processo, e à medida que as gerações progridem, a busca deve se concentrar mais profundamente nas melhores soluções identificadas até o momento.

Cada indivíduo  $x = (x_1, \dots, x_n)$  pode sofrer uma mutação dada por

$$x + \xi = (x_1 + \xi_1, \dots, x_n + \xi_n), \quad (5.1.1)$$

onde  $\xi_i \in \mathcal{N}(0, \sigma_i)$  para  $i = 1, \dots, n$ . Aqui,  $\mathcal{N}(0, \sigma_i)$  denota a distribuição normal com média 0 e desvio padrão  $\sigma_i$ . No AEA,  $\sigma_i$  é adaptativo, o que significa que ele assume diferentes valores para cada  $x_i$ . Seu valor em cada geração  $k$  é dado pelo Teorema 5.3 (ou

Teorema 5.5) ou Corolário 5.4, ver Seção 5.3.

Depois desse processo, se algum indivíduo for inactível, ele deve ser ajustado adequadamente. Como a TMC é analisada depois de um tempo suficientemente grande, é esperado que  $\sigma_i$  seja suficientemente pequeno e que o indivíduo ótimo da população esteja suficientemente próximo do ótimo do problema (e portanto factível). Este ajuste, quando necessário, não interferirá na linearidade da TMC do algoritmo. Além disso, se considerarmos uma taxa de mutação, isso alterará ligeiramente a estimativa do limitante inferior da TMC, mas ainda permanecerá linear. Na Seção 5.4, detalhamos os ajustes feitos em cada exemplo.

#### 5.1.4 Seleção

A seleção de indivíduos para o processo de evolução é realizada por meio de um torneio: três indivíduos são escolhidos aleatoriamente da população, e o melhor entre eles é selecionado.

#### 5.1.5 Elitismo

O operador de elitismo garante que o melhor indivíduo da população sempre sobreviva para a próxima geração. Portanto, o elitismo é um operador que impede que a aptidão da próxima geração seja pior do que a aptidão da geração anterior. No AEA, o elitismo é empregado porque, entre o indivíduo selecionado e a prole gerada, apenas os indivíduos com a melhor avaliação sobrevivem.

#### 5.1.6 Critério de parada

Os critérios de parada empregados incluem o número máximo de gerações  $n_g$  e estagnação. Isso significa que o algoritmo pára quando um número pré-definido de gerações é alcançado ou quando não há melhoria na aptidão da melhor solução por um determinado número de gerações consecutivas.

## 5.2 Noções preliminares

Nesta seção, apresentamos notações e resultados cruciais para o desenvolvimento deste trabalho. Um conceito central é a TMC, denotado por  $R_t$ , definido por HE & LIN (2015) e também utilizado por CHEN & HE (2021). Esta métrica mede a velocidade do algoritmo em alcançar a solução ótima (ou uma solução suficientemente próxima do ótimo global) através de uma expressão do erro de aproximação.

Considere um problema de minimização definido da seguinte forma:

$$\min_{x \in \mathfrak{A}} f(x), \quad (5.2.1)$$

onde

$$f : \mathfrak{A} \subset \mathbb{R}^n \rightarrow \mathbb{R}$$

é uma função contínua definida em um conjunto compacto  $\mathfrak{A}$ . Note que o problema (5.2.1) engloba problemas de otimização de propósito geral, que podem envolver restrições de igualdade e desigualdade, bem como restrições limitando as variáveis. Um indivíduo  $x \in \mathfrak{A}$  é um único ponto, e uma população  $X$  no algoritmo é uma coleção de indivíduos. Denotamos por  $x^*$  a solução ótima do problema. Se  $X_t$ ,  $t = 0, 1, \dots$ , representa uma população na  $t$ -ésima iteração, o erro de aproximação é dado por

$$e(X_t) = |f(X_t) - f(x^*)|,$$

onde

$$f(X_t) = \min\{f(x); x \in X_t\}$$

é o valor de aptidão da população  $X_t$ .

Dessa forma,  $\{e(X_t); t = 0, 1, \dots\}$  pode ser considerado um processo estocástico definido em um espaço de probabilidade  $(\Omega, \mathcal{F}, \mathbb{P})$ . Observa-se que, como o algoritmo utiliza o elitismo, temos que  $e(X_{t+1}) \leq e(X_t)$  quase sempre (q.s.) para qualquer  $t = 0, 1, \dots$ . Denotando a esperança da função  $f(X_t)$  como

$$f_t = \mathbb{E}[f(X_t)],$$

o erro aproximado do algoritmo é dado por

$$e_t = |f_t - f(x^*)|.$$

Dizemos que a sequência  $e(X_0), e(X_1), \dots$  converge em média para zero se  $e_t = \mathbb{E}[e(X_t)]$  tende a zero conforme  $t$  tende ao infinito. Como o objetivo principal é analisar a TMC, é assumido ao longo do trabalho que  $\lim_{t \rightarrow \infty} e_t = 0$ .

Conforme mencionado em HE & LIN (2015), a taxa de variação do erro dada por  $e_t/e_{t-1}$ , conforme utilizado em RUDOLPH (1997a), é altamente sensível a testes computacionais, tornando sua análise desafiadora. Portanto, a TMC é definida calculando-se a média geométrica da taxa de redução do erro ao longo de  $t$  gerações consecutivas:

$$\begin{aligned} R_t &= 1 - \left( \left| \frac{f_1 - f(x^*)}{f_0 - f(x^*)} \right| \cdots \left| \frac{f_t - f(x^*)}{f_{t-1} - f(x^*)} \right| \right)^{1/t} \\ &= 1 - \left( \left| \frac{f_t - f(x^*)}{f_0 - f(x^*)} \right| \right)^{1/t} \\ &= 1 - \left( \frac{e_t}{e_0} \right)^{1/t}. \end{aligned} \tag{5.2.2}$$

É importante observar que, por conta do elitismo no AEA, temos que  $R_t \in [0, 1]$ .

Além disso, como uma vantagem adicional, em HE & LIN (2015), os autores apresentam a Taxa Média de Convergência Alternativa (TMCA), denotada por  $R_t^\dagger$ , como uma forma de medir a TMC para casos em que o ótimo global é desconhecido. A TMCA é definida como:

$$R_t^\dagger := 1 - \left| \frac{f_{t+\tau} - f_t}{f_t - f_{t-\tau}} \right|^{1/\tau}, \quad t > \tau \tag{5.2.3}$$

onde  $\tau$  é um intervalo de tempo apropriado. Para valores suficientemente grandes de  $\tau$ , tanto  $R_t$  quanto  $R_t^\dagger$  convergem para o mesmo valor.

Neste trabalho, utiliza-se as propriedades de  $R_t$  conforme  $t \rightarrow \infty$ . Observe que se  $R_t$  possui um limitante inferior positivo, então existe  $0 < \gamma \leq R_t \leq 1$ , para todo  $t \geq 0$ , tal que

$$e_t = (1 - R_t)^t e_0 \leq (1 - \gamma)^t e_0,$$

e, portanto, o erro aproximado  $e_t$  converge rapidamente para 0. Nesse caso, dizemos que a TMC é linear. Por outro lado, se  $R_t$  converge para 0, então  $e_t$  converge lentamente para 0 e temos uma TMC sublinear. Portanto, para garantir a eficiência da convergência, ajustes precisam ser feitos no algoritmo de forma a estabelecer um limitante inferior positivo para  $R_t$ . Com esse propósito, introduz-se a seguir definições e resultados adicionais.

Se  $\mathcal{S}$  é o conjunto de todas as populações do problema e  $\mathcal{S}^*$  é o conjunto de todas as populações ótimas ( $X \in \mathcal{S}^*$  se e somente se  $x^* \in X$ ), então definimos a região  $\rho$ -promissora de  $X$  como

$$\mathcal{S}(X, \rho) := \{Y \in \mathcal{S} \mid e(Y) < \rho e(X)\},$$

onde  $\rho \in (0, 1]$  é um fator de contração. A seguir, vamos discutir sobre o operador de mutação presente em algoritmos evolutivos. Vamos fornecer sua classificação e resultados que ajudarão a melhorar o desempenho do AEA, tornando a TMC linear.

É importante lembrar que, dada uma população  $X \in \mathcal{S}$  e um conjunto de populações  $\mathcal{A} \subset \mathcal{S}$ , a probabilidade de transição  $P_g(X; \mathcal{A})$  é definida como

$$P_g(X; \mathcal{A}) = \int_{\mathcal{A}} p_g(X; Y) dY,$$

onde  $p_g(X; Y)$  é a função de densidade de probabilidade de transição de  $X$  para  $Y$ . De acordo com CHEN & HE (2021), a mutação que transforma  $x$  em  $y = x + \xi$  pode ser classificada em dois grupos: invariante se a função de densidade de probabilidade  $p_g(0, \xi)$  for independente de  $X$  e, caso contrário, adaptativa.

Além disso, como pode ser visto no Lema 5.1, os autores demonstram que em algoritmos evolutivos, para obter uma TMC linear, não deve-se utilizar uma mutação invariante.

**Lema 5.1.** [CHEN & HE (2021)] Se o número de soluções ótimas é finito e se a mutação é invariante, então a probabilidade de transição para a região promissora satisfaz

$$\inf\{P_g(X; \mathcal{S}(X, 1)); X \notin \mathcal{S}^*\} = 0, \quad (5.2.4)$$

onde  $\inf$  denota o ínfimo do conjunto.

*Demonstração.* Para provar 5.2.4, é suficiente mostrar que  $\lim_{e(X) \rightarrow 0} P_g(X; \mathcal{S}(X, 1)) = 0$ . Isto é,  $\forall \epsilon > 0, \exists \delta > 0, \forall X \in \mathcal{A}(\mathcal{S}^*, \delta) \setminus \mathcal{S}^*$ , temos que

$$P_g(X; \mathcal{S}(X, 1)) < \epsilon, \quad (5.2.5)$$

onde  $\mathcal{A}(\mathcal{S}^*, \delta) = \{X; e(X) \leq \delta\}$ . Seja  $m(\mathcal{A})$  a medida de Lebesgue de um conjunto Lebesgue mensurável  $\mathcal{A} \subset \mathcal{S}$ . Como a função de distribuição de probabilidade  $p_z(Z)$  é contínua e limitada, a probabilidade de  $X + Z$  pertencer a uma pequena área é pequena,

para um  $X$  fixo. Então,  $\forall \epsilon > 0$ ,  $\exists \delta' > 0$  ( $\delta' = \epsilon / \sup p_z(Z)$ ), tal que

$$\mathbb{P}(X + Z \in \mathcal{A}) = \int_{Z: X+Z \in \mathcal{A}} p_z(X + Z) dZ < \epsilon, \quad (5.2.6)$$

$\forall \mathcal{A} \subset \mathcal{S} : m(\mathcal{A}) \leq \delta'$  e  $\forall X \in \mathcal{S}$ . Como o número de soluções ótimas é finito, temos que  $m(\mathcal{S}^*) = 0$  e como  $f$  é contínua, para o conjunto  $\mathcal{A}(\mathcal{S}^*, \delta)$  pode-se escolher  $\delta$  suficientemente pequeno tal que  $m(\mathcal{A}(\mathcal{S}^*, \delta)) \leq \delta'$ .

Ainda pela continuidade de  $f$ , podemos escolher  $\delta$  suficientemente pequeno tal que  $\forall X \in \mathcal{A}(\mathcal{S}^*, \delta)$  e  $Y \notin \mathcal{A}(\mathcal{S}^*, \delta)$ ,  $f(X) < f(Y)$ . Isso implica que  $\mathcal{S}(X, 1) \subset \mathcal{A}(\mathcal{S}^*, \delta)$ . Por 5.2.6 e como  $m(\mathcal{A}(\mathcal{S}^*, \delta)) \leq \delta'$ ,  $\forall X \in \mathcal{A}(\mathcal{S}^*, \delta) \setminus \mathcal{S}^*$ , temos que

$$\mathbb{P}(X + Z \in \mathcal{A}(\mathcal{S}^*, \delta)) < \epsilon.$$

Como  $\mathcal{S}(X, 1) \subset \mathcal{A}(\mathcal{S}^*, \delta)$ , tem-se que

$$P_g(X; \mathcal{S}(X, 1)) \leq \mathbb{P}(X + Z \in \mathcal{A}(\mathcal{S}^*, \delta)) < \epsilon.$$

Então obtém-se 5.2.5 e a prova está completa.  $\square$

Entre as mutações adaptativas, existem dois grupos: dizemos que a mutação é zero-adaptativa se

$$\inf\{P_g(X; \mathcal{S}(X, 1)); X \notin \mathcal{S}^*\} = 0,$$

e ela é positiva-adaptativa se existir  $\rho \in (0, 1)$  tal que

$$C_\rho = \inf\{P_g(X; \mathcal{S}(X, \rho)); X \notin \mathcal{S}^*\} > 0.$$

Ainda em CHEN & HE (2021), os autores demonstram que se o número de soluções ótimas é finito, a sequência  $\{e_t; t = 0, 1, \dots\}$  converge para 0, e a mutação é zero-adaptativa, então existe uma sequência  $\{X_t; t = 0, 1, \dots\}$  tal que  $\lim_{t \rightarrow \infty} R_t = 0$ . Finalmente, eles mostram que para a TMC ser linear, devemos ter um operador positivo-adaptativo, conforme evidenciado no seguinte resultado:

**Proposição 5.2** (CHEN & HE (2021)). Se a sequência  $\{e_t; t = 0, 1, \dots\}$  converge para 0 e a mutação é positiva-adaptativa com um fator de contração  $\rho \in (0, 1)$ , então  $R_t \geq (1 - \rho)C_\rho > 0$ .

### 5.3 TMC linear para otimização de problemas com funções Lipschitz contínuas

Nesta seção, consideramos o problema de minimização (5.2.1) onde  $f$  é uma função Lipschitz contínua, ou seja, existe uma constante positiva  $L_2$  tal que

$$|f(x) - f(y)| \leq L_2 \|x - y\|_2, \quad (5.3.1)$$

para todos os  $x, y$  no conjunto compacto  $\mathfrak{R} \subset \mathbb{R}^n$ .

Neste trabalho,  $\|\cdot\|_2$  denota a norma euclidiana, ou seja, para  $x = (x_1, \dots, x_n)$ ,

$$\|x\|_2 = \left( \sum_{i=1}^n x_i^2 \right)^{1/2},$$

e

$$B(p, r) = \{x \in \mathbb{R}^n : \|x - p\|_2 < r\}$$

é a bola aberta de raio  $r > 0$  centrada em um ponto  $p$  em  $\mathbb{R}^n$ . Nas demonstrações dos principais resultados deste trabalho, usaremos

$$B_\infty(p, r) = \{x \in \mathbb{R}^n : \|x - p\|_\infty < r\},$$

onde  $\|\cdot\|_\infty$  denota a norma máxima, ou seja,

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

É bem conhecido que

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq n\|x\|_\infty, \quad (5.3.2)$$

para todo  $x \in \mathbb{R}^n$ , onde  $\|x\|_1 := \sum_{i=1}^n |x_i|$ .

Para estabelecer nosso próximo resultado, consideramos primeiramente a seguinte suposição:

( $\mathcal{A}_1$ ) Para toda solução ótima  $x^*$ , existe um  $\delta > 0$  suficientemente pequeno tal que  $B(x^*, \delta) \subset \mathfrak{R}$ , e existe uma constante  $L_1 > 0$  tal que

$$f(x^*) + L_1 \|x - x^*\|_2 \leq f(x)$$

para todo  $x \in B(x^*, \delta)$ .

Em particular, devido à condição de Lipschitz (5.3.1), temos que

$$L_1\|x - x^*\|_2 \leq f(x) - f(x^*) \leq L_2\|x - x^*\|_2 \quad (5.3.3)$$

para todo  $x \in B(x^*, \delta)$ . Para um problema de dimensão  $n = 2$  ( $x \in \mathfrak{R} \subset \mathbb{R}^2$ ), por exemplo, esta condição implica que o gráfico de  $f$  numa pequena vizinhança de  $x^*$  está entre dois cones com inclinação  $L_1$  e  $L_2$  com o mesmo vértice comum em  $f(x^*)$ .

Considere a função crescente

$$\Phi(s) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^s e^{-t^2/2} dt.$$

O seguinte resultado apresenta uma escolha específica para a mutação adaptativa que torna a TMC linear e ainda fornece um limitante inferior para  $R_k$  em termos da constante de Lipschitz  $L_2$  e da constante  $L_1$  na condição  $(\mathcal{A}_1)$ .

**Teorema 5.3.** Para o Problema (5.2.1) e o Algoritmo 2 que satisfazem a condição  $(\mathcal{A}_1)$ , se o desvio padrão da mutação para um indivíduo da  $k$ -ésima geração  $x_k = (x_{k1}, \dots, x_{kn})$  é dado por

$$\sigma_{ki} = |x_{ki} - x_i^*|,$$

então o operador de mutação é positivo-adaptativo com um fator de contração  $\rho \in (0, 1)$ . Além disso, obtemos a seguinte estimativa para o limitante inferior da TMC: existe  $k_0 > 0$ , tal que

$$R_k \geq (1 - \rho) \cdot \left[ \Phi \left( 1 + \frac{\rho L_1}{n L_2} \right) - \Phi \left( 1 - \frac{\rho L_1}{n L_2} \right) \right]^n,$$

para todo  $k \geq k_0$ .

*Demonstração.* Primeiramente, demonstramos a existência de uma constante positiva  $\mathcal{K}_\rho$  tal que  $P_g(X_k; \mathcal{S}(X_k, \rho)) \geq P_g(x_k; \mathcal{S}(x_k, \rho)) > \mathcal{K}_\rho$ , onde  $x_k = (x_{k1}, \dots, x_{kn})$  representa o melhor indivíduo da população  $X_k \notin \mathcal{S}^*$ .

Lembramos que

$$\begin{aligned} \mathcal{S}(x_k, \rho) &= \{y \in \mathfrak{R} : |f(y) - f(x^*)| < \rho|f(x_k) - f(x^*)|\} \\ &= \{y \in \mathfrak{R} : f(y) \in (f(x^*) - \epsilon_k, f(x^*) + \epsilon_k)\}, \end{aligned}$$

onde  $\epsilon_k = \rho|f(x_k) - f(x^*)|$ .

Como  $f$  é Lipschitz contínua, para um dado  $\epsilon_k > 0$ , sabemos que para  $\delta_k = \frac{\epsilon_k}{nL_2}$

temos

$$f(B_\infty(x^*, \delta_k)) \subset (f(x^*) - \epsilon_k, f(x^*) + \epsilon_k).$$

Como  $\epsilon_k$  é suficientemente pequeno quando  $k$  é grande, e para a análise da taxa, consideramos o cenário conforme  $k$  tende ao infinito, sem perda de generalidade, podemos assumir que  $B_\infty(x^*, \delta_k) \subset B(x^*, \delta)$  (bola aberta da condição  $(\mathcal{A}_1)$ ) e que o melhor indivíduo  $x_k$  permanece dentro dessa bola aberta para todo  $k$  maior que algum  $k_0 > 0$ .

Portanto, a probabilidade de mover-se para a região promissora  $\mathcal{S}(x_k, \rho)$  é dada por

$$\begin{aligned} P_g(x_k; \mathcal{S}(x_k, \rho)) &= \int_{\mathcal{S}(x_k, \rho)} p_g(x_k, y) dy \\ &\geq \int_{B_\infty(x^*, \delta_k)} p_g(x_k, y) dy \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_{ki}} \int_{x_i^* - \delta_k}^{x_i^* + \delta_k} e^{-\frac{(y_i - x_{ki})^2}{2(\sigma_{ki})^2}} dy_i \end{aligned}$$

(fazendo uma mudança de variável)

$$\begin{aligned} &= \prod_{i=1}^n \left[ \Phi\left(\frac{x_i^* + \delta_k - x_{ki}}{\sigma_{ki}}\right) - \Phi\left(\frac{x_i^* - \delta_k - x_{ki}}{\sigma_{ki}}\right) \right] \\ &= \prod_{i=1}^n \left[ \Phi\left(\frac{x_i^* - x_{ki}}{\sigma_{ki}} + \frac{\rho|f(x_k) - f(x^*)|}{nL_2\sigma_{ki}}\right) - \Phi\left(\frac{x_i^* - x_{ki}}{\sigma_{ki}} - \frac{\rho|f(x_k) - f(x^*)|}{nL_2\sigma_{ki}}\right) \right] \end{aligned}$$

Observamos que se  $\sigma_{ki}$  fosse escolhido como constante, essa última expressão tenderia a zero conforme  $k$  se aproxima do infinito.

Por hipótese  $\sigma_{ki} = |x_{ki} - x_i^*|$  e se  $x_{ki} \neq x_i^*$  para todo  $i \in \{1, \dots, n\}$ , então

$$\begin{aligned} P_g(x_k; \mathcal{S}(x_k, \rho)) &\geq \prod_{i=1}^n \left[ \Phi\left(\frac{x_i^* - x_{ki}}{|x_{ki} - x_i^*|} + \frac{\rho|f(x_k) - f(x^*)|}{nL_2|x_{ki} - x_i^*|}\right) \right. \\ &\quad \left. - \Phi\left(\frac{x_i^* - x_{ki}}{|x_{ki} - x_i^*|} - \frac{\rho|f(x_k) - f(x^*)|}{nL_2|x_{ki} - x_i^*|}\right) \right] \end{aligned}$$

(como  $\Phi(-s) = 1 - \Phi(s)$ )

$$\begin{aligned} &= \prod_{i=1}^n \left[ \Phi\left(1 + \frac{\rho|f(x_k) - f(x^*)|}{nL_2|x_{ki} - x_i^*|}\right) - \Phi\left(1 - \frac{\rho|f(x_k) - f(x^*)|}{nL_2|x_{ki} - x_i^*|}\right) \right] \\ &\geq \prod_{i=1}^n \left[ \Phi\left(1 + \frac{\rho L_1}{nL_2}\right) - \Phi\left(1 - \frac{\rho L_1}{nL_2}\right) \right] \end{aligned}$$

$$:= \mathcal{K}_\rho$$

onde na última igualdade usamos a condição  $(\mathcal{A}_1)$  e o fato que

$$\frac{\rho L_1}{nL_2} \leq \frac{\rho L_1 \|x_k - x^*\|_2}{nL_2 |x_{ki} - x_i^*|} \leq \frac{\rho |f(x_k) - f(x^*)|}{nL_2 |x_{ki} - x_i^*|}. \quad (5.3.4)$$

Pela Proposição 5.2 tem-se que

$$R_k \geq (1 - \rho)C_\rho \geq (1 - \rho)\mathcal{K}_\rho > 0,$$

para  $k \geq k_0$ . □

Para a mesma classe de funções como no Teorema 5.3, propomos um desvio padrão alternativo que considera a distância entre  $f(x_k)$  e  $f(x^*)$ . Em seguida, demonstramos que a TMC permanece linear, porém obtemos uma estimativa diferente para o seu limitante inferior.

**Corolário 5.4.** Para o Problema (5.2.1) e o Algoritmo 2 satisfazendo as condições  $(\mathcal{A}_1)$ , se o desvio padrão associado à mutação para um indivíduo da  $k$ -ésima geração  $x_k = (x_{k1}, \dots, x_{kn})$  é dado por

$$\sigma_{ki} = |f(x_k) - f(x^*)|, \quad \text{para todo } i = 1, \dots, n,$$

então o operador de mutação é positivo-adaptativo com um fator de contração  $\rho \in (0, 1)$ .

Além disso, obtemos a seguinte estimativa para o limitante inferior da TMC:

$$R_k \geq (1 - \rho) \cdot \left[ \Phi \left( \frac{1}{L_1} + \frac{\rho}{nL_2} \right) - \Phi \left( \frac{1}{L_1} - \frac{\rho}{nL_2} \right) \right]^n.$$

*Demonstração.* Sob as mesmas condições do Teorema 5.3, tem-se que

$$P_g(x_k; \mathcal{S}(x_k, \rho)) \geq \prod_{i=1}^n \left[ \Phi \left( \frac{x_i^* - x_{ki}}{\sigma_{ki}} + \frac{\rho |f(x_k) - f(x^*)|}{nL_2 \sigma_{ki}} \right) - \Phi \left( \frac{x_i^* - x_{ki}}{\sigma_{ki}} - \frac{\rho |f(x_k) - f(x^*)|}{nL_2 \sigma_{ki}} \right) \right].$$

Como  $\sigma_{ki} = |f(x_k) - f(x^*)| > 0$  vem que

$$P_g(x_k; \mathcal{S}(x_k, \rho)) \geq \prod_{i=1}^n \left[ \Phi \left( \frac{x_i^* - x_{ki}}{|f(x_k) - f(x^*)|} + \frac{\rho |f(x_k) - f(x^*)|}{nL_2 |f(x_k) - f(x^*)|} \right) - \Phi \left( \frac{x_i^* - x_{ki}}{|f(x_k) - f(x^*)|} - \frac{\rho |f(x_k) - f(x^*)|}{nL_2 |f(x_k) - f(x^*)|} \right) \right] \quad (5.3.5)$$

$$\begin{aligned} &= \prod_{i=1}^n \left[ \Phi \left( \frac{x_i^* - x_{ki}}{|f(x_k) - f(x^*)|} + \frac{\rho}{nL_2} \right) - \Phi \left( \frac{x_i^* - x_{ki}}{|f(x_k) - f(x^*)|} - \frac{\rho}{nL_2} \right) \right] \\ &\geq \prod_{i=1}^n \left[ \Phi \left( \frac{1}{L_1} + \frac{\rho}{nL_2} \right) - \Phi \left( \frac{1}{L_1} - \frac{\rho}{nL_2} \right) \right] \\ &:= \mathcal{K}_\rho \end{aligned} \quad (5.3.6)$$

onde a última desigualdade decorre da condição  $(\mathcal{A}_1)$  e do fato que

$$\frac{|x_i^* - x_{ki}|}{|f(x_k) - f(x^*)|} \leq \frac{|x_i^* - x_{ki}|}{L_1 \|x^* - x_k\|} \leq \frac{1}{L_1}$$

e que a função

$$\alpha(s) = \Phi \left( s + \frac{\rho}{nL_2} \right) - \Phi \left( s - \frac{\rho}{nL_2} \right)$$

é crescente em  $[-1/L_1, 0)$  e decrescente em  $(0, 1/L_1]$ .

Novamente, pela Proposição 5.2, existe  $k_0 \geq 0$  tal que

$$R_k \geq (1 - \rho)C_\rho \geq (1 - \rho)\mathcal{K}_\rho > 0$$

para  $k \geq k_0$ . □

Apenas para ilustrar o Teorema 5.3 e o Corolário 5.4, aqui estão alguns exemplos de funções definidas em um retângulo  $\mathfrak{R}$  contendo o ponto de mínimo  $a = (a_1, \dots, a_n)$  que satisfazem a condição  $(\mathcal{A}_1)$ :

1.  $f(x_1, \dots, x_n) = k + \sum_{i=1}^n c_i |x_i - a_i|$ ;
2.  $f(x_1, \dots, x_n) = \sqrt{\sum_{i=1}^n c_i (x_i - a_i)^2} + \sum_{i=1}^n d_i (x_i - a_i)^2$ ;
3.  $f(x_1, \dots, x_n) = k + \sum_{i=1}^n c_i |\text{sen}(x_i - a_i)|$ ,

onde  $c_1, \dots, c_n$  são constantes positivas, e  $k, d_1, \dots, d_n$  são constantes não negativas. Em geral, podemos tomar como exemplo qualquer outra função cujo gráfico nas proximidades do ótimo global fique acima da superfície de um cone generalizado com inclinação  $L_1$  (condição  $\mathcal{A}_1$ ), e onde as inclinações das linhas secantes conectando dois pontos no gráfico da função sejam limitadas por  $L_2$ . Vale ressaltar que essas funções são mais gerais do que as apresentadas em CHEN & HE (2021).

Observamos que funções Lipschitz contínuas definidas em um conjunto compacto, que se comportam aproximadamente como um parabolóide elíptico na vizinhança do ponto mínimo  $x^*$ , não satisfazem a condição  $(\mathcal{A}_1)$ . Portanto, nossa estratégia é apresentar um resultado semelhante ao Teorema 5.3, mas que possa acomodar funções como  $f(x) = x^2$ ,  $g(x, y) = x^4 + y^4$ ,  $h(x, y, z) = 2x^6 + y^6 + 3z^6$ , etc., ou funções que, pelo menos na vizinhança de seu mínimo global, exibam comportamento semelhante ao mínimo dessas funções. Mais precisamente, consideraremos funções objetivo do Problema (5.2.1) que satisfaçam a seguinte condição:

$(\mathcal{A}_2)$  Para qualquer solução ótima  $x^*$ , existe um pequeno  $\delta > 0$  tal que  $B(x^*, \delta) \subset \mathfrak{X}$ , e existem constantes  $L_3 > 0$  e  $L_4 > 0$  tais que

$$L_3 \|x - x^*\|_2^{2j} \leq f(x) - f(x^*) \leq L_4 \|x - x^*\|_2^{2j}$$

para todo  $x \in B(x^*, \delta)$  e  $j$  um inteiro positivo.

Para um problema de dimensão  $n = 2$  e  $j = 1$ , por exemplo, esta condição implica que o gráfico de  $f$  em uma pequena vizinhança de  $x^*$  está situado entre dois paraboloides elípticos com seu mínimo em  $f(x^*)$ . Além disso, vale ressaltar que como  $j$  pode ser qualquer inteiro positivo, abrangemos uma ampla classe de funções. Por exemplo, a função  $f(x, y) = x^4 + y^4$  satisfaz a condição  $(\mathcal{A}_2)$  com  $j = 2$  e  $(x^*, y^*) = (0, 0)$  pois  $2x^2y^2 \leq x^4 + y^4$ , e portanto

$$\frac{1}{2} \|(x, y)\|_2^4 \leq f(x, y) \leq \|(x, y)\|_2^4, \quad (5.3.7)$$

para todo  $(x, y) \in \mathbb{R}^2$ . No entanto, existem funções que satisfazem a condição  $(\mathcal{A}_1)$  mas não satisfazem  $(\mathcal{A}_2)$ . Por exemplo, considere a função  $f(x) = |x|$ . Em outras palavras, cada condição considera uma classe específica de funções baseada em seu comportamento numa vizinhança do ótimo global.

**Teorema 5.5.** Para o Problema (5.2.1) e o Algoritmo 2 satisfazendo a condição  $(\mathcal{A}_2)$  para algum inteiro positivo  $j$ , se o desvio padrão da mutação para um indivíduo na  $k$ -ésima geração  $x_k = (x_{k1}, \dots, x_{kn})$  é dado por

$$\sigma_{ki} = |x_{ki} - x_i^*|, \quad \text{para todo } i = 1, \dots, n,$$

então o operador de mutação é positivo-adaptativo com fator de contração  $\rho \in (0, 1)$ . Além disso, obtemos a seguinte estimativa para o limitante inferior da TMC: existe  $k_0 > 0$  tal que

$$R_k \geq (1 - \rho) \left[ \Phi \left( 1 + \frac{1}{n} \left( \frac{\rho L_3}{L_4} \right)^{\frac{1}{2j}} \right) - \Phi \left( 1 - \frac{1}{n} \left( \frac{\rho L_3}{L_4} \right)^{\frac{1}{2j}} \right) \right]^n,$$

para todo  $k \geq k_0$ .

*Demonstração.* Utilizando a mesma configuração da prova do Teorema 5.3, se agora considerarmos  $\delta_k = \frac{1}{n} \left( \frac{\epsilon_k}{L_4} \right)^{\frac{1}{2j}}$ , obtemos pela condição  $(\mathcal{A}_2)$  e pela desigualdade (5.3.2) que

$$|f(x) - f(x^*)| \leq n^{2j} L_4 \|x - x^*\|_\infty^{2j} < \epsilon_k, \quad (5.3.8)$$

para todo  $x \in B_\infty(x^*, \delta_k)$ . Portanto,

$$f(B_\infty(x^*, \delta_k)) \subset (f(x^*) - \epsilon_k, f(x^*) + \epsilon_k), \quad (5.3.9)$$

onde  $\epsilon_k = \rho |f(x_k) - f(x^*)|$ .

Agora, usando novamente a condição  $(\mathcal{A}_2)$ , observa-se que

$$\left( \frac{\rho}{n^{2j} L_4} \right)^{\frac{1}{2j}} \frac{|f(x_k) - f(x^*)|^{\frac{1}{2j}}}{|x_{ki} - x_i^*|} \geq \left( \frac{\rho L_3}{n^{2j} L_4} \right)^{\frac{1}{2j}} \frac{\|x_k - x^*\|_2}{|x_{ki} - x_i^*|} \geq \left( \frac{\rho L_3}{n^{2j} L_4} \right)^{\frac{1}{2j}}.$$

Portanto, seguindo os mesmos passos da prova do Teorema 5.3, temos

$$\begin{aligned} P_g(x_k; \mathcal{S}(x_k, \rho)) &\geq \mathcal{K}_\rho \\ &= \prod_{i=1}^n \left[ \Phi \left( 1 + \left( \frac{\rho L_3}{n^{2j} L_4} \right)^{\frac{1}{2j}} \right) - \Phi \left( 1 - \left( \frac{\rho L_3}{n^{2j} L_4} \right)^{\frac{1}{2j}} \right) \right] \end{aligned}$$

e pela Proposição 5.2, obtém-se que existe  $k_0 \geq 0$  tal que

$$R_k \geq (1 - \rho) \mathcal{K}_\rho > 0,$$

para todo  $k \geq k_0$ . □

É importante notar que se tomarmos  $\sigma_{ki} = |f(x_k) - f(x^*)|$  para uma função  $f$  que satisfaz a condição  $(\mathcal{A}_2)$ , não podemos garantir a linearidade da taxa usando as mesmas técnicas propostas na prova do Corolário 5.4. Por exemplo, considerando  $f(x) = x^2$ , a expressão (5.3.5) tende a zero conforme  $x_k \rightarrow 0$ .

## 5.4 Aplicações do AEA

Os resultados teóricos relacionados à TMC podem ser aplicados a várias funções de benchmark. Nesta seção, os testes serão limitados aos casos particulares de problemas de otimização que visam minimizar a função Ackley, a função Alpine e  $f(x, y) = x^4 + y^4$ . Além disso, como a função custo do PDEPV não satisfaz a hipótese  $(\mathcal{A}_1)$ , adaptações nas hipóteses foram feitas para que fosse possível mostrar que a TMC é linear neste problema, ou seja, serão incluídos mais dois resultados (Corolários 5.6 e 5.7). Para os testes numéricos, foram implementados dois casos particulares de AEA, um deles baseado no AG e outro baseado na Estratégia Evolutiva (EE).

### 5.4.1 Algoritmos Implementados

Para que possamos fazer os testes computacionais utilizamos a Lei dos Grandes Números, um conceito fundamental na teoria das probabilidades e estatística, que fornece uma base para a inferência estatística e a tomada de decisões com base em dados. Ela nos permite fazer estimativas confiáveis com base em amostras e é fundamental para o entendimento de fenômenos aleatórios, uma vez que estabelece que, à medida que o número de experimentos independentes aumenta, a média desses experimentos tende a se aproximar do valor esperado ou médio do evento em questão. Isso significa que, mesmo que haja variações nos resultados individuais, a tendência geral será de se aproximar do resultado esperado, conforme mais execuções são realizadas.

Por exemplo, considere o lançamento de um dado justo (não viciado). O valor esperado para cada lançamento é de 3.5 (a média aritmética dos números de 1 a 6). Se lançarmos o dado várias vezes e formos fazendo a média entre os resultados obtidos, a Lei dos Grandes Números garante que esta média se aproximará cada vez mais de 3.5, à

medida que o número de lançamentos aumenta.

Para os testes envolvendo a TMC, a Lei dos Grandes Números garante que o valor aproximado de  $f_k = \mathbb{E}(x_k)$  pode ser calculado computacionalmente como

$$\frac{1}{T} \left( f(x_k^{[1]}) + \dots + f(x_k^{[T]}) \right),$$

onde  $T$  é um número suficientemente grande de execuções, e  $f(x_k^{[j]})$  representa o custo do melhor indivíduo da população na  $k$ -ésima iteração da  $j$ -ésima execução. Portanto, para obter uma aproximação para  $f_k$ , devemos executar o algoritmo um número suficientemente grande de vezes. Para cada iteração  $k$ , o melhor indivíduo de cada execução deve ser salvo, e o custo médio na iteração  $k$  deve ser calculado.

O AG é um tipo de AE que gera uma nova população através de um cruzamento (crossover) entre os indivíduos, visando diversificar ainda mais sua busca. Para ilustrar os resultados apresentados neste trabalho, propomos o Algoritmo Genético Adaptativo (AGA) para resolver alguns problemas de otimização. O AGA incorpora o crossover ao AE descrito na Seção 5.1, e seu pseudocódigo é fornecido no Algoritmo 3. Como as variáveis dos problemas são canalizadas, o ajuste é feito da seguinte forma: sempre que uma variável estiver fora de seus limites, ela é ajustada para o valor do limitante mais próximo.

A seleção, que consiste em escolher os pais que passarão pelo processo de crossover, ocorre de forma aleatória e, tipicamente, indivíduos com melhores avaliações têm uma chance maior de serem escolhidos. Os novos indivíduos (descendentes) são gerados por meio do cruzamento dos indivíduos selecionados (pais), visando uma evolução contínua dos indivíduos em cada geração. De acordo com ENGELBRECHT (2007), esse cruzamento pode ser assexual (onde um dos pais gera toda a prole), sexual (onde dois pais são usados para produzir um ou dois descendentes) ou de múltipla recombinação (onde mais de dois pais são usados para gerar um ou mais descendentes). Neste trabalho, optamos pelo cruzamento sexual. Uma vez selecionados os pais, escolhemos aleatoriamente um ponto de cruzamento e recombinaos os indivíduos, gerando dois descendentes.

A EE é outro tipo de AE que se concentra em estratégias de mutação para guiar a busca no espaço de solução. Nesta seção, além do AGA, utilizamos a EEA para resolver estes problemas de otimização. O pseudocódigo do EEA é fornecido no Algoritmo 4, a

---

**Algoritmo 3** Algoritmo Genético Adaptativo - AGA
 

---

**Input**  $p_M, n_I, n_g$   
 $k \leftarrow 0$   
 Gere a população inicial  $X = \{x_0, \dots, x_{n_I-1}\}$  aleatoriamente;  
**while**  $k < n_g$  **do**  
   Avalie os indivíduos em  $X$   
   **for**  $t = 0, 1, \dots, (n_I - 1)/2$  **do**  
     Selecione dois indivíduos (pais)  $x_i$  e  $x_j$   
     Aplique o crossover, gerando  $y_i$  e  $y_j$  (filhos)  
      $r_t \sim \mathcal{N}(0, 1)$   
     **if**  $r_t \leq p_M$  **then**  
       Determine  $\sigma_i$  e  $\sigma_j$ ;  
        $y_i = y_i + \xi_i$  onde  $\xi_i \in \mathcal{N}(0, \sigma_i)$   
        $y_j = y_j + \xi_j$  onde  $\xi_j \in \mathcal{N}(0, \sigma_j)$   
     **end if**  
     Ajuste e avalie  $y_i$  e  $y_j$   
     Crie uma lista ordenada de  $[x_i, x_j, y_i, y_j]$  com base na aptidão (fitness)  
     Substitua  $x_i$  e  $x_j$  pelos dois indivíduos com a melhor aptidão na lista ordenada  
   **end for**  
    $k = k + 1$   
**end while**  
 Retorne o melhor indivíduo de  $X$  como solução do problema.

---

geração de uma nova população é feita de forma assexual e cada indivíduo na população tem uma probabilidade  $p_M$  de sofrer a mutação adaptativa. Os ajustes de indivíduos ineficazes foram feitos exatamente da mesma forma que no AGA.

### 5.4.2 Função Ackley

Um problema de otimização com a função Ackley como sua função objetivo torna-se um desafio interessante para algoritmos de otimização devido à presença de múltiplos mínimos locais, frequentemente fazendo com que muitos algoritmos fiquem estagnados nesses pontos. Essa característica torna a função Ackley interessante para testar a capacidade desses algoritmos de lidar com problemas complexos e multi-modais.

O problema de otimização com a função objetivo definida pela função de Ackley

---

**Algoritmo 4** Estratégia Evolutiva Adaptativa - EEA
 

---

**Input**  $p_M, n_I, n_g$   
 $k \leftarrow 0$   
 Gere a população inicial  $X = \{x_0, \dots, x_{n_I-1}\}$  aleatoriamente;  
**while**  $k < n_g$  **do**  
   Avalie os indivíduos em  $X$   
   **for**  $t = 0, 1, \dots, (n_I - 1)$  **do**  
     Selecione um indivíduo  $x_i$   
      $r_i \sim (0, 1)$   
     **if**  $r_i \leq p_M$  **then**  
       Determine  $\sigma_i$ ;  
        $y_i = x_i + \xi_i$  onde  $\xi_i \in \mathcal{N}(0, \sigma_i)$   
       Ajuste  $y_i$   
     **end if**  
     **if**  $f(y_i) \leq f(x_i)$  **then**  
        $x_i \leftarrow y_i$   
     **end if**  
   **end for**  
    $k = k + 1$   
**end while**  
 Retorne o melhor indivíduo de  $X$  como solução do problema.

---

pode ser modelado da seguinte forma:

$$\begin{cases} \min_x a - a \cdot \exp\left(-b\sqrt{\frac{1}{n}\sum_{i=1}^n x_i^2}\right) + e - \exp\left(\frac{1}{n}\sum_{i=1}^n \cos(cx_i)\right) \\ s.a. \\ x_i^{\min} \leq x_i \leq x_i^{\max}, \end{cases} \quad (5.4.1)$$

onde  $a, b$  e  $c$  são constantes positivas. A função  $\exp$  representa a função exponencial, e  $e = \exp(1)$ .

Para aplicar os resultados obtidos na seção anterior utiliza-se o caso bidimensional de (5.4.1). Destaca-se que os resultados são válidos para qualquer dimensão  $n$ . Em outras palavras, nosso objetivo é minimizar a função multi-modal

$$f(x, y) = a - a \exp\left(\frac{-b}{\sqrt{2}}\sqrt{x^2 + y^2}\right) + e - \exp\left(\frac{\cos(cx) + \cos(cy)}{2}\right) \quad (5.4.2)$$

no retângulo  $\mathfrak{R} = [x^{\min}, x^{\max}] \times [y^{\min}, y^{\max}]$  contendo o ponto  $(0, 0)$ .

Para determinar as estimativas para o limitante inferior da TMC (Teorema 5.3 e Teorema 5.4), primeiro precisamos determinar a constante de Lipschitz  $L_2$ , que será obtida por meio das seguintes observações:

- (a) Se duas funções  $g_1 : \mathfrak{R} \rightarrow \mathbb{R}$  e  $g_2 : \mathfrak{R} \rightarrow \mathbb{R}$  são Lipschitz contínuas com constantes  $k_1$  e  $k_2$  respectivamente, então a função  $g_1 + g_2$  é Lipschitz contínua com constante  $k_1 + k_2$ . E se  $c$  é uma constante, então a função  $cg_1$  é Lipschitz contínua com constante  $|c|k_1$ ;
- (b) Se  $g_1 : \mathbb{R} \rightarrow \mathbb{R}$  e  $g_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$  são Lipschitz contínuas com constantes  $k_1$  e  $k_2$  respectivamente, então a função  $g_1 \circ g_2$  é Lipschitz contínua com constante  $k_1k_2$ . De fato,

$$\begin{aligned} |g_1 \circ g_2(x, y) - g_1 \circ g_2(u, v)| &\leq k_1 |g_2(x, y) - g_2(u, v)| \\ &\leq k_1 k_2 \|(x, y) - (u, v)\|_2; \end{aligned}$$

- (c) A função  $\eta(x, y) = \sqrt{x^2 + y^2}$  é Lipschitz contínua com constante 1 e a função  $\psi(x, y) = \cos(cx) + \cos(cy)$  é Lipschitz contínua com constante  $2|c|$ ;
- (d) Pelo Teorema do Valor Médio, tem-se que a função  $\xi : [0, k] \rightarrow \mathbb{R}$  definida por  $\xi(t) = \exp\left(\frac{-b}{\sqrt{2}}t\right)$  é Lipschitz contínua com constante  $\frac{b}{\sqrt{2}}$ . Analogamente, a função  $\alpha : [-1, 1] \rightarrow \mathbb{R}$  definida por  $\alpha(t) = \exp(t)$  é Lipschitz contínua com constante  $e$ .

Portanto,

$$L_2 = \frac{ab}{\sqrt{2}} + ec$$

é a constante de Lipschitz para a função de Ackley  $f$ .

Agora, vamos verificar se  $f$  satisfaz a hipótese  $(\mathcal{A}_1)$ , ou seja,

$$f(x, y) \geq f(x^*, y^*) + L_1 \|(x, y) - (x^*, y^*)\|_2,$$

onde

$$L_1 = \frac{ab}{\sqrt{2}} \exp\left(\frac{-b\delta}{\sqrt{2}}\right).$$

De fato, primeiro, observamos que

$$f(x, y) \geq a \left[ 1 - \exp\left(\frac{-b}{\sqrt{2}} \sqrt{x^2 + y^2}\right) \right], \quad \forall (x, y) \in \mathfrak{R}.$$

Seja  $\gamma(t) = a \left[ 1 - \exp\left(\frac{-b}{\sqrt{2}}t\right) \right]$ ,  $t = \sqrt{x^2 + y^2}$  e  $t \in (0, \delta)$  para algum  $\delta > 0$  tal que

$B(0, \delta) \subset \mathfrak{R}$ . Aplicando o Teorema do Valor Médio, temos que existe  $\theta \in (0, \delta)$  tal que

$$\begin{aligned}\gamma(t) &\geq \frac{abt}{\sqrt{2}} \exp\left(\frac{-b}{\sqrt{2}}\theta\right) \\ &\geq \frac{abt}{\sqrt{2}} \exp\left(\frac{-b}{\sqrt{2}}\delta\right)\end{aligned}$$

e portanto  $f(x, y) \geq L_1\|(x, y)\|_2$  conforme desejado.

Para os testes numéricos, foram considerados os seguintes parâmetros:  $a = 20$ ,  $b = 0.2$ ,  $c = 2\pi$ ,  $x^{min} = y^{min} = -32$ , e  $x^{max} = y^{max} = 32$ . Para este exemplo particular, uma probabilidade de mutação  $p_M = 0.1$  foi adotada. Foram realizadas 100 execuções, cada uma delas com  $n_g = 2000$  gerações. O tamanho da população foi definido em  $n_I = 50$  indivíduos e para ambas mutações, usamos o ótimo global conhecido  $x^* = (0, 0)$ . Para o caso invariante,  $\sigma_{ki} = 1$  foi usado. Com estes dados, o AGA retornou a solução  $f(5.572 \cdot 10^{-13}, 3.588 \cdot 10^{-11}) = 3.813 \cdot 10^{-11}$ .

Para as estimativas, foi considerado  $\rho = 0.9$  e  $\delta = 10^{-4}$ . Então

$$L_1 = 2.828 \quad \text{and} \quad L_2 = 19.908,$$

e portanto

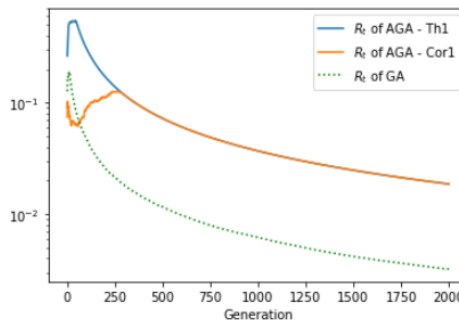
$$Est_1 := (1 - \rho) \cdot \left[ \Phi\left(1 + \frac{\rho L_1}{2L_2}\right) - \Phi\left(1 - \frac{\rho L_1}{2L_2}\right) \right]^2 = 9.573 \cdot 10^{-5}$$

e

$$Est_2 := (1 - \rho) \cdot \left[ \Phi\left(\frac{1}{L_1} + \frac{\rho}{2L_2}\right) - \Phi\left(\frac{1}{L_1} - \frac{\rho}{2L_2}\right) \right]^2 = 2.870 \cdot 10^{-5}.$$

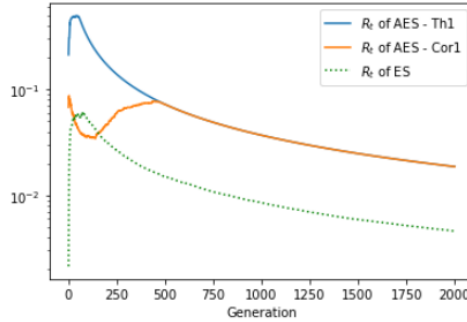
Neste exemplo foi implementado o AGA e o EEA com as duas opções de mutação adaptativa (Teorema 5.3 e Corolário 5.4) e o AG (mutação invariante).

Figura 5.1: Taxa Média de Convergência do AGA (Teorema 5.3 e Corolário 5.4) e do AG em 2000 gerações, na resolução do Problema 5.4.1.



Fonte: Elaborada pelo autor.

Figura 5.2: Taxa Média de Convergência do EEA (Teorema 5.3 e Corolário 5.4) e do EE em 2000 gerações, na resolução do Problema 5.4.1.



Fonte: Elaborada pelo autor.

Tanto na Figura 5.1 quanto na Figura 5.2, podemos ver que, inicialmente, o algoritmo adaptativo que emprega a mutação proposta no Teorema 5.3 supera aquele com a mutação sugerida no Corolário 5.4, mas após certo ponto, as taxas se igualam. Além disso, é claro que o AGA e o EEA consistentemente exibem uma TMC mais alta do que o AG e o EE, respectivamente, destacando que as modificações introduzidas, particularmente a incorporação de mecanismos de busca adaptativos, melhoraram sua capacidade de convergir para a solução ótima do problema.

### 5.4.3 Função Alpine

A função Alpine é uma função de referência amplamente utilizada para testar algoritmos de otimização e a capacidade desses algoritmos de lidar com problemas complexos e multi-modais. Se a considerarmos como a função objetivo, podemos modelar o problema da seguinte forma:

$$\left\{ \begin{array}{l} \min_x \sum_{i=1}^n |x_i \sin(x_i) + 0.1x_i| \\ s.a. \\ x_i^{min} \leq x_i \leq x_i^{max}. \end{array} \right. \quad (5.4.3)$$

Nesta seção, considera-se uma região factível em torno da origem, sabendo que a função Alpine tem um ótimo global em  $f(x^*) = f(0, \dots, 0) = 0$ .

Para determinar a constante de Lipschitz  $L_2$  da função Alpine neste domínio, além das observações feitas anteriormente, observe que:

- (a) Se duas funções  $g_1 : \mathfrak{R} \rightarrow \mathbb{R}$  e  $g_2 : \mathfrak{R} \rightarrow \mathbb{R}$  são Lipschitz contínuas com constantes  $k_1$  e  $k_2$  respectivamente, então a função produto  $g_1 \cdot g_2$  é Lipschitz contínua com constante  $k_2 \cdot \sup_{x \in \mathfrak{R}} |g_1(x)| + k_1 \cdot \sup_{x \in \mathfrak{R}} |g_2(x)|$ ;
- (b) A função  $\psi(x) = \text{sen}(x)$  é Lipschitz contínua com constante 1.

Portanto, podemos ver que

$$L_2 = \sum_{i=1}^n \left( x_i^{max} + \sup_{x_i \in [x_i^{min}, x_i^{max}]} |\text{sen}(x_i)| + 0.1 \right)$$

é uma constante de Lipschitz para a função Alpine.

Não é difícil ver através do Teorema do Valor Médio que, para  $n = 1$ , a constante  $L_1 = 0.09$  é suficiente para garantir a condição  $(\mathcal{A}_1)$ . Portanto, para qualquer  $n \geq 1$ , temos:

$$f(x) = \sum_{i=1}^n |x_i \text{sen}(x_i) + 0.1x_i| \geq \sum_{i=1}^n 0.09|x_i| = 0.09\|x\|_1 \geq 0.09\|x\|_2.$$

A seguir são apresentados os resultados obtidos a partir da aplicação do AGA e do EEA para resolver o Problema (5.4.3) com dimensão  $n = 10$ . Foram utilizados os seguintes parâmetros:  $x_i^{min} = -10$ ,  $x_i^{max} = 10$ ,  $p_M = 0.1$ , população com  $n_I = 50$  indivíduos e 100 execuções foram realizadas, cada uma delas com  $n_g = 2000$  gerações. Para ambas mutações, usou-se o ótimo global  $x^* = (0, \dots, 0)$ . Para o caso invariante,  $\sigma_{ki} = 1$  foi usado.

Similarmente à aplicação anterior, para os parâmetros escolhidos, temos  $L_2 = 111$ , e portanto,

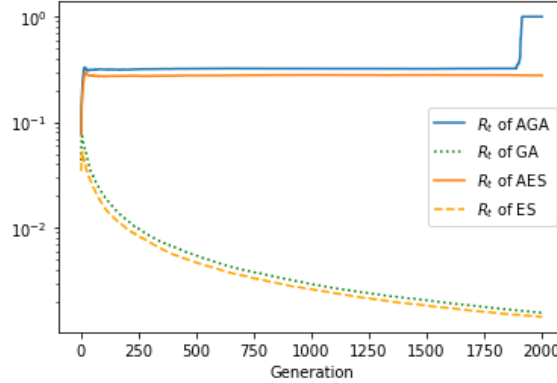
$$Est_1 := 3.017 \cdot 10^{-46} \text{ e } Est_2 = 2.913 \cdot 10^{-158}.$$

Apesar das estimativas baixas, principalmente devido à complexidade da função Alpine, elas ainda garantem a linearidade da TMC, como provamos no Teorema 5.3 e Corolário 5.4.

Para avaliar o impacto das modificações propostas no AE e comparar os dois algoritmos implementados, a Figura 5.3 exibe a TMC para o AG, AGA (Teorema 5.3), EE e EEA (Teorema 5.3) ao longo de 2000 gerações. É evidente que os algoritmos adaptativos consistentemente exibem uma TMC mais alta do que o algoritmo invariante, destacando

que as modificações introduzidas, particularmente a incorporação de mecanismos de busca adaptativos, aprimoraram sua capacidade de convergir para a solução ótima do problema. Além disso, note que no AGA, a TMC atinge o valor 1, indicando que em todas as 100 execuções, a solução ótima foi encontrada pelo algoritmo.

Figura 5.3: Uma comparação entre a TMC do AGA (Teorema 5.3), EEA (Teorema 5.3), AG e EE em 2000 gerações na resolução do Problema (5.4.3) com  $n = 10$ .



Fonte: Elaborada pelo autor.

#### 5.4.4 Função $f(x, y) = x^4 + y^4$

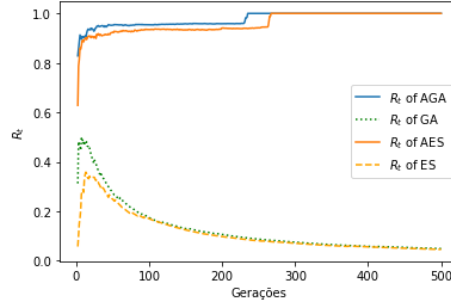
Finalmente, para ilustrar numericamente o Teorema 5.5, utilizamos o AGA e o EEA com a função de teste  $f(x, y) = x^4 + y^4$ . Para essas simulações, focamos no retângulo  $[-10, 10] \times [-10, 10]$ . De acordo com (5.3.7), temos que  $f(x, y)$  satisfaz a condição  $(\mathcal{A}_2)$ , com  $L_3 = \frac{1}{2}$  e  $L_4 = 1$ . Com essas constantes, para  $\rho = 0.9$ , temos:

$$R_k \geq 0.392 \cdot 10^{-2}.$$

Para essas simulações, definimos os seguintes parâmetros:  $p_M = 0.1$ , o tamanho da população é  $n_I = 50$  indivíduos e foram realizadas 100 execuções, cada uma com  $n_g = 500$  gerações. Para ambos os algoritmos, utilizamos o conhecido ótimo global  $x^* = (0, 0)$  e, para o caso invariante, foi usado  $\sigma_{ki} = 1$ .

A Figura 5.4 exibe a TMC do AG, do AGA (Teorema 5.5), do EE e do EEA (Teorema 5.5) ao longo de 500 gerações. Podemos notar, assim como nos exemplos anteriores, que os algoritmos adaptativos exibem uma TMC mais alta do que os algoritmos invariantes. Além disso, vale destacar que ambos os métodos adaptativos atingem o valor de 1, indicando que em todas as 100 execuções, a solução ótima foi encontrada pelos algoritmos.

Figura 5.4: Comparação entre a TMC do AGA (Teorema 5.5), EEA (Teorema 5.5), AG and EE em 500 gerações na minimização de  $f(x, y) = x^4 + y^4$ .



Fonte: Elaborada pelo autor.

### 5.4.5 Discussão: $x^*$ desconhecido

Para os casos em que o ótimo global  $x^*$  é desconhecido, uma alternativa é encontrar um desvio padrão  $\sigma_{ki}$  de mutação para um indivíduo da  $k$ -ésima geração  $x_k = (x_{k1}, \dots, x_{kn})$  que não dependa de  $x^*$ . Nesta seção, algumas possibilidades são apresentadas, comparações gráficas de suas TMC são realizadas e analisa-se a dificuldade inerente em encontrar uma adaptação que independa de  $x^*$  e que ainda seja possível obter uma TMC linear.

Para isso, faz-se primeiramente a seguinte observação. Considere um operador dado por

$$\sigma_{ki}^{\psi} := |\psi(x_{ki}, x_{(k-1)i}, \dots, x_{0i})|,$$

onde  $\psi$  é uma função definida em termos da sequência aleatória até a  $k$ -ésima geração. Seguindo a demonstração do Teorema 5.3, mas agora utilizando  $\sigma_{ki}^{\psi}$  na desigualdade (5.3.4), obtemos

$$P_g(x_k; \mathcal{S}(x_k, \rho)) \geq \prod_{i=1}^n \left[ \Phi \left( \frac{x_i^* - x_{ki}}{\sigma_{ki}^{\psi}} + \frac{\rho}{nL_2} \frac{|f(x_k) - f(x^*)|}{\sigma_{ki}^{\psi}} \right) - \Phi \left( \frac{x_i^* - x_{ki}}{\sigma_{ki}^{\psi}} - \frac{\rho}{nL_2} \frac{|f(x_k) - f(x^*)|}{\sigma_{ki}^{\psi}} \right) \right]. \quad (5.4.4)$$

Seja

$$c_{ki} := \frac{|x_{ki} - x_i^*|}{\sigma_{ki}^{\psi}}. \quad (5.4.5)$$

Usando os mesmos argumentos de (5.3.4), tem-se que

$$P_g(x_k; \mathcal{S}(x_k, \rho)) \geq \prod_{i=1}^n \left[ \Phi \left( c_{ki} \left[ 1 + \frac{\rho L_1}{n L_2} \right] \right) - \Phi \left( c_{ki} \left[ 1 - \frac{\rho L_1}{n L_2} \right] \right) \right] \quad (5.4.6)$$

para  $k$  suficientemente grande. Precisamos encontrar um  $\sigma_{ki}^\psi$  que impeça o lado direito da expressão (5.4.6) de tender a zero, de modo que  $c_{ki}$  não tende a zero e nem ao infinito. Além disso, para assegurar a linearidade da taxa utilizando a mesma configuração das demonstrações dos Teoremas 1 e 2, ainda é necessário garantir que  $\inf_k \{c_{ki}\}$  seja igual a alguma constante positiva. Acreditamos que encontrar tal  $\sigma_{ki}^\psi$  é um tema relevante e desafiador para trabalhos futuros. A seguir, apresentamos algumas possibilidades de mutações adaptativas utilizando operadores da forma  $\sigma_{ki}^\psi$ , que tendem a zero à medida que  $k$  se aproxima do infinito. Graficamente, pode-se observar que, com as mutações adaptativas sugeridas, o desempenho do algoritmo é superior em comparação com o algoritmo com  $\sigma_{ki}$  constante. Essas mutações propostas têm a vantagem de não requerer conhecimento do ótimo global.

A primeira alternativa apresentada envolve a média aritmética entre os termos da sequência  $x_{ki}$ . Esta alternativa foi pensada devido ao fato que se  $x_{ki} \rightarrow x_i^*$  q.s. conforme  $k \rightarrow \infty$ , temos que a sequência das médias parciais dos termos  $x_{ki}$  também converge para o mesmo valor, ou seja,

$$\mu_{ki} := \frac{1}{k} \sum_{j=0}^{k-1} x_{ji} \rightarrow x_i^*, \quad q.s.$$

Este resultado é bastante conhecido e sua demonstração pode ser vista no apêndice, Proposição A.

Portanto, para uma  $\sigma_{ki}$  que não dependa de  $x^*$ , implementamos o seguinte desvio padrão:

$$\bar{\sigma}_{ki} = |x_{ki} - \mu_{ki}|. \quad (5.4.7)$$

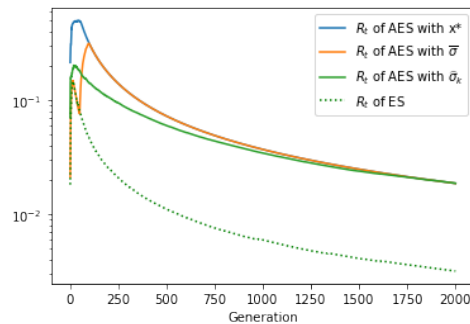
Para observar o impacto que essa adaptação tem sobre  $R_k$ , durante as primeiras 50 iterações  $\sigma_{ki}$  foi mantido constante em 1, e somente a partir desse ponto  $\bar{\sigma}_{ki}$  foi considerado conforme em (5.4.7).

Outro teste realizado foi com

$$\hat{\sigma}_{ki} = |(1 - \alpha)x_{ki} - \alpha x_{(k-1)i}|,$$

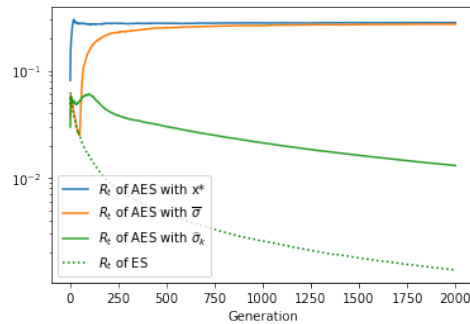
onde  $\alpha \in (0, 1)$ . Tanto  $\bar{\sigma}_{ki}$  quanto  $\hat{\sigma}_{ki}$  foram avaliados usando o EEA para minimizar a função de Ackley (Figura 5.5), a função Alpine (Figura 5.6) e  $f(x, y) = x^4 + y^4$  (Figura 5.7). As comparações foram feitas com o cenário em que o ótimo é conhecido (Teorema 5.3 ou 5.5) e com EE ( $\sigma_{ki} = 1$ ). Para esses testes numéricos, os parâmetros foram definidos da seguinte forma:  $\alpha = 0.5$ ,  $n_I = 50$ ,  $n_g = 2000$ , e 100 execuções foram realizadas para cada teste. Para o caso invariante, foi usado  $\sigma_{ki} = 1$ .

Figura 5.5: Comparação entre a TMC de diferentes mutações (Teorema 5.3,  $\bar{\sigma}$  e  $\hat{\sigma}$ ) aplicadas ao EEA e EE na resolução do Problema 5.4.1.



Fonte: Elaborada pelo autor.

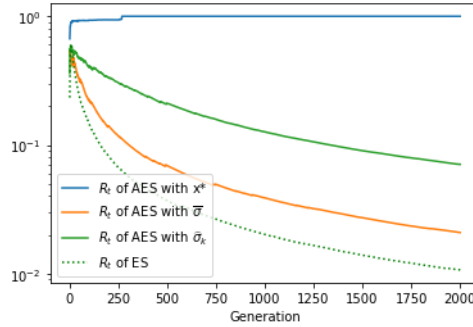
Figura 5.6: Comparação entre a TMC de diferentes mutações (Teorema 5.3,  $\bar{\sigma}$  e  $\hat{\sigma}$ ) aplicadas ao EEA e EE na resolução do Problema (5.4.3).



Fonte: Elaborada pelo autor.

Nas Figuras 5.5 e 5.6, é evidente que o uso de  $\bar{\sigma}$  resulta em uma melhoria significativa na TMC em comparação ao caso invariante, e também supera  $\hat{\sigma}$ . Além disso, rapidamente atinge os mesmos valores que o caso onde o ótimo  $x^*$  é conhecido, demonstrando sua eficácia nesses casos. Em contraste, na Figura 5.7, o uso de  $\hat{\sigma}$  resulta em uma melhor TMC do que  $\bar{\sigma}$ , embora ambos proporcionem uma melhoria substancial em relação ao caso invariante.

Figura 5.7: Comparação entre a TMC de diferentes mutações (Teorema 5.3,  $\bar{\sigma}$  e  $\hat{\sigma}$ ) aplicadas ao EEA e EE na minimização de  $f(x, y) = x^4 + y^4$ .



Fonte: Elaborada pelo autor.

### 5.4.6 PDEPV

Para que o AEA pudesse ser aplicado na resolução do PDEPV, como a função objetivo não satisfaz a condição  $\mathcal{A}_1$ , algumas adaptações se fizeram necessárias para mostrar que a TMC continua sendo linear. Nesta seção apresenta-se tais resultados e prova-se que as duas alternativas para a mutação do AEA tornam a TMC linear. Também foram feitos testes computacionais para sistemas com 3, 5 e 40 geradores.

Como descrito na Seção 2.2, o PDEPV pode ser modelado da seguinte forma:

$$\begin{aligned} \min_x \quad & \sum_{i=1}^n F_i(x_i) \\ \text{s.a:} \quad & \sum_{i=1}^n x_i = D \\ & x_i^{\min} \leq x_i \leq x_i^{\max}, \quad i = 1, \dots, n, \end{aligned} \quad (5.4.8)$$

onde

$$F_i(x_i) = a_i x_i^2 + b_i x_i + c_i + |e_i \sin(f_i(x_i^{\min} - x_i))|, \quad (5.4.9)$$

$n$  é o número de unidades geradoras no sistema,  $D$  é a demanda a ser atendida,  $x_i$  é a potência ativa gerada na unidade  $i$  e as constantes  $a_i, b_i, c_i, e_i, f_i$  são os coeficientes da função de custo  $F_i : \mathfrak{R}_i \rightarrow \mathbb{R}$ .

Assumimos que a sequência  $x_k \rightarrow x^*$  com probabilidade 1 conforme  $k \rightarrow \infty$ , e portanto temos que cada  $x_{ki} \rightarrow x_i^*$ . Consequentemente, devido à continuidade da função  $F_i$ , segue que  $F_i(x_{ki}) \rightarrow F_i(x_i^*)$  para cada  $i = 1, \dots, n$ . Assim, uma melhoria na taxa

de convergência média da sequência  $x_k$  associada ao modelo (5.4.8) pode ser alcançada ajustando o processo de mutação, levando em consideração a função  $F_i$ .

Com este objetivo em mente, e visando aplicar os dois resultados previamente obtidos (Teorema 5.3 e Corolário 5.4), examina-se inicialmente a seguinte propriedade da função  $F_i$ :

$$l_i |x_i - y_i| \leq |F_i(x_i) - F_i(y_i)| \leq L_i |x_i - y_i|, \quad (5.4.10)$$

para quaisquer  $x_i, y_i \in \mathfrak{R}_i$ , onde

$$l_i = |2a_i x_i^{\min} + b_i - 2e_i f_i|,$$

e

$$L_i = 2a_i (x_i^{\max})^2 + b_i x_i^{\max} + e_i f_i.$$

De fato, uma vez que as funções  $f(x) = \text{sen}(x)$ ,  $g(x) = |x|$ , e a função polinomial  $h(x) = \sum_{i=0}^n q_i x^i$  são Lipschitz contínuas, com constantes dadas respectivamente por 1, 1, e  $\sum_{i=0}^n |q_i| \cdot i \cdot c^{i-1}$ , onde  $c$  é o limite superior do domínio de  $h$ , e a soma e composição de funções Lipschitz permanecem Lipschitz, temos que a função  $F_i$  é Lipschitz contínua com constante  $L_i = 2a_i (x_i^{\max})^2 + b_i x_i^{\max} + e_i f_i$ . Para a primeira desigualdade em (5.4.10), observe que

$$\begin{aligned} & |F_i(x_i) - F_i(y_i)| \\ &= |a_i(x_i^2 - y_i^2) + b_i(x_i - y_i) + e_i [|\text{sen}(f_i(x_i^{\min} - x_i))| - |\text{sen}(f_i(x_i^{\min} - y_i))|]| \\ &\geq |2a_i x_i^{\min}(x_i - y_i) + b_i(x_i - y_i) - 2e_i f_i |x_i - y_i|| \\ &\geq |2a_i x_i^{\min} + b_i - 2e_i f_i| |x_i - y_i| \end{aligned}$$

No que segue, assume-se que

$$l_i = |2a_i x_i^{\min} + b_i - 2e_i f_i| \neq 0,$$

$i = 1, \dots, n$ . Com essa condição, a função  $F_i$  é denominada bi-Lipschitz.

**Corolário 5.6.** Para o Problema (5.4.8) e Algoritmo (2), se o desvio padrão da mutação para um indivíduo da  $k$ -ésima geração  $x_k = (x_{k1}, \dots, x_{kn})$  é dado por

$$\sigma_{ki} = |x_{ki} - x_i^*|,$$

então a TMC é linear e existe um  $k_0 > 0$  tal que

$$R_k \geq (1 - \rho) \prod_{i=1}^n \left[ \Phi \left( 1 + \frac{\rho l_i}{L_i} \right) - \Phi \left( 1 - \frac{\rho l_i}{L_i} \right) \right],$$

para todo  $k \geq k_0$ .

*Demonstração.* Para começar, observemos que

$$P_g(X_k; \mathcal{S}(X_k, \rho)) \geq \prod_{i=1}^n P_g(x_{ki}; \mathcal{S}_i(x_{ki}, \rho))$$

onde  $x_k = (x_{k1}, \dots, x_{kn})$  é o melhor indivíduo da população  $X_k \notin \mathcal{S}^*$ , e  $\mathcal{S}_i(x_{ki}, \rho)$  é a região promissora para  $x_{ki}$ . Sob as mesmas condições do Teorema 5.3, temos

$$\mathcal{S}_i(x_{ki}, \rho) = \{y \in \mathfrak{D}_i : |F_i(y) - F_i(x_i^*)| < \rho |F_i(x_{ki}) - F_i(x_i^*)|\}.$$

Como

$$0 < l_i \leq \frac{|F_i(x_{ki}) - F_i(x_i^*)|}{|x_i^* - x_{ki}|} \leq L_i,$$

segue que

$$\begin{aligned} P_g(x_{ki}; \mathcal{S}_i(x_{ki}, \rho)) &\geq \Phi \left( 1 + \frac{\rho |F_i(x_{ki}) - F_i(x_i^*)|}{L_i |x_i^* - x_{ki}|} \right) - \Phi \left( 1 - \frac{\rho |F_i(x_{ki}) - F_i(x_i^*)|}{L_i |x_i^* - x_{ki}|} \right) \\ &\geq \Phi \left( 1 + \frac{\rho l_i}{L_i} \right) - \Phi \left( 1 - \frac{\rho l_i}{L_i} \right) \\ &:= \mathcal{K}_\rho^i. \end{aligned}$$

Portanto,  $R_k \geq (1 - \rho) \prod_{i=1}^n \mathcal{K}_\rho^i > 0$ . □

**Corolário 5.7.** Para o Problema (5.4.8) e Algoritmo (2), se o desvio padrão da mutação para um indivíduo da  $k$ -ésima geração  $x_k = (x_{k1}, \dots, x_{kn})$  é dado por

$$\sigma_{ki} = |F_i(x_{ki}) - F_i(x_i^*)|,$$

então a TMC é linear e

$$R_k \geq (1 - \rho) \prod_{i=1}^n \left[ \Phi \left( \frac{1}{l_i} + \frac{\rho}{L_i} \right) - \Phi \left( \frac{1}{l_i} - \frac{\rho}{L_i} \right) \right].$$

*Demonstração.* Como

$$\frac{|x_i^* - x_{ki}|}{|F_i(x_{ki}) - F_i(x_i^*)|} \leq \frac{|x_i^* - x_{ki}|}{l_i |x_i^* - x_{ki}|} \leq \frac{1}{l_i},$$

similarmente à prova do Corolário 5.6 e do Teorema 5.4, temos o resultado esperado.  $\square$

Para ilustrar os resultados obtidos, o AGA e o EEA foram implementados na resolução do PDEPV com 3, 5 e 40 geradores, utilizando as duas versões de mutação (Corolários 5.6 e 5.7). Comparações foram feitas entre as duas versões e os métodos convencionais nos quais nos baseamos (AG e EE).

#### 5.4.6.1 Sistema com 3 geradores

Para esta instância foi considerada uma demanda de  $D = 201.5\text{MW}$ , população de 20 indivíduos e foram realizadas 50 execuções, cada uma delas com 100 gerações. Na Tabela 5.1 apresenta-se os valores das constantes  $l_i$  e  $L_i$ , em cada unidade geradora  $i$ .

Tabela 5.1: Constantes para o PDEPV com 3 geradores.

Gerador	1	2	3
$l_i$	6.698	6.275	2.260
$L_i$	2769.650	1952.362	2434.728

Fonte: Elaborada pelo autor.

De posse dessas constantes, pode-se determinar as estimativas para o limitante inferior da TMC, nas duas propostas de mutação. Considerando  $\rho = 0.9$ , tem-se:

$$Est_1 = 5.961 \cdot 10^{-11}$$

e

$$Est_2 = 2.490 \cdot 10^{-12}.$$

A Tabela 5.2 apresenta a melhor solução, seu custo, custo médio, desvio padrão e tempo computacional obtidos pelo AGA (nas duas versões, denotadas por AGA-Cor1 e AGA-Cor2), AG, EEA (nas duas versões, denotadas por EEA-Cor1 e EEA-Cor2) e EE. Pode-se notar que a melhor solução foi determinada pelos métodos que realizaram a mutação de forma adaptativa e estes obtiveram a mesma solução. O menor custo médio foi obtido pelo EEA-Cor1. Vale notar que a solução encontrada pelos métodos com mutação adaptativa é uma singularidade do problema.

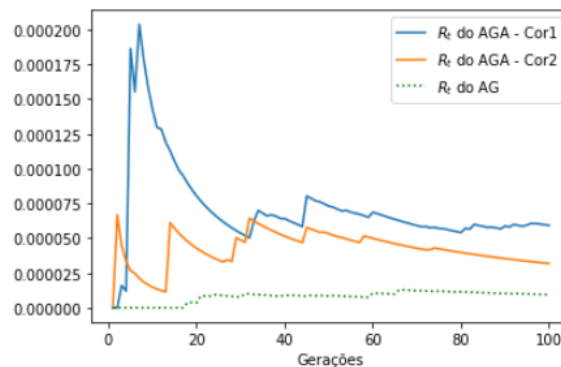
Tabela 5.2: Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 3 unidades geradoras.

	AG	AGA-Cor1	AGA-Cor2	EE	EEA-Cor1	EEA-Cor2
$P_1(MW)$	50.318	50.0	50.0	50.0	50.0	50.0
$P_2(MW)$	38.567	37.5	37.5	106.50	37.5	37.5
$P_3(MW)$	112.615	114.0	114.0	45.0	114.0	114.0
$Custo(\$ / h)$	3017.64	<b>3007.66</b>	<b>3007.66</b>	3019.09	<b>3007.66</b>	<b>3007.66</b>
Custo médio ( $\$/h$ )	3070.49	3070.29	3069.48	3070.52	<b>3062.80</b>	3065.16
$DP$	27.76	30.47	28.46	25.99	32.01	31.03
tempo (s)	0.884	0.342	0.383	0.163	0.163	0.203

Fonte: Elaborada pelo autor.

A Figura 5.8 mostra a TMC do AGA-Cor1, do AGA-Cor2 e do AG ao longo de 100 gerações. Note que o AGA-Cor1 e o AGA-Cor2 possuem uma velocidade de convergência sempre maior que o AG. Entre os métodos adaptativos, temos que o AGA-Cor1 apresenta TMC melhor que o AGA-Cor2.

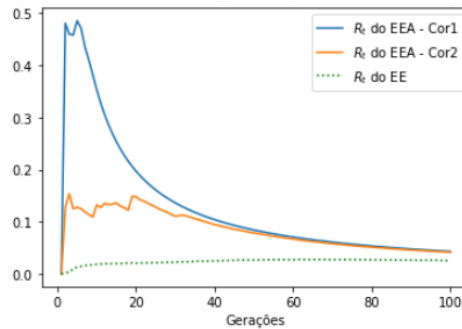
Figura 5.8: Comparação entre a TMC do AGA (Corolários 5.6 e 5.7) e AG em 100 gerações na resolução do PDEPV com 3 unidades geradoras.



Fonte: Elaborada pelo autor.

Na Figura 5.9 tem-se a TMC do EEA-Cor1, do EEA-Cor2 e do EE em 100 gerações. Neste caso, os dois métodos adaptativos apresentam uma maior TMC durante todo o processo evolutivo, mostrando que tornar o método adaptativo das formas propostas realmente melhora a velocidade de convergência do algoritmo.

Figura 5.9: Comparação entre a TMC do EEA (Corolários 5.6 e 5.7) e EE em 100 gerações na resolução do PDEPV com 3 unidades geradoras.



Fonte: Elaborada pelo autor.

#### 5.4.6.2 Sistema com 5 geradores

Para esta instância foi considerada uma demanda de  $D = 292.9\text{MW}$ , população de 30 indivíduos e foram realizadas 50 execuções, cada uma delas com 100 gerações. Na Tabela 5.3 apresenta-se os valores das constantes  $l_i$  e  $L_i$ , em cada unidade geradora  $i$ .

Tabela 5.3: Constantes para o PDEPV com 5 geradores.

Gerador	1	2	3	4	5
$l_i$	9.605	0.53	2.225	2.117	2.2
$L_i$	280.84	175.96	882.55	343.994	517.45

Fonte: Elaborada pelo autor.

Considerando  $\rho = 0.9$ , as estimativas para o limitante inferior da TMC são:

$$Est_1 = 1.065 \cdot 10^{-14}$$

e

$$Est_2 = 3.008 \cdot 10^{-16}.$$

Portanto, neste caso, a mutação proposta no Corolário 5.6 se mostra mais vantajosa do que a proposta no Corolário 5.7.

A Tabela 5.4 apresenta a melhor solução, seu custo, custo médio, desvio padrão e tempo computacional obtidos pelo AGA (nas duas versões, denotadas AGA-Cor1 e AGA-Cor2), AG, EEA (nas duas versões, denotadas EEA-Cor1 e EEA-Cor2) e EE. Pode-se notar que a melhor solução e o menor custo médio foram determinados pelo EEA-Cor1.

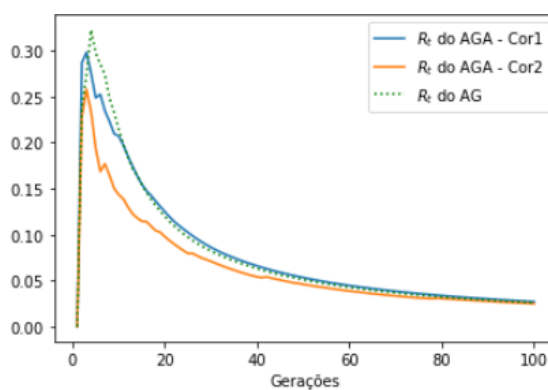
Tabela 5.4: Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 5 unidades geradoras.

	AG	AGA-Cor1	AGA-Cor2	EE	EEA-Cor1	EEA-Cor2
$P_1(MW)$	95.154	96.073	99.322	19.854	96.073	26.935
$P_2(MW)$	49.111	49.998	48.113	21.182	50.0	22.260
$P_3(MW)$	19.174	19.093	29.008	33.186	19.091	58.777
$P_4(MW)$	65.157	75.987	70.851	27.962	75.987	27.085
$P_5(MW)$	63.305	51.749	45.605	190.716	51.749	157.84
$Custo(\$ / h)$	1019.01	999.01	1022.93	1016.82	<b>998.98</b>	1008.91
Custo médio ( $\$/h$ )	1087.29	1080.46	1100.67	1096.07	<b>1071.04</b>	1081.91
$DP$	36.64	41.74	32.73	38.95	49.17	41.60
tempo (s)	0.819	0.891	0.996	0.399	0.471	0.516

Fonte: Elaborada pelo autor.

A Figura 5.10 mostra a TMC do AGA-Cor1, do AGA-Cor2 e do AG ao longo de 100 gerações. Note que os três métodos obtiveram uma TMC semelhante, porém o AGA-Cor1 foi o que obteve a melhor solução entre eles.

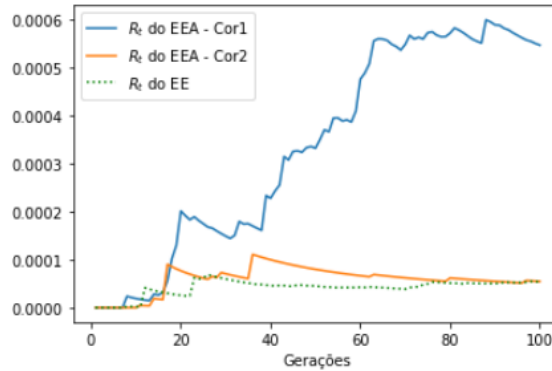
Figura 5.10: Comparação entre a TMC do EEA (Corolários 5.6 e 5.7) e EE em 100 gerações na resolução do PDEPV com 3 unidades geradoras.



Fonte: Elaborada pelo autor.

Na Figura 5.11 tem-se a TMC do EEA-Cor1, do EEA-Cor2 e do EE em 100 gerações. Neste caso, o EEA-Cor1 mostrou uma TMC bem superior aos outros dois, que obtiveram uma TMC bem semelhante no decorrer das gerações.

Figura 5.11: Comparação entre a TMC do EEA (Corolários 5.6 e 5.7) e EE em 100 gerações na resolução do PDEPV com 3 unidades geradoras.



Fonte: Elaborada pelo autor.

### 5.4.6.3 Sistema com 40 geradores

Para esta instância foi considerada uma demanda de  $D = 10500\text{MW}$ , população de 50 indivíduos e foram realizadas 100 execuções, cada uma delas com 200 gerações. As constantes  $l = (l_1, \dots, l_{40})$  e  $L = (L_1, \dots, L_{40})$  obtidas para cada uma das 40 unidades geradoras são:

$$\begin{aligned}
 l &= (9.573, 9.573, 7.296, 9.2138, 12.058, 7.197, 7.985, 8.482, 8.653, 2.327, 2.932, 2.930, \\
 &7.447, 10.280, 10.080, 10.080, 11.653, 11.673, 11.515, 11.515, 12.856, 12.856, \\
 &12.897, 12.897, 12.493, 12.493, 4.725, 4.725, 4.725, 12.058, 12.278, 12.278, \\
 &12.278, 7.832, 8.162, 8.162, 8.995, 8.995, 8.995, 11.515); \\
 L &= (954.965, 954.965, 1440.864, 2243.774, 742.715, 1583.064, 3060.000, 2991.000, \\
 &3019.800, 4967.400, 6294.3375, 6408.7125, 8365.500, 8190.500, 8125.500, \\
 &8125.500, 5560.500, 5550.500, 6287.650, 6287.650, 5459.900, 5459.900, \\
 &5391.700, 5391.700, 5591.350, 5591.350, 23964.540, 23964.540, \\
 &23964.540, 742.715, 1346.670, 1346.670, 1346.670, 1806.400, \\
 &1740.400, 1740.400, 1044.260, 1044.260, 1044.260, 6287.650).
 \end{aligned}$$

Considerando  $\rho = 0.9$ , as estimativas para o limitante inferior da TMC são:

$$Est_1 = 5.036 \cdot 10^{-120}$$

e

$$Est_2 = 4.003 \cdot 10^{-149}.$$

Assim como no PDEPV com 3 geradores, neste caso a mutação proposta no Corolário 5.6 se mostra mais vantajosa do que a proposta no Corolário 5.7.

Na Tabela 5.5 apresenta-se a melhor solução, seu custo, custo médio, desvio padrão e tempo computacional obtidos pelo EEA-Cor1, EEA-Cor2 e EE. Pode-se notar que o EEA-Cor1 apresenta a melhor solução e menor custo médio, mostrando-se melhor que os outros dois métodos utilizados na comparação.

Tabela 5.5: Comparação entre custo mínimo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 40 unidades geradoras.

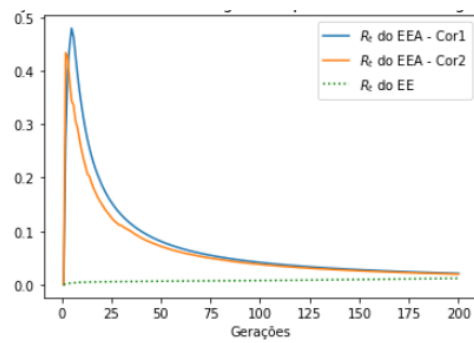
	EE	EEA (Cor1)	EEA (Cor2)
$P_1(MW)$	87.824	111.601	36.831
$P_2(MW)$	111.510	112.895	56.238
$P_3(MW)$	81.295	113.575	69.833
$P_4(MW)$	146.425	178.822	120.687
$P_5(MW)$	51.516	95.540	88.194
$P_6(MW)$	120.233	140.0	109.707
$P_7(MW)$	255.585	295.557	151.249
$P_8(MW)$	148.032	292.586	199.180
$P_9(MW)$	266.477	283.95	191.704
$P_{10}(MW)$	163.586	130.0	163.219
$P_{11}(MW)$	319.351	167.297	244.604
$P_{12}(MW)$	343.402	171.820	206.310
$P_{13}(MW)$	363.683	215.941	269.982
$P_{14}(MW)$	432.738	394.034	204.481
$P_{15}(MW)$	492.753	215.586	372.945
$P_{16}(MW)$	414.380	395.040	326.894
$P_{17}(MW)$	487.977	491.223	267.191
$P_{18}(MW)$	288.302	489.537	463.785
$P_{19}(MW)$	458.554	514.733	380.181
$P_{20}(MW)$	483.271	512.950	318.509
$P_{21}(MW)$	453.548	524.320	449.974
$P_{22}(MW)$	474.830	529.177	269.709

$P_{23}(MW)$	405.880	525.397	529.776
$P_{24}(MW)$	270.547	524.731	540.10
$P_{25}(MW)$	333.520	526.641	445.423
$P_{26}(MW)$	338.775	532.590	366.981
$P_{27}(MW)$	90.852	10.599	140.492
$P_{28}(MW)$	70.693	10.0	92.262
$P_{29}(MW)$	131.450	10.005	47.477
$P_{30}(MW)$	85.919	88.258	71.971
$P_{31}(MW)$	97.989	165.321	109.710
$P_{32}(MW)$	130.940	187.187	91.964
$P_{33}(MW)$	121.656	190.0	60.538
$P_{34}(MW)$	144.887	165.906	135.622
$P_{35}(MW)$	129.516	200.0	136.034
$P_{36}(MW)$	159.795	198.725	133.925
$P_{37}(MW)$	90.031	58.083	87.689
$P_{38}(MW)$	102.178	107.206	81.644
$P_{39}(MW)$	38.507	110.0	58.214
$P_{40}(MW)$	1311.592	513.162	2408.765
$Custo(\$ / h)$	133044.24	<b>122369.09</b>	131389.51
Custo médio ( $\$/h$ )	137631.61	<b>134387.97</b>	137366.41
$DP$	1787.54	5419.11	2032.52
tempo (s)	16.326	17.374	23.877

Na Figura 5.12 tem-se a TMC do EEA-Cor1, EEA-Cor2 e do EE. Nota-se que o desempenho dos dois métodos adaptativos é bem parecido, e muito melhor que o EE, cuja mutação é invariante.

Após os estudos relacionados ao direcionamento da busca do Algoritmo Genético às singularidades do problema e à adaptabilidade de algoritmos evolutivos, sentiu-se a necessidade de reunir essas duas características, tornando adaptativo o AGD. No próximo capítulo propõe-se então essas alterações e realiza-se testes e comparações.

Figura 5.12: Comparação entre a TMC do EEA (Corolários 5.6 e 5.7) e EE em 200 gerações na resolução do PDEPV com 40 unidades geradoras.



Fonte: Elaborada pelo autor.

## Capítulo 6

# Algoritmo Genético Direcionado Adaptativo - AGDA

A fim de potencializar os resultados alcançados tanto no direcionamento da busca do AG quanto na adaptabilidade de sua mutação, neste capítulo sugerimos a fusão dessas duas características, dando origem ao Algoritmo Genético Direcionado Adaptativo (AGDA), que incorpora a mutação adaptativa do AGA ao AGD.

Testes numéricos foram realizados na resolução do PDEPV de 3, 5 e 40 geradores. Para que fosse possível realizar comparações entre o AG, AGD, AGA (5.6) e AGDA, os dados utilizados foram os mesmos das seções anteriores.

### 6.1 Testes numéricos

Esta seção apresenta os resultados obtidos e as comparações realizadas.

#### 6.1.1 Sistema com 3 geradores

Para esta instância, dados relacionados ao custo e capacidade de geração de cada unidade geradora foram os mesmos de 4.1.1.1. Foram conduzidas 50 execuções, cada uma delas consistindo de 100 iterações. A população inicial é composta de 20 indivíduos e ao desvio padrão da mutação dos algoritmos não adaptativos foi atribuído o valor 1.

A Tabela 6.1 mostra a melhor solução obtida em cada método, seu custo, o custo médio das execuções, desvio padrão (DP) e o tempo gasto na execução que forneceu a melhor solução. De acordo com os dados apresentados nessa tabela, o AGDA encontrou a mesma solução que os outros dois métodos propostos neste trabalho e com mesmo custo médio que o AGD, porém com um tempo computacional menor.

Tabela 6.1: Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 3 unidades geradoras.

	AG	AGD	AGA	AGDA
$P_1(MW)$	50.318	50.0	50.0	50.0*
$P_2(MW)$	38.567	37.5	37.5	37.5*
$P_3(MW)$	112.615	114.0	114.0	114.0
$Custo(\$/h)$	3017.64	<b>3007.66</b>	<b>3007.66</b>	<b>3007.66</b>
Custo médio ( $\$/h$ )	3070.49	<b>3007.66</b>	3070.29	<b>3007.66</b>
$DP$	27.76	$4.55 \cdot 10^{-13}$	30.47	$4.56 \cdot 10^{-13}$
tempo (s)	0.884	0.059	0.342	0.019

Fonte: Elaborada pelo autor. Os valores marcados com asterisco (\*) representam singularidades.

### 6.1.2 Sistema com 5 geradores

Para esta instância foi considerada uma demanda de  $D = 292.9MW$ , população de 30 indivíduos e foram realizadas 50 execuções, cada uma delas com 100 gerações.

A Tabela 6.2 apresenta a melhor solução, seu custo, custo médio, desvio padrão e tempo computacional obtidos pelo AG, AGD, AGA (5.6) e AGDA. Pode-se notar que os três métodos propostos obtiveram um custo mínimo bem próximo, sendo o AGD o que obteve o menor dentre eles. Porém o AGDA foi o que apresentou o menor custo médio, desvio padrão e tempo computacional, mostrando-se um método muito bom na resolução do PDEPV e justificando a utilização de cada uma de suas características.

Tabela 6.2: Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 5 unidades geradoras.

	AG	AGD	AGA	AGDA
$P_1(MW)$	95.154	96.089	96.073	96.071*
$P_2(MW)$	49.111	50.0	49.998	50.0*
$P_3(MW)$	19.174	17.907	19.093	19.095
$P_4(MW)$	65.157	76.0	75.987	76.0*
$P_5(MW)$	63.305	49.004	51.749	51.733
$Custo(\$/h)$	1019.01	<b>976.25</b>	999.01	988.96
Custo médio ( $\$/h$ )	1087.29	990.87	1080.46	<b>988.98</b>
$DP$	36.64	19.01	41.74	0.01
tempo (s)	0.819	0.100	0.891	0.084

Fonte: Elaborada pelo autor. Os valores marcados com asterisco (\*) representam singularidades.

### 6.1.3 Sistema com 40 geradores

Para o PDEPV com 40 unidades geradoras, foram utilizados os mesmos dados de 4.1.1.4. Foram realizadas 50 execuções, cada uma delas composta por 200 iterações. Cada geração é composta por 100 indivíduos e ao desvio padrão da mutação dos algoritmos não adaptativos foi atribuído o valor 1.

A Tabela 6.3 mostra o despacho e o custo da melhor solução obtida por cada um dos métodos utilizados para comparação, além do custo médio, desvio padrão e tempo computacional. Analisando a tabela nota-se que o AGDA obteve uma solução melhor que os outros métodos utilizados na comparação. Ainda, apresentou o menor custo médio, desvio padrão e tempo computacional, comprovando mais uma vez que as modificações ao AG propostas neste trabalho realmente trazem melhorias ao método, inclusive em sistemas maiores.

Tabela 6.3: Comparação de custo, custo médio, desvio padrão (DP) e tempo de execução: sistema com 40 unidades geradoras.

	AG	AGD	AGA	AGDA
$P_1(MW)$	87.824	107.953	111.916	112.941
$P_2(MW)$	111.510	112.033	112.092	112.971

$P_3(MW)$	81.295	116.267	120.0	113.588
$P_4(MW)$	146.425	181.434	178.826	178.610
$P_5(MW)$	51.516	95.699	97.0	96.388
$P_6(MW)$	120.232	139.674	140.0	140.0*
$P_7(MW)$	255.585	299.669	295.759	256.666
$P_8(MW)$	148.032	284.546	296.547	294.408
$P_9(MW)$	266.477	287.934	289.296	283.925
$P_{10}(MW)$	163.586	201.967	130.0	130.0*
$P_{11}(MW)$	319.351	95.383	168.390	169.871
$P_{12}(MW)$	343.402	167.687	95.045	169.780
$P_{13}(MW)$	363.683	126.343	216.050	214.428*
$P_{14}(MW)$	432.738	392.205	393.144	395.151
$P_{15}(MW)$	492.753	393.087	228.574	215.605
$P_{16}(MW)$	414.380	304.287	394.80	395.213
$P_{17}(MW)$	487.977	490.808	485.508	491.381
$P_{18}(MW)$	288.302	489.101	490.086	489.554*
$P_{19}(MW)$	458.554	513.818	516.326	514.538
$P_{20}(MW)$	483.271	516.514	497.708	510.344
$P_{21}(MW)$	453.548	528.629	525.229	524.313
$P_{22}(MW)$	474.830	523.636	529.095	525.256
$P_{23}(MW)$	405.880	526.518	525.927	525.390
$P_{24}(MW)$	270.547	535.597	524.760	524.713
$P_{25}(MW)$	333.520	525.219	526.640	526.639
$P_{26}(MW)$	338.775	533.308	550.0	521.588
$P_{27}(MW)$	90.852	10.0	10.0	10.60
$P_{28}(MW)$	70.693	11.164	10.0	10.0*
$P_{29}(MW)$	131.450	10.517	10.666	10.633
$P_{30}(MW)$	85.919	87.525	86.927	88.220
$P_{31}(MW)$	97.989	189.529	173.781	165.068
$P_{32}(MW)$	130.940	190.0	187.189	187.187
$P_{33}(MW)$	121.656	162.378	190.0	190.0*

$P_{34}(MW)$	144.887	163.130	165.703	165.830
$P_{35}(MW)$	129.516	164.735	200.0	200.0*
$P_{36}(MW)$	159.795	199.010	199.467	199.402
$P_{37}(MW)$	90.031	110.0	101.924	110.0*
$P_{38}(MW)$	102.178	109.699	102.045	107.680
$P_{39}(MW)$	38.507	91.071	109.307	108.930
$P_{40}(MW)$	1311.592	511.920	514.239	513.187
$Custo(\$ / h)$	133044.24	122188.72	122679.02	<b>122070.62</b>
Custo médio ( $\$/h$ )	137631.61	122802.62	134032.77	<b>122253.64</b>
$DP$	1787.54	418.58	4266.20	65.28
tempo (s)	8.938	6.137	9.701	3.579

## Capítulo 7

# PDEPVZ com restrições de rede (PDEPVZ-RR)

Na literatura, existem algumas variantes do modelo PDEPVZ, incorporando representações adicionais de certas características do sistema elétrico. Na maioria dos modelos de PDEPVZ, o sistema de transmissão é representado principalmente por meio de suas perdas. No entanto, os limites de fluxo de potência no sistema de transmissão geralmente são negligenciados. A introdução tanto das perdas de transmissão quanto dos limites de fluxo de potência levaria a um modelo de PDEPVZ com restrição de rede (PDEPVZ-RR) que, até onde se tem conhecimento, ainda não foi explorado no contexto de abordagens meta-heurísticas. Algumas variações do modelo PDEPVZ exploram a natureza multi-objetivo do problema. Por exemplo, em ALTINOZ (2019); HUSSAIN ET AL. (2019), os autores resolvem um PDEPVZ Multiobjetivo (PDEPVZ-MO) envolvendo a minimização dos objetivos econômico e de emissões, enquanto em SRIVASTAVA & DAS (2020), o foco está em otimizar os objetivos econômico e de calor combinados. Modelos de PDEPVZ com representação de múltiplas áreas (PDEPVZ-MA) também foram descritos usando abordagens meta-heurísticas. No PDEPVZ-MA, o sistema de transmissão é dividido em áreas operacionais interconectadas, com limites de fluxo de potência introduzidos apenas nas linhas de interconexão entre essas áreas, enquanto os limites de fluxo de potência dentro de cada área não são aplicados. Abordagens meta-heurísticas específicas para resolver o PDEPVZ-MA são propostas em GHASEMI ET AL. (2016); MOHAMMADIAN ET AL. (2018).

Este capítulo apresenta um modelo linear para o sistema de transmissão no PDEPVZ-RR, no qual tanto as perdas quanto os limites de fluxo de potência são representados. A solução para este modelo deve ser aplicada a todos os indivíduos da população, não apenas aos indivíduos membros da elite. O modelo PDEPVZ-RR resultante é mais representativo do que os modelos de despacho econômico de múltiplas áreas (PDEPVZ-MA) (GHASEMI ET AL., 2016; MOHAMMADIAN ET AL., 2018), pois considera as restrições de rede para todos os ramos (linhas de transmissão) do sistema, ao contrário do PDEPVZ-MA, que considera restrições de fluxo de potência apenas nos ramos que conectam áreas vizinhas (linhas de interconexão).

## 7.1 Modelo

A introdução da representação da rede no contexto dos estudos de PDEPVZ tradicionalmente se concentrou apenas nas perdas de potência, enquanto os limites de fluxo de potência na rede muitas vezes são negligenciados. Essa simplificação é justificada pelas complexidades envolvidas na solução do PDEPVZ, que é um problema não convexo, multimodal e descontínuo. A introdução de uma representação mais detalhada da rede aumentaria significativamente os tempos computacionais necessários para resolver o problema utilizando abordagens meta-heurísticas.

No contexto de utilização de meta-heurísticas para a solução do PDEPVZ, a representação das perdas tem sido feita tradicionalmente por meio da matriz de coeficientes  $\mathbf{B}$ . Em GRAINGER & STEVENSON JR (1994), os autores mostram que a matriz de coeficientes  $\mathbf{B}$  é obtida para um ponto de carregamento operacional específico usando um algoritmo de fluxo de potência. Portanto, para pontos de carregamento que se desviam significativamente desse ponto operacional específico, a precisão da representação das perdas pode se deteriorar. Isso implica que, em situações operacionais práticas, onde os pontos de carregamento podem variar consideravelmente, os métodos que dependem da matriz de coeficientes  $\mathbf{B}$  para representar as perdas de rede podem calcular soluções que se desviam significativamente da otimalidade devido a essa imprecisão. Além disso, métodos baseados na matriz de coeficientes  $\mathbf{B}$  não consideram os limites de fluxo de potência no sistema de transmissão.

Nesta seção, propomos um modelo de sistema de transmissão capaz de representar limites nos fluxos de potência ativa na rede e permitir a representação de perdas de transmissão independentemente de um ponto de carregamento operacional específico. Embora essa forma de representação de perdas já tenha sido usada em estudos de sistemas de transmissão, não foi encontrado nenhum trabalho onde ela tenha sido aplicada no contexto de problemas de PDEPVZ. O modelo de PDEPVZ resultante da integração desta modelagem do sistema de transmissão é denominado PDEPVZ com restrição de rede (PDEPVZ-RR).

Para obter um modelo linearizado para o sistema de transmissão, inicia-se com as equações não lineares para o fluxo de potência ativa no ramo  $k - m$  dadas por:

$$P_{km} = V_k^2 g_{km} - V_k V_m g_{km} \cos(\theta_{km}) - V_k V_m b_{km} \sin(\theta_{km}), \forall k \in K, \forall m \in \Omega_k, \quad (7.1.1)$$

onde  $K$  é o conjunto de todas as barras do sistema;  $\Omega_k$  é o conjunto de barras vizinhas a  $k$ ;  $V_k$ ,  $V_m$  e  $\theta_k$ ,  $\theta_m$  são as magnitudes e ângulos de tensão nas barras  $k$  e  $m$ , respectivamente; e  $g_{km}$  e  $b_{km}$  são a condutância e susceptância série da linha  $k - m$ , respectivamente.

As seguintes aproximações são adotadas para os fluxos de potência ativa em (7.1.1) para reescrevê-los como fluxos de potência linearizados:

- (i) as perdas são inicialmente desprezadas definindo-se  $g_{km} = 0$ ;
- (ii) as magnitudes de tensão são aproximadas por  $V_k \cong V_m \cong 1$ ;
- (iii) a função seno é aproximada por  $\sin\theta_{km} \cong \theta_{km}$ ;
- (iv) a susceptância é aproximada por  $b_{km} \cong -\frac{1}{x_{km}}$ , onde  $x_{km}$  é a reatância da linha  $k - m$ .

De acordo com GRAINGER & STEVENSON JR (1994), essas aproximações são razoáveis para sistemas de transmissão, onde ocorrem altos níveis de tensão. Ao adotar essas aproximações, os fluxos de potência ativa são expressos como equações lineares dadas por:

$$P_{km} = \frac{\theta_k - \theta_m}{x_{km}}, \forall k \in K, \forall m \in \Omega_k. \quad (7.1.2)$$

A equação de balanço de potência ativa em um nó  $k$  pode ser escrita como:

$$P_k^{net} = \sum_{m \in \Omega_k} P_{km}, \quad \forall k \in K, \quad (7.1.3)$$

onde  $P_k^{net} = P_k - P_k^C$  é a potência injetada líquida na barra  $k$ , e  $P_k$  e  $P_k^C$  são a potência ativa gerada e consumida na barra  $k$ , respectivamente.

Utilizando o conjunto de equações (7.1.2) em (7.1.3) e escrevendo o balanço de potência para todas as barras  $k \in K$  (mas excluindo a barra de referência) em forma matricial resulta no sistema linear:

$$\mathbf{P}^{net} = \mathbf{B}'\theta, \quad (7.1.4)$$

onde  $\theta \in \mathbb{R}^{nb-1}$  é o vetor de ângulos de tensão, com  $nb$  sendo o número de barras do sistema;  $\mathbf{P}^{net} \in \mathbb{R}^{nb-1}$  é o vetor de injeções líquidas de potência ativa (ou seja, a potência gerada menos a consumida em cada barra); e  $\mathbf{B}' \in \mathbb{R}^{(nb-1) \times (nb-1)}$  é a matriz de susceptância nodal, cujos elementos são dados por:

$$B'_{km} = -\frac{1}{x_{km}}, \quad \forall k \in K, \forall m \in \Omega_k \quad (7.1.5)$$

e

$$B'_{kk} = \sum_{m \in \Omega_k} \frac{1}{x_{km}}, \quad \forall k \in K. \quad (7.1.6)$$

Note que as dimensões de  $\mathbf{B}'$  e os vetores em (7.1.4) correspondem a  $nb - 1$ , e não  $nb$ , porque as linhas e colunas associadas à barra de referência não estão incluídas em  $\mathbf{B}'$  e nesses vetores. Este procedimento de construção adotado para  $\mathbf{B}'$  e os vetores é necessário para garantir tanto a consistência matemática quanto física para o sistema linear (7.1.4). A consistência matemática é alcançada porque este procedimento resulta em  $\mathbf{B}'$  sendo positiva definida, e a consistência física é mantida porque uma barra de referência (referência angular) é adotada para considerar as perdas do sistema. Como resultado, a matriz  $\mathbf{B}'$  é invertível e possui posto completo, garantindo que o sistema linear (7.1.4) seja possível determinado, ou seja, sempre tem solução e é única.

Embora o sistema linear (7.1.4) imponha o balanço de potência ativa em cada barra, ele desconsidera as perdas de potência ativa. Neste trabalho, propomos uma abordagem

para incorporar as perdas de potência ativa nas restrições de balanço de potência em (7.1.4). Para atingir este objetivo, considere as perdas de potência  $p_{km}^{Loss}$  na linha  $k - m$ , representadas pela equação não linear (7.1.7), conforme descrito em GRAINGER & STEVENSON JR (1994):

$$p_{km}^{Loss} = g_{km} (V_k^2 + V_m^2 - 2V_k V_m \cos(\theta_{km})), \quad \forall k \in K, \forall m \in \Omega_k. \quad (7.1.7)$$

Adotando as aproximações (i) as magnitudes de tensão são aproximadas por  $V_k \cong V_m \cong 1$  e (ii) a função cosseno é aproximada pela expansão de sua série de potências de segunda ordem, i.e.,  $\cos(\theta_{km}) \cong 1 - \frac{\theta_{km}^2}{2}$  em (7.1.7), tem-se:

$$p_{km}^{Loss} = g_{km} \theta_{km}^2, \quad \forall k \in K, \forall m \in \Omega_k. \quad (7.1.8)$$

Para obter uma nova equação de balanço de potência que inclua as perdas de potência, primeiro aplicam-se as aproximações (i) e (ii) mencionadas acima em (7.1.1), bem como as aproximações anteriormente usadas para a função senoidal, ou seja,  $\sin \theta_{km} \cong \theta_{km}$ , e para a susceptância série, ou seja,  $b_{km} \cong -\frac{1}{x_{km}}$ . Finalmente, substituímos o  $P_{km}$  obtido dessas aproximações em (7.1.3), resultando em:

$$P_k^{net} - g_{km} \frac{\theta_{km}^2}{2} = \frac{\theta_{km}}{x_{km}}, \quad \forall k \in K, \forall m \in \Omega_k. \quad (7.1.9)$$

Observe que metade das perdas quadráticas para a linha  $k - m$  dadas em (7.1.8) aparecem no lado esquerdo de (7.1.9). Esta expressão sugere que metade das perdas na linha  $k - m$  são tratadas como uma carga (devido ao sinal negativo) na barra  $k$ , enquanto a outra metade aparece como uma carga na barra terminal  $m$ . Portanto, as perdas em cada linha podem ser tratadas como cargas adicionais conectadas nas barras terminais da linha. Finalmente, a expressão (7.1.9) pode ser reinterpretada em forma matricial, para que as perdas de potência, formuladas como cargas adicionais, sejam incluídas da seguinte forma:

$$\mathbf{P}^{net} - \mathbf{P}^{Loss} = \mathbf{B}'\theta, \quad (7.1.10)$$

onde  $\mathbf{P}^{Loss}$  é o vetor composto por todas as perdas acumuladas como cargas adicionais

em cada barra  $k$ , conforme descrito no procedimento acima.

Agora, podemos reformular o problema de PDEPVZ para incluir a representação do sistema de transmissão e descrever o problema de PDEPVZ com restrição de rede (PDEPVZ-RR):

$$\begin{aligned}
 \min_{\mathbf{P}, \theta} \quad & \sum_{i=1}^n C_i(P_i) & (a) \\
 \text{s.a. :} \quad & & \\
 & \mathbf{P}^{\text{net}} - \mathbf{P}^{\text{Loss}} = \mathbf{B}'\theta & (b) \quad (7.1.11) \\
 & -P_{km}^{\text{max}} \leq \frac{\theta_k - \theta_m}{x_{km}} \leq P_{km}^{\text{max}}, \forall k \in K, \forall m \in \Omega_k & (c) \\
 & P_i \in \bigcup_{p=0}^{z_i} [P_{i,p}^u, P_{i,p+1}^l], \forall i = 1, \dots, n. & (d)
 \end{aligned}$$

No problema PDEPVZ-RR descrito em (7.1.11), as restrições de balanço de potência são consideradas para cada barra do sistema, por meio das restrições (7.1.11)(b), que implicitamente representam as perdas do sistema. O conjunto de restrições (7.1.11)(c) representa os limites nos fluxos de potência ativa para cada linha do sistema, enquanto as zonas de operação proibidas são descritas em (7.1.11)(d).

Vale notar que todas as hipóteses assumidas na prova de convergência (Capítulo 4) permanecem válidas para o PDEPVZ-RR, tornando todos os resultados válidos também neste problema.

Um novo operador genético, chamado operador de balanço de potência, é necessário especificamente para calcular as perdas do sistema dadas por  $\mathbf{P}^{\text{Loss}}$  e impor as restrições de balanço de potência dadas em (7.1.11)(b). O operador de balanço de potência é introduzido na abordagem proposta do AGD sempre que o algoritmo se destina a resolver o PDEPVZ-RR. Este operador também permite a solução do PDEPVZ-RR com pequenos acréscimos no tempo computacional. A descrição detalhada do operador de balanço de potência é fornecida na próxima seção.

## 7.2 Adaptações ao AGD para resolução do PDEPVZ-RR

Assim como no PDEPVZ, no caso do PDEPVZ-RR existe uma estreita relação entre as soluções do problema e suas singularidades. Com o intuito de aproveitar essa característica para melhorar a busca por soluções ótimas do problema, nesta seção propõem-se modificações ao AGD para que este possa ser utilizado na resolução do PDEPVZ-RR. As principais mudanças estão relacionadas à criação do operador de balanço de potência, que será responsável por calcular as perdas do sistema e realizar o balanço de potências, e à função avaliação, que deverá penalizar os indivíduos inactivos com relação ao fluxo de potência e à potência atribuída à barra de referência. Os outros operadores permanecem como descrito em 3.1. Nas seções a seguir, o pseudocódigo revisado do AGD, bem como os novos operadores genéticos utilizados para resolver o PDEPVZ-RR são descritos em detalhes.

### 7.2.1 Pseudocódigo

A busca direcionada do AGD é refletida na criação de indivíduos e na inclusão dos indivíduos pivôs. Na solução do PDEPVZ-RR, observou-se que a taxa de convergência do algoritmo é consideravelmente melhorada quando se utiliza um conjunto de indivíduos pivôs, ao invés de se utilizar apenas um único indivíduo pivô, conforme discutido anteriormente, para a solução do PDEPVZ. O pseudocódigo do AGD revisado para a resolução do PDEPVZ-RR é apresentado no Algoritmo 5, onde os operadores modificados são descritos nesta seção.

### 7.2.2 Operador de Balanço de Potência

O operador de balanço de potência é um operador genético específico proposto para garantir que cada indivíduo  $I^j$  na população tenha um despacho de potência que atenda às restrições de balanço de potência, assegurando que nenhum indivíduo inactivo seja gerado na população.

Para garantir que  $I^j$  seja um indivíduo activo no que se refere às restrições de

---

**Algoritmo 5** Pseudocódigo do AGD
 

---

- 1: Inicialize os parâmetros:  $t = 0$ ,  $n_I$  e  $p_M$ ;
  - 2: Determine todas as singularidades factíveis de cada gerador do sistema;
  - 3: Gere a população inicial  $POP$ ;
  - 4: Gere os indivíduo pivôs e acrescente à  $POP$ ;
  - 5: Aplique o operador de balanço de potência à  $POP$ , de acordo com 7.2.2;
  - 6: Avalie cada indivíduo de  $POP$  de acordo com Seção 7.2.3;
  - 7: **while** condição de parada não é satisfeita **do**
  - 8:     **for**  $i = 1, \dots, \frac{n_I}{2}$  **do**
  - 9:         Selecione dois indivíduos de  $POP$ ;
  - 10:         Aplique o crossover;
  - 11:         Aplique a mutação com probabilidade  $p_M$ ;
  - 12:         Aplique o elitismo;
  - 13:         Aplique o operador de balanço de potência a  $POP$ ;
  - 14:     **end for**
  - 15:     Faça a mutação do indivíduo pivô  $I_p$ , compare com o pivô anterior e acrescente o melhor deles em  $POP$ ;
  - 16:     Aplique o operador de balanço de potência a  $I_p$ ;
  - 17:     Avalie cada indivíduo de  $POP$  de acordo com Seção 7.2.3;
  - 18:      $t = t + 1$
  - 19: **end while**
  - 20: Retorne o melhor indivíduo de  $POP$  como solução.
- 

balanço de potência do PDEPVZ-RR, um novo algoritmo é necessário para calcular as perdas e impor o balanço nodal de potência, como mostrado no Algoritmo 6.

---

**Algoritmo 6** Pseudocódigo para o operador de balanço de potência para o problema PDEPVZ-RR
 

---

Utilizando as potências geradas fornecidas por  $I^j$ , obtenha o vetor  $\mathbf{P}^{\text{net}}$ , que é calculado como a potência gerada menos a demanda em cada barra.

Resolva  $\mathbf{P}^{\text{net}} = \mathbf{B}'\hat{\theta}$ , obtendo  $\hat{\theta}$ .

Use  $\hat{\theta}$  em (7.1.8) para calcular as perdas  $p_{km}^{\text{Loss}}$ .

Calcule  $\mathbf{P}^{\text{Loss}}$  alocando incrementalmente metade de  $p_{km}^{\text{Loss}}$  para cada barra terminal da linha  $k - m, \forall k \in K, \forall m \in \Omega_k$ .

Resolva  $\mathbf{P}^{\text{net}} - \mathbf{P}^{\text{Loss}} = \mathbf{B}'\theta$ , obtendo  $\theta$ .

Use  $\theta$  em (7.1.8) para calcular as perdas finais  $p_{km}^{\text{Loss}}$ .

Atualize a potência correspondente à barra de referência (*slack*) em  $I^j$  para garantir o cumprimento da restrição de balanço de potência naquela barra.

Retorne  $I^j$  e as perdas  $p_{km}^{\text{Loss}}$ .

---

É importante enfatizar alguns aspectos-chave deste algoritmo: (i) Os dois sistemas lineares resolvidos sempre fornecem uma solução única, pois a matriz  $\mathbf{B}'$  é de posto completo e definida positiva. (ii) Diferentemente da matriz de coeficientes  $\mathbf{B}$  (geralmente usada para a aproximação de perdas de potência), a matriz  $\mathbf{B}'$  não depende de um ponto operacional específico. Portanto, as perdas calculadas não dependem de um ponto de carregamento específico. (iii) A matriz  $\mathbf{B}'$  é constante para todos os indivíduos da população, permitindo cálculos eficientes usando multiplicação matriz-vetor com os fatores da matriz computados apenas uma vez; (iv) A solução é extremamente rápida, permitindo sua aplicação a todos os indivíduos da população com tempo de computação adicional razoavelmente reduzidos.

### 7.2.3 Função de Avaliação

Cada indivíduo  $I^j$  deve ser avaliado de maneira que permita estabelecer parâmetros de adaptabilidade (ou *fitness*) para selecionar os indivíduos. Para o AGD na solução do PDEPVZ-RR, um indivíduo tem uma boa avaliação de *fitness* quando: i) seus custos de produção correspondentes são baixos; ii) seus fluxos de potência ativa na rede estão dentro dos limites estabelecidos e iii) os limites na geração de potência na barra *slack* também estão dentro dos limites estabelecidos.

Para avaliar os itens ii) e iii), utilizam-se as funções auxiliares (7.2.1)-(7.2.3) para ajudarem na formulação da função avaliação. A equação

$$g_{km}^F(\theta) = \left| \frac{\theta_k - \theta_m}{x_{km}} \right| - P_{km}^{\max}, \quad \forall k \in K, \forall m \in \Omega_k, \quad (7.2.1)$$

representa os limites de cada linha  $k - m$  do sistema, enquanto equações

$$g_s^{Pmax}(P_s) = P_s - P_s^{\max} \quad (7.2.2)$$

e

$$g_s^{Pmin}(P_s) = P_s^{\min} - P_s \quad (7.2.3)$$

representam os limites máximos e mínimos das potências de saída  $P_s$  na barra *slack*  $k = s$ .

Usando as funções auxiliares mencionadas acima e tratando-se os limites associados

por meio da abordagem de penalização, uma função de custo estendida é definida por

$$C^A(I^j) = \sum_{i=1}^n C_i(P_i^j) + \sum_{k \in K, m \in \Omega_k} \mu [\max\{0, g_{km}^F(\theta)\}]^2 \quad (7.2.4)$$

$$+ \nu \left[ (\max\{0, g_s^{Pmax}(P_s)\})^2 + (\max\{0, g_s^{Pmin}(P_s)\})^2 \right],$$

onde  $\mu$  e  $\nu$  são parâmetros de penalidade.

Note que, ao usar o operador max em (7.2.4), uma restrição associada é introduzida na função de custo estendida apenas quando ela viola seu limite associado. Finalmente, a função de avaliação para o PDEPVZ-RR é definida como

$$F^A(I^j) = (C^A(I^j))^{-1}. \quad (7.2.5)$$

Os resultados numéricos associados à solução do PDEPVZ-RR são discutidos na seção a seguir, para os sistemas testes do IEEE de 118 e 300 barras.

### 7.3 Testes numéricos

O Algoritmo 5 foi implementado em Python, e testes numéricos foram realizados para resolver o PDEPVZ-RR para o sistema IEEE-118 barras e para resolver o PDEPV-RR para o sistema IEEE-300 barras.

Neste capítulo, todos os parâmetros do AGD foram ajustados com o auxílio do *software Optuna*. Para fins de comparação, assim como em testes anteriores, os seguintes métodos foram implementados: o Algoritmo Genético clássico com codificação real (AG), Evolução Diferencial Colaborativa Otimizada (EDOC), Otimização por Enxame de Partículas (PSO), Estratégia Evolutiva de Adaptação de Matriz de Covariância (EEAMC), Evolução Diferencial Adaptativa Caótica com Arquivo Externo Opcional (CJADE) e Estratégia Evolutiva de Amostragem Híbrida (HSES).

Os dados para os sistemas elétricos foram obtidos do banco de dados do pacote *pandapower*, incluindo os coeficientes das funções de custo  $a_i$ ,  $b_i$  e  $c_i$ . Os termos  $e_i$  e  $f_i$ , relacionados aos efeitos de pontos de carregamento de válvula, foram estimados conforme proposto em PINHEIRO ET AL. (2022). As zonas de operação proibidas foram obtidas de POURAKBARI-KASMAEI & MANTOVANI (2018).

Na resolução do PDEPVZ-RR, para o CJADE, utilizou-se  $p = 0,05$  e  $d = 0,1$ ; para o HSES,  $hmcr = 0,9$ ,  $par = 0,3$  e  $bw = 0,01$ ; e para o PSO, definiu-se  $c_1 = c_2 = 1,49618$  e  $w = 0,7298$ . Os demais parâmetros foram definidos com os mesmos valores determinados pelo *Optuna* para o AGD.

### 7.3.1 Sistema IEEE-118 barras

Para esta instância, a barra 88 é definida como a barra de referência (*slack*), e 10 indivíduos pivôs foram criados. Foram realizadas 25 execuções, cada uma consistindo de 100 iterações. A população inicial foi composta por 84 indivíduos, com uma taxa de mutação de 0,1127, uma taxa de cruzamento de 0,896, e  $\sigma_i = 1$ . Para a penalização da função de aptidão, foi utilizado um peso de  $\mu = \nu = 1.5458$ .

A Tabela 7.1 apresenta os resultados comparativos obtidos entre todos os métodos testados, incluindo AGD, CJADE, HSES, AG, EDOC, EEAMC e PSO. Esta tabela mostra a melhor solução encontrada, o custo em  $\$/h$ , os valores médios de custo, o desvio padrão e o tempo de execução para cada solução. Observa-se que a melhor solução é determinada pelo AGD, que, além de ser factível, possui várias entradas posicionadas em singularidades (os valores de potência de saída localizados em pontos de singularidade de cada gerador são destacados na tabela por meio do asterisco), justificando a técnica proposta neste trabalho para direcionar a busca para as singularidades do problema.

Tabela 7.1: Comparação de custo mínimo, custo médio, desvio padrão (DP) e tempo de execução: sistema IEEE 118 barras.

barra	CJADE	HSES	AG	EDOC	EEAMC	PSO	AGD
0	58.00	21.00	71.20	52.52	298.10	52.80	85.00*
3	31.00	92.00	85.00	99.39	78.00	61.10	20.00*
5	52.00	35.00	48.00	96.27	96.60	32.20	20.00*
7	35.00	17.00	42.00	25.04	-52.30	319.00	50.00*
9	303.00	20.00	22.00	74.31	277.80	72.50	410.00*
11	10.00	166.00	366.00	278.41	76.00	24.70	93.00
14	34.00	59.00	121.00	95.09	61.20	11.30	60.00
17	12.00	25.00	57.00	99.44	4.50	7.10	25.00
18	67.00	18.00	7.00	74.85	103.20	29.50	100.00*
23	19.00	35.00	48.00	49.86	1.90	174.30	25.00*
24	5.00	43.00	27.00	24.71	161.40	188.10	160.00
25	405.00	264.00	189.00	99.25	200.20	0.00	207.00
26	76.00	112.00	197.00	99.60	-51.30	35.50	75.00*
30	31.00	14.00	0.00	99.95	1.80	34.10	0.00*

31	90.00	78.00	0.00	0.05	189.10	75.70	25.00
33	92.00	62.00	0.00	74.96	27.50	40.90	20.00
35	64.00	34.00	0.00	75.51	121.20	53.50	0.00*
39	26.00	73.00	44.00	77.16	103.80	62.40	100.00*
41	94.00	86.00	0.00	49.20	157.20	57.90	55.00*
45	31.00	60.00	27.00	49.59	33.10	65.30	30.00*
48	45.00	106.00	3.00	28.29	225.80	62.00	152.00
53	1.00	103.00	148.00	98.89	53.40	12.50	37.00
54	30.00	112.00	100.00	74.50	-7.40	49.40	0.00*
55	22.00	74.00	0.00	75.23	242.70	79.50	100.00*
58	68.00	87.00	0.00	50.51	64.20	154.80	95.00
60	93.00	126.00	260.00	99.90	80.60	33.40	130.00
61	45.00	256.00	100.00	99.99	-141.20	200.50	50.00
64	129.00	45.00	71.00	24.94	21.20	312.60	368.00
65	240.00	306.00	280.00	99.69	257.80	28.40	246.00*
69	29.00	36.00	362.00	99.18	-95.60	66.30	0.00*
71	19.00	39.00	34.00	25.26	-123.10	49.20	25.00
72	61.00	100.00	0.00	55.78	89.00	81.80	25.00*
73	10.00	56.00	0.00	72.47	97.80	59.30	50.00*
75	61.00	94.00	0.00	51.80	166.90	51.10	0.00*
76	87.00	65.00	21.00	99.89	45.30	378.00	0.00*
79	361.00	76.00	18.00	99.82	126.90	40.30	433.00
84	83.00	331.00	311.00	99.98	23.60	61.00	20.00
86	16.00	28.00	10.00	96.23	0.00	35.10	0.00*
88	727.00	97.00	3.00	4.95	565.00	37.26	538.00
89	45.00	90.00	11.00	74.90	28.90	36.30	25.00*
90	65.00	78.00	23.00	69.04	-25.9	32.20	0.00*
91	4.00	67.00	19.00	74.33	86.00	45.40	25.00*
98	77.00	50.00	6.00	99.65	94.30	21.60	75.00*
99	19.00	100.00	161.00	97.00	260.50	268.00	88.00*
102	11.00	88.00	0.00	39.62	32.10	80.60	35.00*
103	25.00	47.00	100.00	24.89	37.10	26.50	75.00*
104	56.00	55.00	28.00	25.20	-13.40	61.50	25.00*
106	13.00	37.00	0.00	21.61	-50.50	48.00	0.00*
109	73.00	47.00	0.00	27.03	72.60	45.50	25.00*
110	14.00	30.00	20.00	32.33	40.50	63.40	68.00*
111	89.00	56.00	5.00	78.51	-52.00	67.60	0.00*
112	84.00	59.00	81.00	76.79	91.60	45.80	5.00
115	4.00	78.00	74.00	75.90	203.30	30.80	100.00*
Custo (\$/h)	150553.01	167429.29	179910.01	181714.88	148222.31	161606.65	<b>141315.84</b>
Custo médio (\$/h)	156792.32	168644.13	188814.19	194882.82	149550.02	168277.44	148228.82
DP	4109.03	1893.82	9458.16	8934.39	741.54	8650.11	8273.65
tempo (s)	432.9	1451.71	917.93	1181.72	713.43	768.48	807.89

<sup>[0]</sup> Os valores marcados com asterisco (\*) representam singularidades.

### 7.3.2 Sistema IEEE-300 barras

Para esta instância, a barra 97 é definida como a barra de referência (*slack*), e 15 indivíduos pivôs foram criados. Foram realizadas 25 execuções, cada uma consistindo de 200 iterações. A população inicial foi composta por 100 indivíduos, com uma taxa de mutação de 0.1033, uma taxa de cruzamento de 0.899 e  $\sigma_i = 1.5594$ . Para a penalização da função de aptidão, valores de penalidade de  $\mu = \nu = 1.344$ .

A Tabela 7.2 apresenta os resultados comparativos obtidos entre todos os métodos testados. Ela mostra a melhor solução encontrada, o custo em  $\$/h$ , os valores médios dos custos, o desvio padrão e o tempo de execução para esta solução. Observa-se que a melhor solução foi alcançada pelo AGD, seguido pelo PSO. Similarmente aos resultados do sistema de 118 barras, várias entradas estão posicionadas em singularidades, indicadas por um asterisco. Isso reforça a vantagem de realizar a busca de forma direcionada às singularidades do problema.

Tabela 7.2: Comparação de custo mínimo, custo médio, desvio padrão (DP) e tempo de execução: sistema IEEE 300 barras.

barra	CJADE	HSES	AG	EDOC	EEAMC	PSO	AGD
7	43.90	84.34	25.04	65.43	440.60	89.50	100.00*
9	65.90	75.42	79.90	23.2	445.00	95.90	100.00*
18	78.20	8.93	96.14	1.44	549.40	40.50	100.00*
54	47.60	27.30	100.00	76.44	903.90	34.00	100.00*
62	59.30	67.40	100.00	62.37	39.10	58.90	0.00*
68	107.50	326.87	37.73	362.09	6.20	354.50	475.00*
75	164.10	93.67	473.37	9.93	42.90	196.50	191.25*
76	269.10	222.34	76.50	326.48	122.40	225.60	390.00*
79	165.20	140.65	168.00	104.08	-46.00	115.50	84.00*
87	206.10	161.56	100.31	122.26	163.20	72.80	217.00*
97	5790.30	7724.43	213.69	5782.33	2548.20	4110.80	4407.70
102	305.10	82.39	340.00	212.37	243.40	294.70	340.00
103	47.10	90.62	0.00	86.44	515.60	657.40	100.00*
116	16.50	5.30	0.00	58.69	36.80	69.00	100.00
119	216.70	381.0	275.46	221.34	-8.80	299.50	190.50
121	9.80	90.71	395.18	788.58	-166.30	517.20	381.00
124	96.40	138.19	77.33	139.24	-105.50	105.30	138.00*
125	280.40	132.86	18.02	140.21	-113.80	125.00	158.50*
127	119.70	203.0	37.37	143.6	-48.80	103.90	101.50*
130	351.10	307.08	222.20	183.72	-119.40	270.60	118.00*
131	56.50	42.86	278.93	72.32	69.20	93.70	237.00*
134	71.10	99.97	33.73	59.69	421.40	50.40	0.00*
148	193.20	50.42	304.57	283.1	291.60	159.30	100.00

149	88.10	100.0	79.83	92.91	-186.60	81.00	100.00*
154	6.90	113.99	200.15	112.14	-168.10	169.80	100.00
155	5.00	137.95	95.38	97.09	-77.00	155.60	184.00*
163	117.10	295.38	264.02	63.49	135.50	270.70	150.00*
164	1041.10	1196.04	1260.80	1294.32	275.50	1002.20	650.00*
165	315.40	723.53	215.48	1275.45	158.50	1126.70	325.00*
168	408.00	575.0	575.00	373.47	852.50	590.10	575.00*
169	2045.70	1605.76	1357.03	1667.55	326.30	1645.30	575.00
176	362.90	3.48	524.00	374.71	652.80	484.50	524.00*
191	368.40	2.5	372.00	220.65	176.50	542.50	0.00*
198	138.60	200.0	174.68	10.67	51.30	220.70	200.00*
199	514.50	168.46	373.87	250.12	560.70	431.00	200.00
200	247.40	87.28	183.97	236.18	351.30	38.00	350.00*
205	277.20	261.28	306.51	273.05	412.00	295.80	350.00
208	331.90	344.96	384.17	166.04	474.20	335.60	403.00*
211	80.50	395.67	400.00	376.68	391.90	260.30	400.00
214	544.00	241.38	555.68	117.15	350.80	568.70	0.00*
216	236.20	350.0	340.03	175.13	363.90	197.40	305.45
217	492.20	191.44	517.89	449.07	803.20	519.70	350.00*
219	600.20	112.21	460.15	470.87	-177.90	364.60	675.43
220	228.60	171.72	270.00	175.13	262.80	212.20	270.00
221	148.30	1.84	66.90	50.65	129.90	191.00	184.00
246	170.40	567.0	464.11	88.57	707.10	245.70	184.00*
247	630.20	723.0	723.00	459.78	736.50	148.90	567.00*
248	972.70	237.72	1310.00	785.09	1658.10	430.00	1310.00
249	135.90	334.0	334.00	59.9	332.60	258.80	334.00
250	141.10	55.9	0.00	332.83	494.50	434.10	236.00*
251	416.20	185.22	430.00	195.72	539.70	317.60	107.50*
252	243.40	256.0	279.97	190.31	281.50	153.90	71.25*
253	444.80	332.88	350.69	427.06	619.90	277.80	285.00
254	302.70	88.64	600.00	101.84	753.80	424.30	510.00
255	114.50	46.42	0.00	2.62	60.00	81.30	68.50*
257	0.30	73.97	145.00	112.97	70.10	167.50	137.00
258	22.60	265.0	265.00	161.57	214.60	89.00	265.00*
259	5.30	463.19	500.00	349.82	532.40	432.20	0.00*
260	465.50	5.11	500.00	419.08	502.20	332.40	250.00*
261	194.80	216.0	204.25	173.41	161.90	191.40	216.00*
262	1316.30	1362.64	1301.25	1337.6	1745.70	1023.30	1392.00*
263	194.80	101.71	800.00	428.75	814.00	145.30	80.00
264	456.20	35.1	465.10	318.99	5.50	543.40	489.75*
266	85.90	61.34	64.59	40.19	451.10	85.70	100.00*
291	12.30	86.41	100.00	15.37	288.20	89.70	100.00*
293	14.90	9.6	100.00	57.26	453.70	92.00	50.00*
294	147.30	105.4	0.00	41.55	67.80	22.70	100.00
295	0.50	105.4	0.00	99.5	7.20	16.70	27.00*
Custo (\$/h)	854840.00	967974.58	851199.48	851786.79	872782.42	798198.62	<b>765020.49</b>
Custo médio (\$/h)	887894.16	1023986.44	881184.26	890057.03	879342.14	862974.97	768087.40

<i>DP</i>	28280.28	38653.63	20890.31	27847.37	4048.71	28158.17	31015.03
tempo (s)	3652.93	1971.85	2474.96	1946.51	2142.35	2141.50	1648.69

Para avaliar o desempenho do AGD em comparação com os métodos selecionados, utilizou-se o teste de postos sinalizados de Wilcoxon para amostras pareadas. Este teste não paramétrico é adequado para identificar diferenças nas tendências centrais de dados pareados (para mais detalhes, ver DERRAC ET AL. (2011)). Foram comparados o custo mínimo, o custo médio e o tempo computacional de cada algoritmo nas cinco instâncias consideradas neste trabalho. Os resultados, apresentados na Tabela 7.3, para um nível de significância de 0,05, indicam se há uma diferença estatisticamente significativa entre o AGD e os outros métodos. Como mostrado, o AGD, em comparação com os métodos selecionados, demonstrou ser competitivo em termos de alcançar o custo mínimo, o custo médio e o tempo computacional.

Métrica	Algoritmo Comparado	Valor-p	Significância
Custo Mínimo	CJADE	0.0625	Não
Custo Mínimo	AG	0.0625	Não
Custo Mínimo	PSO	0.0625	Não
Custo Mínimo	EEAMC	0.0625	Não
Custo Mínimo	EDOC	0.0625	Não
Custo Mínimo	HSES	0.0625	Não
Custo Médio	CJADE	0.0625	Não
Custo Médio	AG	0.0625	Não
Custo Médio	PSO	0.0625	Não
Custo Médio	EEAMC	0.0625	Não
Custo Médio	EDOC	0.0625	Não
Custo Médio	HSES	0.0625	Não
Tempo Computacional	CJADE	0.3125	Não
Tempo Computacional	AG	0.125	Não
Tempo Computacional	PSO	1	Não
Tempo Computacional	EEAMC	0.3125	Não
Tempo Computacional	EDOC	0.125	Não
Tempo Computacional	HSES	0.0625	Não

Tabela 7.3: Resultados do teste de Wilcoxon comparando o AGD com outros algoritmos em diferentes instâncias, com base no custo mínimo, custo médio e tempo computacional.

Além disso, para possibilitar uma comparação abrangente do desempenho de todos os métodos, aplicou-se o teste de Friedman para o custo mínimo, uma técnica não paramétrica comumente usada para detectar diferenças estatisticamente significativas entre vários métodos em experimentos com medidas repetidas. O teste de Friedman resultou em uma estatística de 28,0 e um valor p de 0,00001, indicando que há evidências estatísticas para sugerir uma diferença significativa entre pelo menos um par de métodos. Com base neste resultado, utilizou-se o teste de Nemenyi para uma análise *post-hoc* a fim

<sup>[0]</sup> Os valores marcados com asterisco (\*) representam singularidades.

de identificar especificamente quais métodos diferiram significativamente dos outros. A Tabela 7.4 fornece uma visualização detalhada dos resultados do teste de Nemenyi, onde se pode observar que há uma diferença estatisticamente significativa entre o AGD e o EDOC (valor  $p = 0.017163 < 0.05 = \alpha$ ). Dado que o custo obtido pelo AGD é melhor do que o apresentado pelo EDOC, podemos concluir que o AGD é superior ao EDOC em termos dos custos mínimos alcançados. Em relação aos demais métodos, nenhuma diferença significativa foi detectada por este teste.

	AGD	CJADE	AA	PSO	EEAMC	EDOC	HSES
AGD	1.000000	0.137562	0.191774	0.738735	0.191774	<b>0.017163</b>	0.096003
CJADE	0.137562	1.000000	0.900000	0.900000	0.900000	0.900000	0.900000
AG	0.191774	0.900000	1.000000	0.900000	0.900000	0.900000	0.900000
PSO	0.738735	0.900000	0.900000	1.000000	0.900000	0.523543	0.867850
EEAMC	0.191774	0.900000	0.900000	0.900000	1.000000	0.900000	0.900000
EDOC	<b>0.017163</b>	0.900000	0.900000	0.523543	0.900000	1.000000	0.900000
HSES	0.096003	0.900000	0.900000	0.867850	0.900000	0.900000	1.000000

Tabela 7.4: Resultados do Teste de Nemenyi para todas as comparações de métodos.

Com os resultados obtidos pelo AGD adaptado para a resolução do PDEPVZ-RR e os testes estatísticos realizados, é possível afirmar que o AGD apresentou um desempenho excelente e mostrou-se competitivo em comparação com os demais métodos escolhidos.

É importante destacar que, até onde se tem conhecimento, a representação das perdas nos sistemas de transmissão não foi discutida na literatura no contexto de solução do PDEPVZ-RR por meio de meta-heurísticas, de forma que todos os indivíduos da população possuem uma representação da rede.

Vale também salientar que o PDEPVZ com restrições de rede não linear já foi abordado através de algoritmos de otimização convencionais com formulações de programação não linear inteira mista, como as apresentadas em POURAKBARI-KASMAEI & MANTOVANI (2018) e POURAKBARI-KASMAEI ET AL. (2019) para sistemas de até 118 barras. No entanto, as soluções desses algoritmos podem ficar presas em ótimos locais. Essa é uma questão particularmente crítica para o PDEPVZ-RR, dado que o problema é multimodal e pode conter centenas de milhares de ótimos locais.

Portanto, o algoritmo proposto para o PDEPVZ-RR é capaz de representar a transmissão, ainda que de forma linear, e de explorar o espaço de busca de maneira mais eficiente, evitando com maior facilidade a convergência para soluções locais. Além disso, o modelo proposto foi capaz de resolver sistemas de até 300 barras, o que demonstra sua robustez e aplicabilidade em problemas de maior escala.

## Capítulo 8

# Considerações finais e perspectivas futuras

O Problema de Despacho Econômico com Efeito de Pontos de Carregamento de Válvula e Zonas Proibidas (PDEPVZ) e o PDEPVZ com Restrição de Rede (PDEPVZ-RR) apresentam dificuldades para serem resolvidos por métodos determinísticos, uma vez que são problemas multimodais, fazendo com que estes métodos convirjam com frequência para mínimos locais, e cujas funções a serem minimizadas não são diferenciáveis em todo domínio. Por outro lado, mesmo as meta-heurísticas (como o AG) apresentando melhores resultados na exploração do espaço de busca, podem apresentar lentidão devido à dimensão da instância a ser resolvida. Buscando diminuir o tempo computacional por meio da redução do espaço de busca e aproveitar características específicas destes problemas, como o conhecimento de que suas soluções estão localizadas em singularidades, mas ainda mantendo as características do AG, neste trabalho foi proposto o AGD, um método desenvolvido especificamente para abordar esses problemas e capaz de melhorar o desempenho do algoritmo genético ao reduzir o espaço de busca e introduzir novos operadores, como o elitismo e o indivíduo pivô.

Além disso, pela necessidade de provas formais de convergência para algoritmos que visam a resolução destes problemas, utilizou-se a teoria de análise estocástica para interpretar formalmente a dinâmica do espaço de busca e provou-se a convergência do AGD. As provas da Proposição 4.22 e do Teorema 4.24 destacam a importância dos operadores: elitismo, mutação e do indivíduo pivô na garantia de convergência do algoritmo.

Ademais, duas variações de mutação positiva adaptativa para Algoritmos Evolutivos (EAs) foram introduzidas, resultando em um Algoritmo Evolutivo Adaptativo (AEA), cuja taxa média de convergência  $R_t$  (Chen et al. (2021)) foi analisada e mostrada ser linear para funções Lipschitz contínuas que satisfazem a condição  $(\mathcal{A}_1)$  ou  $(\mathcal{A}_2)$  (descrita na Seção 5.3). Ainda, foram obtidos limites inferiores para  $R_t$  para cada mutação proposta (Teorema 5.3, Corolário 5.4 e Teorema 5.5). A estimativa depende da dimensão do problema, da constante de Lipschitz e da condição relacionada ao comportamento da função objetivo na vizinhança do ótimo global (condições  $(\mathcal{A}_1)$  e  $(\mathcal{A}_2)$ ). Uma adaptação de tais teoremas foi feita para que o AEA aplicado na resolução do PDEPV ainda tivesse uma  $R_t$  linear. Como o AGD e AEA apresentaram bons resultados, foi proposto o AGDA, que incorpora a mutação do AEA à busca direcionada do AGD.

Por fim, devido a necessidade de modelar tanto as perdas de transmissão quanto os limites de fluxo de potência na rede por meio de algoritmos eficientes, que possam ser aplicados a todos os indivíduos da população, foi proposta uma nova representação da rede de transmissão no contexto do PDEPVZ-RR. Adaptações foram feitas ao AGD, para que este pudesse ser utilizado na resolução do PDEPVZ-RR. A principal delas foi a criação do operador que calcula as perdas e realiza o balanço de potências.

A fim de validar os métodos propostos, foi implementado o AGD em Python para resolver PDEPV e o PDEPVZ para sistemas de potência com 3, 5, 6 e 40 geradores. Além disso, o AGD também foi implementado na resolução do PDEPVZ-RR para o sistema IEEE-118 barras e PDEPV-RR para o sistema IEEE-300 barras. Foi possível notar que, comparado ao AG e aos outros métodos escolhidos para comparações, os métodos propostos tiveram seu desempenho melhorado, mas o melhor resultado foi obtido da junção das características de ambos (AGDA).

Uma área promissora para estudos posteriores é a incorporação de múltiplos combustíveis, tornando o problema mais representativo. No entanto, para manter a validade da prova de convergência diante dessa inclusão, seria necessário tratar a descontinuidade causada na função objetivo. Com relação aos resultados envolvendo a TMC, uma pesquisa futura poderia ser desenvolvida buscando uma melhor estimativa para o limitante inferior da TMC e propor uma mutação positiva-adaptativa que não dependa do ótimo global do problema e que ainda seja possível mostrar que resulta em uma TMC linear.

## Artigos completos publicados em congressos:

BORGES, L.T.M.; MATSUMOTO, I. H.; SOLER, E. M. *Otimização por Colônia de Formigas para Resolver o Problema do Caixeiro Viajante*. 12<sup>o</sup> Concistec - Congresso Científico da Semana Nacional de Ciência e Tecnologia do IFSP, 2022.

BORGES, L.T.M.; PAVAN, R.; SOLER, E. M. *Método de Otimização por Colônia de Formigas aplicado a Problemas não Lineares Multimodais*. XXIX Simpep - Simpósio de Engenharia de Produção, 2022.

BORGES, L.T.M.; PAVAN, R.; SOLER, E. M. *Ant Colony Optimization Method with Directed Singularity Search for Solving the Economic Dispatch Problem*. CLAGTEE 2022, XIV LATIN-AMERICAN CONGRESS ON ELECTRICITY GENERATION AND TRANSMISSION, Rio de Janeiro, Brazil.

## Artigos submetidos em periódicos:

BORGES, L.T.M.; SILVA, F.B; SOLER, E.M.; NEPOMUCENO, L. A Directed Genetic Algorithm for Solving Economic Dispatch and Network-Constrained Economic Dispatch Problems: Convergence Proof.

SILVA, F.B; BORGES, L.T.M.; NEPOMUCENO, L.; SOLER, E.M. Adaptive evolutionary algorithms with linear average convergence rate for optimization problems with Lipschitz continuous functions.

## Referências Bibliográficas

- AKIMOTO, Y., AUGER, A., & HANSEN, N. (2020). Quality gain analysis of the weighted recombination evolution strategy on general convex quadratic functions. *Theoretical Computer Science*, 832, 42–67.
- ALLEN, L. J. (2010). *An introduction to stochastic processes with applications to biology*. CRC press.
- ALTINOZ, O. T. (2019). The distributed many-objective economic/emission load dispatch benchmark problem. *Swarm and Evolutionary Computation*, 49, 102–113.
- ARENALES, M., ARMENTANO, V., & OTHERS (2006). *Pesquisa operacional*. Elsevier Brasil.
- 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, 53–61.
- AUGER, A. (2005). Convergence results for the  $(1, k)$ -sa-es using the theory of irreducible markov chains. *Theoret. Comput. Sci.*, 334, 35–69.
- AUGER, A. & HANSEN, N. (2016). Linear convergence of comparison-based step-size adaptive randomized search via stability of markov chains. *SIAM J. Optim.*, 26, 1589–1624.
- 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.

- BARRY, J. (2008). *Probabilidade: um curso em nível intermediário*.
- BERGSTRA, J., BARDENET, R., BENGIO, Y., & KÉGL, B. (2011). Algorithms for hyperparameter optimization. *Advances in neural information processing systems*, 24.
- BERTSEKAS, D. P. (1997). Nonlinear programming. *Journal of the Operational Research Society*, 48(3), 334–334.
- BEYER, H. G. (2013). *The Theory of Evolution Strategies*. Springer Science & Business Media.
- BROWNLEE, J. (2011). *Clever algorithms: nature-inspired programming recipes*. Jason Brownlee.
- CABRAL, M. A. (2010). Introdução a teoria da medida e integral de Lebesgue. *Rio de Janeiro-RJ: Universidade Federal do Rio de Janeiro*.
- CHAVEZ, J. C. S., ZAMORA-MENDEZ, A., PATERNINA, M. R. A., HEREDIA, J. F. Y., & CARDENAS-JAVIER, R. (2019). A hybrid optimization framework for the non-convex economic dispatch problem via meta-heuristic algorithms. *Electric Power Systems Research*, 177, 105999.
- CHEN, Y. & HE, J. (2021). Average convergence rate of evolutionary algorithms in continuous optimization. *Information Sciences*, 562, 200–219.
- DE OLIVEIRA, K. W., NASCIMENTO JR, 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.
- DERRAC, J., GARCÍA, S., MOLINA, D., & HERRERA, F. (2011). A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm and Evolutionary Computation*, 1(1), 3–18.
- DHIVYAPRABHA, T., SUBASHINI, P., & KRISHNAVENI, M. (2018). Synergistic fibroblast optimization: a novel nature-inspired computing algorithm. *Front. Inf. Technol. Electron. Eng.*, 19, 815–833.
- DING, L. & KANG, L. (2001a). Convergence rates for a class of evolutionary algorithms with elitist strategy. *Acta Mathematica Scientia*, 21(4), 531–540.

- DING, L. & KANG, L. (2001b). Convergence rates for a class of evolutionary algorithms with elitist strategy. *Acta Math. Sci.*, 21, 531–540.
- DONG, H., LI, T., DING, R., & SUN, J. (2018). A novel hybrid genetic algorithm with granular information for feature selection and optimization. *Appl. Soft Comput.*, 65, 33–46.
- ELON, L. (2009). Curso de análise vol. 2. *Projeto Euclides IMPA*.
- ENGELBRECHT, A. P. (2007). *Computational intelligence: an introduction*. John Wiley & Sons.
- EXLER, O. & SCHITTKOWSKI, K. (2007). A trust region sqp algorithm for mixed-integer nonlinear programming. *Optimization Letters*, 1(3), 269–280.
- FAN, J., ZHANG, C., LIU, Q., SHEN, W., & GAO, L. (2022). An improved genetic algorithm for flexible job shop scheduling problem considering reconfigurable machine tools with limited auxiliary modules. *Journal of Manufacturing Systems*, 62, 650–667.
- GAO, S., YU, Y., WANG, Y., WANG, J., CHENG, J., & ZHOU, M. (2019). Chaotic local search-based differential evolution algorithms for optimization. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(6), 3954–3967.
- GASPAR-CUNHA, A., TAKAHASHI, R., & ANTUNES, C. H. (2012). *Manual de computação evolutiva e metaheurística*. Imprensa da Universidade de Coimbra/Coimbra University Press.
- GHASEMI, M., AGHAEI, J., AKBARI, E., GHAVIDEL, S., & LI, L. (2016). A differential evolution particle swarm optimizer for various types of multi-area economic dispatch problems. *Energy*, 107, 182–195.
- GHOLAMGHASEMI, M., AKBARI, E., ASADPOOR, M. B., & GHASEMI, M. (2019). A new solution to the non-convex economic load dispatch problems using phasor particle swarm optimization. *Applied Soft Computing*, 79, 111–124.
- GONÇALVES, E., BALBO, A. R., SILVA, D. N., NEPOMUCENO, L., BAPTISTA, E. C., & SOLER, E. M. (2019). Deterministic approach for solving multi-objective non-smooth environmental and economic dispatch problem. *International Journal of Electrical Power & Energy Systems*, 104, 880–897.

- GRAINGER, J. J. & STEVENSON JR, W. D. (1994). *Power System Analysis*. McGraw-Hill Science/Engineering.
- HAPP, H. (1977). Optimal power dispatch: A comprehensive survey. *IEEE Transactions on Power Apparatus and Systems*, 96(3), 841–854.
- HE, J. & KANG, L. (1999a). On the convergence rate of genetic algorithms. *Theoret. Comput. Sci.*, 229, 23–39.
- HE, J. & KANG, L. (1999b). On the convergence rates of genetic algorithms. *Theoretical Computer Science*, 229(1-2), 23–39.
- HE, J. & LIN, G. (2015). Average convergence rate of evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 20(2), 316–321.
- HE, J. & YU, X. (2001a). Conditions for the convergence of evolutionary algorithms. *J. Syst. Arch.*, 47, 601–612.
- HE, J. & YU, X. (2001b). Conditions for the convergence of evolutionary algorithms. *Journal of systems architecture*, 47(7), 601–612.
- HOSSEINNEZHAD, V. & BABAEI, E. (2013). Economic load dispatch using  $\theta$ -pso. *International journal of electrical power & energy systems*, 49, 160–169.
- HOTA, P. K. & SAHU, N. C. (2015). Non-convex economic dispatch with prohibited operating zones through gravitational search algorithm. *International Journal of Electrical and Computer Engineering*, 5(6).
- HUSSAIN, S., AL-HITMI, M., KHALIQ, S., HUSSAIN, A., & ASGHAR SAQIB, M. (2019). Implementation and comparison of particle swarm optimization and genetic algorithm techniques in combined economic emission dispatch of an independent power plant. *Energies*, 12(11), 2037.
- JÄGERSKÜPPER, J. (2006). How the (1+ 1) es using isotropic mutations minimizes positive definite quadratic forms. *Theoretical Computer Science*, 361(1), 38–56.
- JÄGERSKÜPPER, J. (2007). Algorithmic analysis of a basic evolutionary algorithm for continuous optimization. *Theoretical Computer Science*, 379(3), 329–347.

- JANIGA, D., CZARNOTA, R., STOPA, J., & WOJNAROWSKI, P. (2019). Self-adapt reservoir clusterization method to enhance robustness of well placement optimization. *J. Petrol. Sci. Eng.*, 173, 37–52.
- LI, T., DONG, H., & SUN, J. (2019). Binary differential evolution based on individual entropy for feature subset optimization. *IEEE Access*, 7, 24109–24121.
- LI, X. & GAO, L. (2016). An effective hybrid genetic algorithm and tabu search for flexible job shop scheduling problem. *International Journal of Production Economics*, 174, 93–110.
- LIU, T., JIAO, L., MA, W., MA, J., & SHANG, R. (2016). Cultural quantum-behaved particle swarm optimization for environmental/economic dispatch. *Applied Soft Computing*, 48, 597–611.
- MANDAL, B. & ROY, P. K. (2021). Dynamic economic dispatch problem in hybrid wind based power systems using oppositional based chaotic grasshopper optimization algorithm. *Journal of Renewable and Sustainable Energy*, 13(1), 013306.
- MELIÁN, B., PÉREZ, J. A. M., & VEGA, J. M. M. (2003). Metaheurísticas: Una visión global. *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial*, 7(19), 0.
- MOHAMMADIAN, M., LORESTANI, A., & ARDEHALI, M. M. (2018). Optimization of single and multi-areas economic dispatch problems based on evolutionary particle swarm optimization algorithm. *Energy*, 161, 710–724.
- MOLINA-PÉREZ, D., MEZURA-MONTES, E., PORTILLA-FLORES, E. A., VEGA-ALVARADO, E., & CALVA-YAÑEZ, B. (2024). A differential evolution algorithm for solving mixed-integer nonlinear programming problems. *Swarm and Evolutionary Computation*, 84, 101427.
- MORINAGA, D., FUKUCHI, K., SAKUMA, J., & AKIMOTO, Y. (2023). Convergence rate of the (1+1)-es on locally strongly convex and lipschitz smooth functions. *IEEE Transactions on Evolutionary Computation*, 28(2), 501–515.
- NESTEROV, Y. (2015). Universal gradient methods for convex optimization problems. *Mathematical Programming*, 152(1), 381–404.

- NETO, J. X. V., REYNOSO-MEZA, G., RUPPEL, T. H., MARIANI, V. C., & COELHO, L. S. (2017). Solving non-smooth economic dispatch by a new combination of continuous grasp algorithm and differential evolution. *International Journal of Electrical Power & Energy Systems*, 84, 13–24.
- OLIVEIRA, A. S. L. D., RIBEIRO, T. S. G., & DA, S. F. B. (2017). Cadeia de markov: modelo probabilístico e convergência das distribuições de probabilidade. *C.Q.D. – Revista Eletrônica Paulista de Matemática*, 11, 49–61.
- PANG, A., LIANG, H., LIN, C., & YAO, L. (2023). A surrogate-assisted adaptive bat algorithm for large-scale economic dispatch. *Energies*.
- PINHEIRO, R. B., BALBO, A. R., CABANA, T. G., & NEPOMUCENO, L. (2022). Solving nonsmooth and discontinuous optimal power flow problems via interior-point lp-penalty approach. *Computers & Operations Research*, 138, 105607.
- 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.
- POURAKBARI-KASMAEI, M., LEHTONEN, M., FOTUHI-FIRUZABAD, M., MARZBAND, M., & MANTOVANI, J. R. S. (2019). Optimal power flow problem considering multiple-fuel options and disjoint operating zones: A solver-friendly minlp model. *International Journal of Electrical Power & Energy Systems*, 113, 45–55.
- POURAKBARI-KASMAEI, M. & MANTOVANI, J. R. S. (2018). Logically constrained optimal power flow: Solver-based mixed-integer nonlinear programming model. *International Journal of Electrical Power & Energy Systems*, 97, 240–249.
- 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.
- RUDOLPH, G. (1997a). Convergence rates of evolutionary algorithms for a class of convex objective functions. *Control and Cybernetics*, 26, 375–390.
- RUDOLPH, G. (1997b). Local convergence rates of simple evolutionary algorithms with cauchy mutations. *IEEE Trans. Evol. Comput.*, 1(4), 249–258.

- RUFFINO, P. R. C. (2012). *Uma iniciação aos sistemas dinâmicos estocásticos*. Publicações Matemáticas-IMPA.
- SA-NGIAMVIBOOL, W., POTHIIYA, 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.
- SECUI, D. C. (2015). A method based on the ant colony optimization algorithm for dynamic economic dispatch with valve-point effects. *International Transactions on Electrical Energy Systems*, 25(2), 262–287.
- SEMENOV, M. & TERKEL, D. (2003). Analysis of convergence of an evolutionary algorithm with self-adaptation using a stochastic lyapunov function. *Evol. Comput.*, 11, 363–379.
- SINGH, D. & DHILLON, J. S. (2019). Ameliorated grey wolf optimization for economic load dispatch problem. *Energy*, 169, 398–419.
- SONG, H. B. & LIN, J. (2021). A genetic programming hyper-heuristic for the distributed assembly permutation flow-shop scheduling problem with sequence dependent setup times. *Swarm and Evolutionary Computation*, 60, 100807.
- SÖRENSEN, K. & GLOVER, F. (2013). Metaheuristics. *Encyclopedia of operations research and management science*, 62, 960–970.
- SRIVASTAVA, A. & DAS, D. K. (2020). A new kho-kho optimization algorithm: An application to solve combined emission economic dispatch and combined heat and power economic dispatch problem. *Engineering Applications of Artificial Intelligence*, 94, 103763.
- STEINBERG, M. J., SMITH, T. H., & OTHERS (1943). Economy loading of power plants and electric systems.
- SUKKERD, W. & WUTTIPORNPUN, T. (2016). Hybrid genetic algorithm and tabu search for finite capacity material requirement planning system in flexible flow shop with assembly operations. *Computers & Industrial Engineering*, 97, 157–169.
- SUN, J., DENG, J., & LI, Y. (2020). Indicator & crowding distance-based evolutionary algorithm for combined heat and power economic emission dispatch. *Applied Soft Computing*, 90, 106158.

- TARLOWSKI, D. (2023). On asymptotic convergence rate of random search. *Journal of Global Optimization*, (pp. 1–31).
- 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.
- WANG, L. & PENG, Z. P. (2020). Solving energy-efficient distributed job shop scheduling via multi-objective evolutionary algorithm with decomposition. *Swarm and Evolutionary Computation*, 58, 100745.
- WANG, W., TIAN, G., ZHANG, H., LI, Z., & ZHANG, L. (2023). A hybrid genetic algorithm with multiple decoding methods for energy-aware remanufacturing system scheduling problem. *Robotics and Computer-Integrated Manufacturing*, 81, 102509.
- WANG, Y., JIAO, Y.-C., & LI, H. (2005). An evolutionary algorithm for solving nonlinear bilevel programming based on a new constraint-handling scheme. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 35(2), 221–232.
- WOOD, A. J., WOLLENBERG, B. F., & SHEBLÉ, G. B. (2013). *Power generation, operation, and control*. John Wiley & Sons.
- XIN-GANG, Z., JI, L., JIN, M., & YING, Z. (2020). An improved quantum particle swarm optimization algorithm for environmental economic dispatch. *Expert Systems with Applications*, 152, 113370.
- YANG, Q., QIAO, Z. Y., XU, P., LIN, X., GAO, X. D., WANG, Z. J., & ZHANG, J. (2024). Triple competitive differential evolution for global numerical optimization. *Swarm and Evolutionary Computation*, 84, 101450.
- ZHAN, J., WU, Q., GUO, C., & ZHOU, X. (2014a). Economic dispatch with non-smooth objectives-part i: Local minimum analysis. *IEEE Transactions on Power Systems*, 30(2), 710–721.
- ZHAN, J., WU, Q., GUO, C., & ZHOU, X. (2014b). Economic dispatch with non-smooth objectives-part ii: Dimensional steepest decline method. *IEEE Transactions on Power Systems*, 30(2), 722–733.

ZHANG, S. X., HU, X. R., & ZHENG, S. Y. (2024). Differential evolution with evolutionary scale adaptation. *Swarm and Evolutionary Computation*, 85, 101481.

# Apêndice A

## Apêndice

**Proposição A.1.** Seja  $x_n$  uma sequência convergente com  $x_n \rightarrow L$ , quando  $n \rightarrow \infty$ . Então a sequência das somas parciais dos termos de  $(x_n)$  é convergente com limite  $L$ .

*Demonstração.* Suponha que  $\lim x_n = L$ , e fixemos  $\epsilon > 0$ . Pela definição de limite, podemos encontrar um  $N_\epsilon$  tal que

$$k \geq N_\epsilon \implies |x_k - L| < \epsilon/2.$$

Se considerarmos a sequência

$$y_n = \frac{x_1 + x_2 + \cdots + x_n}{n},$$

de modo que  $y_n$  representa a média dos primeiros  $n$  termos da nossa sequência original  $x_n$ , então a prova do nosso teorema se reduz a mostrar que  $y_n \rightarrow L$ . Queremos utilizar a nossa definição de limite, então começamos analisando o termo de erro  $|y_n - L|$ .

$$\begin{aligned} |y_n - L| &= \left| \frac{x_1 + x_2 + \cdots + x_n}{n} - L \right| \\ &= \left| \frac{x_1 + x_2 + \cdots + x_n}{n} - \frac{nL}{n} \right| \\ &= \left| \frac{x_1 + x_2 + \cdots + x_n - nL}{n} \right| \\ &= \left| \frac{(x_1 - L) + (x_2 - L) + \cdots + (x_n - L)}{n} \right|. \end{aligned}$$

Agora, podemos aplicar a desigualdade triangular para obter (assumindo que  $n$  é grande o suficiente para ultrapassar  $N_\epsilon$ ):

$$\begin{aligned} |y_n - L| &\leq \frac{|x_1 - L|}{n} + \frac{|x_2 - L|}{n} + \dots + \frac{|x_{N_\epsilon-1} - L|}{n} + \frac{|x_{N_\epsilon} - L|}{n} + \dots + \frac{|x_n - L|}{n} \\ &= \sum_{k=1}^{N_\epsilon-1} \frac{|x_k - L|}{n} + \sum_{k=N_\epsilon}^n \frac{|x_k - L|}{n}. \end{aligned}$$

Como  $|x_k - L| < \epsilon/2$  para  $k \geq N_\epsilon$ , temos:

$$\begin{aligned} |y_n - L| &< \sum_{k=1}^{N_\epsilon-1} \frac{|x_k - L|}{n} + \sum_{k=N_\epsilon}^n \frac{\epsilon}{2n} \\ &= \frac{1}{n} \sum_{k=1}^{N_\epsilon-1} |x_k - L| + \frac{n - N_\epsilon + 1}{2n} \epsilon. \end{aligned}$$

Note que a soma dos erros para os primeiros termos da sequência (primeiros  $N_\epsilon - 1$  termos) é finita, então tomando o limite conforme  $n \rightarrow \infty$ , essa parte tende a zero, ou seja, menor que  $\epsilon/2$ . Para o segundo termo, note que

$$\lim_{n \rightarrow \infty} \frac{n - N_\epsilon + 1}{n} = 1.$$

Assim, para  $n$  suficientemente grande, podemos afirmar que

$$|y_n - L| < \epsilon.$$

Em outras palavras,  $y_n$  satisfaz a definição de limite para  $\lim y_n = L$ , como desejado.  $\square$